Andreas Krause

Theoretical Foundations of Commercial Banking

Andreas Krause

Theoretical Foundations of Commercial Banking

Disclaimer The author does not make any guarantee or other promise as to any results that may be obtained from using the content of this book. You should never make any investment, borrowing, employment or any other decision without first consulting with your own financial and/or legal advisor and conducting your own research and due diligence. To the maximum extent permitted by law, the author disclaim any and all liability in the event any information, commentary, analysis, opinions, advice and/or recommendations contained in this book prove to be inaccurate, incomplete or unreliable, or result in any investment or other losses.

Andreas Krause
Department of Economics
University of Bath
Claverton Down
Bath BA2 7AY
United Kingdom
E-mail: mnsak@bath.ac.uk

Preface

The literature on the theory of banking has become quite extensive in recent times, not least in response to the financial crisis 2008/2009. In response to this crisis many models addressing the contagion of bank failures and liquidity shortages in the banking system have been developed, including regulatory responses. Overall, the literature is dominated by the asymmetric information between borrowers and banks, where borrowers have better information about their own prospects than banks, but also between banks and depositors or between different banks. In addition, moral hazard in that a borrower (bank) chooses an investment (loan) that is too risky to be optimal for the bank (depositor) as also considered alongside or instead of asymmetric information. Many models then address the implications of these market imperfections and how banks have responded to such challenges. Such models provide insights into the behaviour of banks and show the complexity of banking decisions. The majority of these models is concerned with commercial banks, i. e. banks that take deposits and lend these out, while the theoretical literature on investment banking, that facilitates of capital market transactions, is much more limited.

This plethora of theoretical models is accompanied by an ever larger number of empirical investigations, covering similar problems, but in many instances also going beyond the scope of models. While empirical investigations are often easily accessible, this is much less the case for theoretical models. Not only are the mathematical requirements often substantial, but access to these models is hampered by differences in the modelling approach, making relevant similarities or differences much more difficult to identify. Furthermore, it also makes combining different models for a more comprehensive analysis of bank behaviour more challenging. Using different notations further aggravates this problem.

Philosophy of this book

The aim of this book on the theory of banking is to overcome some of these identified shortcomings. The main features are

Comprehensive coverage I cover the full breadth of the theory of banking at considerable depth. Not only are the standard theories of banking covered in more depth than in other books, we also cover topics that are commonly not covered at all or given a very rudimentary treatment. Examples include the competition with non-bank entities, the hiring, remuneration, and promotion of employees.

Consistent modelling In the literature, models differ substantially in their assumptions. This might affect the number of time periods considered, the possible outcomes might be continuous or discrete, outcomes might differ by the probability of success or the return if successful, amongst many others. Here we use the same framework as much as possible for all models we discuss. This allows us to compare the results of these models and even combine different models to get more in-depth insights into bank behaviour. This made it necessary to rewrite many of the existing models in the literature, such that they often only resemble the initial idea intended by its authors. In other cases, relevant models were not able to be translated into the common framework and for that reason excluded. During the process of using a common modelling framework, we also ensured that the notation is as consistent as reasonably possible across models.

Detailed derivation Many books only provide the idea behind models and sketches of proofs before discussing some of their implications. This often leaves readers unable to fully understand the models without referring back to the original publications. All models discussed here are derived step-by-step with all assumptions clearly stated to allow the reader to fully understand the models. Some elements of proofs are omitted, though, as they are often trivial or on the other hand very lengthy without adding to the understanding of the model and its implications. Commonly second order conditions are not considered explicitly as they do not aid the understanding of the model or its implications. Similarly, we frequently do not consider corner solutions by making implicitly assumptions such that these can be excluded. Each model is presented in a way that it can be analysed in isolation of any other model, thus there are no prerequisites for any models in the form of having had to have acquired knowledge of any other model.

Practical problem sets Many books include exercise sets, most of which ask readers, mostly students, to solve variations or extensions of models that have been discussed. In addition, there might be some questions testing the understanding of specific models. The approach taken here is different, readers are exposed to a problem a bank, regulator, or observer faces and is supposed to use the models discussed to offer a solution or explanation. In some instances several models need to be combined to provide a comprehensive answer to the problem, and additional information needs to be extracted from the problem provided. This allows for a more realistic evaluation of actual problems in banking and trains the reader to look beyond the confines of the models by understanding their implications and context of banking decisions.

Prerequisites

Generally, the models used in banking are not very difficult and in most cases knowledge of the principles of microeconomic theory are sufficient. Some more advanced concepts such as game theory or mechanism design are used, but commonly at a level that allows sufficient understanding even without specialist knowledge.

All steps required to understand a model are provided in the text and derivations shown in detail; where necessary this is complemented by additional background material provided in the appendix to aid the understanding of economic theory or mathematical techniques. In general, anyone having acquired the knowledge of a thorough module in microeconomic theory is well equipped to follows this book.

Structure

After having looked at the benefits that banks can bring to an economy, we will explore the lending contract between a bank and its borrowers. We look not only onto the optimal contract specification, but will also analyse the incentives of borrowers to repay loans, the provision of collateral, covenants, the sharing of information about borrowers between banks, and the relationship between these borrowers and their bank. In addition, we will also look at reasons why some borrowers may fail to obtain loans, even if meeting all lending criteria. Looking at deposits, we will investigate situations where deposits get suddenly withdrawn without any discernable reason, how lending between banks can stabilise or destabilise the funding of banks, and what impact deposit insurance has on such arrangements. Other funding sources, such as repurchase agreements, are also considered, alongside payment services banks offer to their customers.

We then continues with the analysis of commercial banks, but focus more on the interaction between banks. We will look at competition between banks themselves as well as with non-bank financial institutions, but also at the spread of bank failures and how regulation affects banks' behaviour and subsequently their propensity to fail. Finally, the way banks treat their employees is considered, alongside the ethical considerations of bank behaviour. These aspects complement the analysis of lending and taking deposits in that rather than focussing on these primary activities of commercial banks directly, the emphasis moves away from the day-to-day running of the banking business to looking at issues that affect decision-making of senior managers, such as the impact of competition or reactions to regulatory constraints, but also the conditions of employees.

The final goal should not only be to derive models of banking and see how they contribute to the overall practice in banking, but also to apply these models to solve problems as they emerge in the day-to-day running of banks, or to analyze a situation in which banks find themselves in, with the aim to guide banks or regulators on resolving these. To this end, I also present a wide range of problem sets that can be solved using the models discussed here.

Using this book

This book is aimed at researchers and students alike. Researchers will naturally seek those models and detailed aspects they are most interested in, while for students a more structured approach needs to be taken. How a teacher might approach this, will largely depend on the aims of the module they are teaching. If looking at an introductory module in banking, either at advanced undergraduate level or beginning graduate level, teachers would most likely select a small number of models across the entire range, while more specialised graduate module might want to explore a small number of topics in much more depth. This book allows for both of these approaches and given that all models are presented self-contained, models can be selected freely as the teacher sees fit. Having acquired some knowledge of the financial system prior to using this book is desirable and will allow the reader to appreciate the importance of the issues discussed here more, but this is not essential.

Contents

List of Figures

Chapter 1

Prologue: Taking deposits and lending

Types of banks

What is loosely referred to as ’banks’ consists of two different types of businesses, commercial banks and investment banks. These two types of banks are quite different and have a very limited overlap in their activities.

Commercial banks are businesses whose main activities involve accepting monies from the general public and lending monies to individuals, companies, and public bodies. The monies accepted are in most cases repayable on demand and are commonly called deposits. Therefore, businesses that finance themselves mainly through the issue of bonds are not classified as commercial banks. While bonds can be traded and their holders that way obtain their monies, they are only repayable by the issuer at maturity and not at any time the bondholder demands.

This definition of a commercial bank excludes a range of institutions that are calling themselves ’bank’ from being classified as a commercial bank. Firstly, central banks generally do neither take monies from the general public, but only from commercial banks and public bodies, and they also only lend to the same group of customers. Secondly, development banks, such as the World Bank and many regional development banks, do accept monies from the general public, but only through issuing bonds, not by taking deposits. Their lending, though, can be to either public bodies only, or they provide loans to companies and even individuals directly, depending on their remit. On the other hand, our definition of commercial banks includes state-owned or publicly owned banks, and institutions not calling themselves banks, such as credit unions, mutual societies, and friendly societies. The ownership structure of commercial banks is irrelevant for their classification, as is their legal form, and we include limited companies, partnerships, and sole traders.

The exact legal definition of a commercial bank is in most jurisdictions much more complex than the definition provided here. These more detailed definitions have the goal to prevent commercial banks from circumventing strict regulations in some of their activities by claiming that these fall outside of the scope of commercial banking and are therefore not subject to these regulations. It is common in the legal context to call commercial banks simply 'banks', a convention that we will follow here unless we need to clarify the type of bank that is referred to. This identification of banks with commercial banks is also in line with the interpretation of the general public.

Investment banks, on the other hand, facilitate capital market transactions. This facilitation can take many forms, such as giving advice on investment decisions in capital markets for individuals or companies, acting as brokers to bring orders to buy or sell securities to the market, acting as market maker to trade on their own account to facilitate a transaction between two market participants, advising on buying and selling companies (mergers & acquisitions), and advising on and underwriting of the issuance of securities (bonds or shares). The final two business lines are seen as the main activities of investment banks. Legal definitions of investment banks are less consistent across jurisdictions as the regulation of investment banks has traditionally focussed more on the regulation of specific activities and their relationship to each other, rather than the regulation of the investment bank as a whole. A key difference to commercial banks is that investment banks do not accept deposits from the general public nor is their main business the provision of loans, even though they might occasionally provide loans to customers as part of capital market transactions, for example securities lending, bridge loans when advising on mergers, and similar occasions. Investment banks may, but rarely do, issue bonds.

In most countries, investment banking and commercial banking activities are conducted within the same legal entity, commonly referred to as 'universal banking'. Combining these two activities allows universal banks to provide their customers with the full range of banking services and advice, from holding their deposits, providing loans to advising on raising funds in capital markets or merging with other companies. However, operationally, these activities are usually distinct by being located in different departments and movement of staff as well as the exchange of information between these departments is unusual.

Therefore, while universal banking is common, we can clearly distinguish between commercial and investment banking activities. Investment banking activities are not considered here, where we exclusively focus on commercial banking activities.

Modelling the banking business

In order to understand the way banks conduct their business, it will be necessary to make many simplifying assumptions on a range of aspects in banking; this may range from simplifying the aspect under investigation itself, the considerations of banks and other market participants in decision-making, to the environment in which such decisions are made. It will be common to focus on a single aspect of banking activities only, and ignoring other, often related aspects. As we will see, the banking business is very complex and we will explore a wide range of facets that cover the range of problems a bank may face. In this context it is important to develop a common framework that allows us to capture the essence of the banking business as this then allows us to compare results and even combine different models to obtain a more holistic view of banking, integrating different and often contradictory results.

In this prologue, we will provide the framework which will be used for commercial banks throughout this book. Even though we will vary some assumptions as needed in the context of the problem on hand, it nevertheless provides an anchor point that allows us to have a consistent approach when addressing the different challenges that banks face. Firstly, we will look at the composition of the balance sheet of a bank as this allows us to identify the main drivers of bank profits, before we then look at the profits of banks, their borrowers and depositors. These profits will mainly drive the decision-making of banks as well as their customers and are therefore of central importance in the analysis of bank behaviour.

Bank balance sheets The key elements of a bank balance sheet are the loans given to companies, private individuals, and public bodies, denoted by L. For simplicity, we will refer to any borrowers as such or as 'companies', but will not exclude the possibility of loans been given to private individuals or public bodies. These loans are typically financed by the raising of deposits D from the general public and we refer to these as depositors and it is often implicitly assumed that they are private individuals, but the models do not require this to be the case and they might well be companies depositing excess funds. Banks typically do not invest all their deposits into loans, but retain same fraction as cash reserves, C, to cover any deposit withdrawals. They might also hold securities, often government bonds, G, rather than cash, in order to obtain some returns from their investment while being able to generate cash at very short notice without making losses. In most settings we will neglect the holding of securities and instead interpret them as cash reserves.

In addition to dealing with the general public as borrowers and depositors, banks may also borrow and lend to other banks in so called interbank markets. The lending to another bank, B, is another use of the funds available to banks. Similarly, banks may want to complement their funding from deposits by borrowing from other banks, $\hat{B}$, allowing them to either invest more into cash reserves, securities or loans. Unless we are concerned with interbank markets, we will neglect this position. Industrial companies also obtain loans from other companies, and give loans, but in contrast to banks, this is usually based on having established a relationship between the two companies through their supply chains for goods or services, with one of the companies being the supplier and the other their customer. The loan can take the form of a customer paying a deposit to their supplier who then supplies the goods or services at a later stage, or a company provides the goods or services, but allows their customer to make payment for these at a later stage. For interbank loans no such relationships exists, the provision of loans is independent of any other business

Assets		Liabilities	
Cash	C	Deposits	D
Securities	G	Central bank loans	M
Interbank lending	B	Interbank borrowing	$\hat{B}$
Loans	L	Equity	K

Fig. 1.1: Bank balance sheet

relationship two banks might have, and in most cases there is not further relationship between banks beyond interbank lending.

Other assets that banks might hold, such as property or long-term investments, are always ignored. All these positions are typically small compared to the amount of loans banks provide and will thus make no material difference to results if excluded. We generally only include assets beyond loans and liabilities beyond deposits if they are important for the outcome of the model, or if they are the focus of the model and the activity on hand.

Banks finance their loans not only by deposits and interbank borrowing, but may also obtain loans from the central bank, M. Again these usually small positions, when compared to deposits, are ignored unless they are the focus of the investigation. The final source of finance by banks is equity, K. As banks normally have very little equity relative to deposits, we again ignore this position in many models, unless equity is a relevant variable to understand the behaviour of banks.

Figure 1.1 shows the balance sheet thus discussed with all its components. As indicated, we will in nearly all cases neglect interbank lending and borrowing, the ownership of government securities as well as the loans obtained from central banks. Therefore we commonly assume that $B = \hat{B} = G = M = 0$. With the obvious exemption of discussing capital regulation or the impact of equity on bank decisions, we will frequently set $K = 0$ as a simplification as well. If the presence of cash reserves is not relevant for the question the model seeks to address, we will neglect these as well by setting $C = 0$. In this case, we have $L = D$, a relationship we will find in many of the models discussed. In the more general case when including cash reserves and equity it would be $L + C = D + K$.

Having clarified the structure of a banks' balance sheet, we can now continue to develop the context in which loans are given and how this generates profits to borrowers, depositors, and banks.

Profit functions We will generally assume that companies (borrowers) use the loan to make an investment of size I. This investment will either succeed with probability π and yield a return of R or it will fail with probability $1 - \pi$ and in this case yield nothing. Companies have only this single investment and due to limited liability will only be able to repay the loan if their investment is successful. If the investment is not successful, the company does not receive any funds, but also does not have to repay the loan. If the loan rate is given by r_L, the companies' expected profits at maturity of the loan are then given as

$$\begin{aligned}\Pi_C &= \pi\left((1+R)\,I - (1+r_L)\,L\right) \\ &= \pi\left((1+R) - (1+r_L)\right)L.\end{aligned} \tag{1.1}$$

It will be in most cases that we assume that companies have no own funds and thus their investment is entirely financed by the loan. If they also do not hold back any monies for other uses, we will have $I = L$, which gives rise to the second line in equation (1.1). Commonly we will use a single time period in which investments are completed and the outcome is known.

As loans are only repaid if investments by companies are successful, they are repaid with probability π, the same as the success rate of the company investments. With companies receiving no funds if their investments fail, banks will in this case receive no payments either. Banks finance their loans using deposits D, on which they have to pay interest r_D. We commonly assume in our models that deposits are repayable at maturity of the loan, i. e. at the end of one time period. To assess the profits of banks, we will make one of two assumptions on the banks' liability. The first possible assumption is that banks have limited liability and no other funds available than the repayment of the loan. In this case, the bank can repay their deposits only if the loan has been repaid, thus the expected returns of banks are given by

$$\Pi_B = \pi\left((1+r_L)\,L - (1+r_D)\,D\right). \tag{1.2}$$

An alternative assumption is that either banks have unlimited liability or other resources to repay depositors and hence will do so, regardless of the loan repayment. In this case the expected profits of the banks are given by

$$\Pi_B = \pi\,(1+r_L)\,L - (1+r_D)\,D. \tag{1.3}$$

This case might be relevant if we investigate the impact of a single loan which is part of a large loan portfolio and the default of this loan will not affect the bank's ability to repay deposits, while the former case would look at the entire loan portfolio. It will be common to have $D = L$ as we neglect equity as well as cash holdings and interbank loans.

For depositors we obtain that they are either repaid in all cases or repaid only if the bank has been repaid the loan, which happens with probability π. If depositors have the possibility to invest into other assets yielding a return r, such as government securities, their expected surplus from using deposits are given by

$$\Pi_D = \pi\,(1+r_D)\,D - (1+r)\,D, \tag{1.4}$$

$$\Pi_D = (1+r_D)\,D - (1+r)\,D, \tag{1.5}$$

for the case of limited and unlimited bank liability, respectively. We will often assume that $r = 0$ for simplicity or that no alternative to deposits is available, apart from holding cash on which no interest is payable.

In most models we will assume that all market participants are solely concerned about their expected profits and seek to either break even in a competitive market, requiring that $\Pi_i \geq 0$, or maximize their expected profits. Implied with that assumption is that market participants are risk neutral.

These base models will be adjusted to suit the needs such that the problem in question is addressed adequately. Therefore, the baseline model presented here serves as a benchmark and starting point that will be modified to allow us to capture the problem we seek to address.

Key challenges for banks and depositors

Banks provide loans to companies, who then seek to invest these monies. However, once banks have provided the loan, they cannot direct the company to make the investment they have committed to, unless mechanisms are in place to provide incentives for companies to do so or other enforcement actions are possible. The same is the case for depositors. Once they have provided the funding (deposits), the bank can use these funds to grant loans as they see fit. Similarly, companies might not provide truthful information to the bank about the investments they seek to make, as much as banks might not be truthful to depositors about their intentions on the types of loans that they will grant. Again, legal constraints might be used to require truthful disclosure, but incentives to be truthful would avoid the problem of enforcing such regulation. Many models will discuss the consequences of these problems that banks and borrowers have.

Here we will briefly discuss the two main manifestations of the resulting problem, namely asymmetric information between companies (banks) and banks (depositors) as well as the moral hazard in the behaviour of companies and banks.

Adverse selection When lending, banks are often in a situation where the borrower is better informed about the prospects of their investment than the bank. This informational asymmetry can be exploited by the borrower. Akerlof (1970) provided a simple model of this adverse selection problem, the 'lemon' problem, which we will use here in the context of lending.

Let us assume there are two types of companies that the bank cannot distinguish, but the companies know their own type. One has a probability of success of their investments of π_H and the other of $\pi_L < \pi_H$. The bank knows that companies of type H are a fraction p of the market. Such companies are called 'high-quality companies', while companies of type L are called 'low quality'/ The bank profits are then given by

$$\begin{aligned}\Pi_B &= p\left(\pi_H\left(1+r_L\right)L-\left(1+r_D\right)D\right)\\ &\quad+\left(1-p\right)\left(\pi_L\left(1+r_L\right)L-\left(1+r_D\right)D\right)\\ &=\left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)L-\left(1+r_D\right)D.\end{aligned} \tag{1.6}$$

The first term represents the expected profits of the bank from lending to high-qualities companies, of which a fraction p populate the market, and the second term the expected profits of lending to low-quality companies, who have a share of $1-p$

in the market. As the bank cannot distinguish between the types of companies, it will have to charge the same loan rate r_L to both types of companies.

If banks are competitive, we find that $\Pi_B = 0$. With $L = D$ for simplicity, this allows us to obtain the loan rate as

$$1 + r_L = \frac{1 + r_D}{p\pi_H + (1 - p)\,\pi_L}. \tag{1.7}$$

For the company to demand a loan, we need that it is profitable to do so, hence

$$\Pi_C^i = \pi_i \left((1 + R_i)\, I - (1 + r_L)\, L\right) \geq 0, \tag{1.8}$$

where R_i denotes the return of a successful investment for a company of type i. We assume now that $\pi_H\,(1 + R_H) > 1 + r_D > \pi_L\,(1 + R_L)$, implying that the expected return of the high-quality company's investment is higher then that of the low-quality company. However, assuming that $R_H < R_L$, which can be interpreted that if the investment is successful, the return of the low-quality company L is higher. This corresponds to a situation where a higher risk, here a lower likelihood of succeeding, attracts a higher return. Furthermore, the investment of the high-quality company is desirable as it is, on average, able to cover its financing costs of the bank in form of deposits, while the investment of the low-quality company does not cover these costs.

With the assumption that companies fully rely on bank loans to finance their investment, $I = L$, and inserting from equation (1.7) for the loan rate, we can solve equation (1.8) for the high-quality company as requiring that

$$p \geq p^* = \frac{(1 + r_D) - \pi_L\,(1 + R_H)}{(\pi_H - \pi_L)\,(1 + R_H)}. \tag{1.9}$$

Hence if the fraction of high-quality companies is too small, there will be no demand for loans by these companies. The profits of low-quality companies are given by

$$\begin{aligned} \Pi_C^L &= \pi_L\left((1 + R_L) - (1 + r_L)\right) L \\ &= \frac{(p\pi_H + (1 - p)\,\pi_L)\,(1 + R_L) - (1 + r_D)}{p\pi_H + (1 - p)\,\pi_L} L, \end{aligned} \tag{1.10}$$

after inserting from equation (1.7) for the loan rate. If we now make the additional assumption that even if $p \leq p^*$, the parameters are such that $\Pi_C^L \geq 0$, low-quality companies will demand loans. As $R_L > R_H$ this is a feasible solution if $1 + R_L \geq 1 + r_L > 1 + R_H$. As high-quality companies will not demand any loans, banks will only be able to lend to low-quality companies. Hence, their profits will be $\Pi_B = \pi_L\,(1 + r_L)\, L - (1 + r_D)\, D$ and even if we set $r_L = R_L$ and extract any surplus from the low-quality company, our assumption that $1 + r_D > \pi_L\,(1 + r_L)$, makes banks unprofitable and they would cease to lend. The market has collapsed.

Hence in the presence of adverse selection, the existence of too many low-quality companies that cannot be identified by the bank, has crowded out the desirable loans to high-quality companies and leads to the collapse of the loan market. The reason is that as low-quality companies become numerous, the loan rate has to increase to compensate the bank for the lower sucess rate of the more common low-quality companies, reducing the profits of the high-quality companies. Once the loan rate has increased sufficiently, the high-quality company is not profitable anymore and will cease to demand loans from the bank. Banks have developed mechanisms to be able to distinguish the different types of companies through the loan contract. By providing specific loan terms such that low-quality companies cannot profitably pretend to be a high-quality company, banks can continue to provide loans in such circumstances, as we will see in future models.

An identical problem can also be constructed where the bank is replaced by the depositor and the company by the bank. In this case the bank is of unknown type to the depositor and the deposit market might collapse in the exactly same way as described above. Again, high-quality banks might find mechanisms to reveal the type of bank they are, such that depositors can fund them.

Moral hazard Another problem faced by banks is that of moral hazard as introduced by Arrow (1963). Borrowers, having obtained a loan, might make different investments from what the bank had anticipated. We will often assume that borrowers can choose between two investments with success probabilities π_H and $\pi_L < \pi_H$, respectively, yielding returns of $R_L > R_H$ if successful. Furthermore, we assume that $\pi_H (1 + R_H) > 1 + r_D > \pi_L (1 + R_L)$, meaning the expected profits of investment H exceeds its funding costs of the deposits used to finance the loan, and is thus viable, while investment L does not earn its costs. This setting is identical to that of adverse selection, but here the company does not have a specific type exogenously given, but can choose the type of investment they make. We therefore call investment H the high-quality investment and investment L the low-quality investment.

If the investment is fully financed by loans with a loan rate r_L, $I = L$ then the company profits for investment of type i are given by

$$\Pi_C^i = \pi_i ((1 + R_i) I - (1 + r_L) L) = \pi_i (R_i - r_L) L, \tag{1.11}$$

where the second equality arises from using $I = L$. For the company to choose the high-quality investment we need $\Pi_C^H \geq \Pi_C^L$, which solves for

$$1 + r_L \leq 1 + r_L^* = \frac{\pi_H (1 + R_H) - \pi_L (1 + R_L)}{\pi_H - \pi_L}, \tag{1.12}$$

implying that the loan rate must not be too high. This requirement limits the profitability of banks and they might seek mechanisms to ensure companies choose high-quality investments. Given that we assume that $1 + r_D > \pi_L (1 + R_L)$, lending to a company that will choose the low-quality investment can never be profitable for the bank, assuming unlimited liability. It will also only be profitable to lend to a high-quality company if $\Pi_B = \pi_H (1 + r_L) L - (1 + r_D) D \geq 0$. If we use the highest possible loan rate r_L^* and use our assumption that loans are fully financed by deposits, $D = L$, this requires that $1 + r_L \geq 1 + r_L^{**} = \frac{1+r_D}{\pi_H}$. Depending on parameter constellations, this condition might not be fulfilled at the same time as the constraint in equation (1.12). We can only find a loan rate that prevents companies choosing low-quality investments and banks being profitable, if we can find a loan rate that fulfills $1 + r^{**} \leq 1 + r_L \leq 1 + r_L*$, hence we require $1 + r^{**} \leq 1 + r_L^*$.

Thus, moral hazard may prevent banks from lending, even if it would be feasible based on the returns that high-quality investments generate. The reason is the incentive of companies to divert their loan to make low-quality investments, which in the corporate finance literature is referred to as 'risk shifting'. The origin of this term is the fact that low-quality investment are riskier, due to their lower probability of success, but higher return in case of success; this allows the company to retain higher profits if successful, but faces the same losses if the investment is not successful. This moral hazard can lead to the collapse of the loan market and banks will seek mechanisms that can prevent such a breakdown of lending. They might provide incentives to companies such that switching to low-quality investments is less desirable or even have contractual terms to prevent these decisions all together.

Depositors face a similar moral hazard problem with banks. Bank will have the same incentives to seek low-quality (riskier) loans over high-quality (less risky) loans and in order to attract deposits, they will have to establish a way to reduce this moral hazard problem.

Addressing adverse selection and moral hazard The assumptions made for the case of adverse selection and moral hazard were nearly identical. Their only difference is that in the case of adverse selection companies are of a certain type and do not make an active choice of their investments, while with moral hazard the company makes the investment decision. The aim of addressing adverse selection would be to exclude low-quality companies from demanding loans in the first place or allow banks other mechanism to identify them and offer them loan terms that suit their type. In contrast, in the case of moral hazard, companies make an active choice to make low-quality investments. Consequently, the aim of addressing moral hazard is to prevent companies from choosing such low-quality investment and choosing high-quality investments instead.

Faced with either adverse selection or moral hazard, most models will use loan conditions to make demanding loans for low-quality investments unprofitable or ensure that the desirable high-quality investment is more profitable. This will frequently be done by using constraints on the behaviour of banks, equivalent to those derived in equations (1.9) and (1.12).

Summary

In this prologue, we have established the basic setup of the banking models we will use in addressing a wide range of aspects of banking. Using such a common framework will allow us some insights into the decision-making by banks and their customers, companies as well as depositors. With the bank and depositors facing adverse selection and moral hazard, a rich variety of challenges are to be found within banks, who offer many, often conflicting solutions to these challenges as the coming chapters will explore. The common framework outlined here, will enable us to compare results across models and combine the results of different models to allow a more comprehensive analysis of bank behaviour.

Part I
The importance of banks

In many instances banks are seen as intermediaries between market participants that have excess funds, e. g. savers or investors, and those that have a shortage of funds, such as companies investing into production or consumers seeking to purchase items without having funds instantly available. The role of banks is then one in which such excess funds are matched to the demand of borrowers, thereby reducing transaction costs of both parties involved. Neglecting these transaction costs in chapter 2.1 we investigate what the impact of banks on the economy would be in a such a scenario. Even if banks have an advantage in monitoring borrowers, chapter 2.3 shows that this would have no direct impact on market outcomes with perfect competition. In addition, banks are seen as storers of value in that temporarily not needed funds can be given to the bank where it will be safe from theft and loss, unlike cash, gold, or other valuable items. Furthermore, banks often offer payment services that allow their customers to transfer funds to other customers of the same or another bank, replacing the sending of cash. The storing of value in an account at a bank as well as the provision of payment services reduces transaction costs to those using this service. While the view of banks as pure intermediaries to match funds and provide the additional services is generally accepted, banks are playing a much wider role in the economy.

As the following chapters will show, banks are having a much more profound impact than a mere intermediary would have. An important difference is that unlike a pure intermediary, banks do not only hold on to funds and pass them on between customers, whether for payments between customers or to match excess funds with borrowers, but instead take an active role in the process of lending. This active role goes beyond that of monitoring of borrowers, which in itself is an added value of banks and is discussed in chapter 3. It is also not limited to the benefits of borrowers forming a cooperative with the aim to reduce the cost of borrowing through a scheme of joint liability that is exercised through a bank as discussed in chapter 3.3. Unlike intermediaries for goods and services, such as retailers and wholesalers, banks do not buy the good (obtain a deposit from a customer) and then sells this good (uses the deposit to grant a loan), but allows the depositor to withdraw its funds, independent of the maturity of the loan. In this sense, banks are different from granting the loan directly as the lender (depositor) in this case cannot withdraw funds prior to the maturity of the loan. A bank may lend money long-term, despite possibility that the deposits that are used to fund this long-term loan, are being withdrawn at any time. This ability to withdraw funds at any time while providing long-term loans is seen as one of the key features of banks and discussed in chapter 4, but it is also one of the causes banks are fragile as we will discover in chapter 15. Banks are also able to provide liquidity to borrowers by allowing for credit lines or overdrafts, that borrowers only use if they require the liquidity. Banks are not only passive in that they provide loans to companies with given characteristics. In chapter 5 we will see how banks can induce companies to alter their characteristics, showing that banks can have a much more pronounced effect on an economy than an intermediary would.

It is important to understand the benefits banks provide an economy with to fully appreciate the relevance of their operations and regulation, which will be covered in the coming parts of this book. This part provides an overview of some of the key contributions banks make to the economy.

Chapter 2

Intermediation

A common view of banks is that they act as intermediaries between lenders (depositors) and borrowers. Using this approach, banks are seen as collecting funds from individuals and companies with excess funds, called deposits, and using these funds to provide loans to individuals and companies. Propagators of such an outlook are interpreting banks as organisations that pass the funds from one group (depositors) to another group (borrowers). In its simplest form, the bank does not add any value itself, but its value arises from providing a platform for these two groups to come together and match any offers from lenders to the needs of borrowers. Similar to market makers in securities markets, banks take a proprietary position by taking the funds of lenders on as deposits, similar to buying securities from investors, and providing loans on their own accounts, equivalent to selling securities to investors. Like the buying and selling of market makers in securities markets, banks seek to have a balance between these two activities such that at all times the amounts are approximately balanced. In contrast to market makers, though, once a transaction has been offset, the position does not vanish from the balance sheet of the bank, the loan and the deposits remain an asset and a liability, respectively, of the bank. Whereas market makers transfer ownership of the securities, banks retain ownership of both depositors and borrowers until the loan is repaid and the deposit is withdrawn.

In such a simplified view, banks are playing no active role in the loan market and, assuming transaction costs are not affected, they will not affect the outcome in the economy. Thus, from an economic perspective, banks are seen as a convenience for borrowers and lenders without any meaningful impact for the outcomes in an economy and their existence could be largely ignored.

Before moving to models that show how banks can improve the efficiency of loan markets and provide a higher social welfare, we will explore this intermediation role and more formally show that banks are irrelevant if seen as pure intermediaries. Firstly, in chapter 2.1 we will explore a setting in which markets are frictionless and banks have no inherent advantage compared to lenders and borrowers negotiating loans directly, before in chapter 2.2 considering the case of banks having market power to set loan and deposit rates. The third model in chapter 2.3 then considers the case where a bank can extract additional surplus from the borrower if it cannot repay the loan in full. In both cases we will see that regardless how funds are raised, via bank loans or directly from lenders, there is no impact on the outcomes of any market participant.

2.1 Frictionless markets

As a benchmark, let us consider banks acting as pure intermediaries between those in need of funds (borrowers) and those with excess funds (savers). The role of banks in this scenario would be to collect the funds of savers and make them available to borrowers, who then in turn use them to make investments and repay the funds from these proceeds, including interest. The bank then uses these repayments of borrowers to pay savers interest on their funds and return them.

Let us consider an economy consisting of consumers (that will act as lenders), companies (who will be borrowers), and commercial banks. Companies do not have any funds and thus rely on loans to finance their investments I, that yield them a return of R with probability π and cause a total loss of the investment otherwise. Lacking any equity, the company will not be able to repay the loan if the investment is not successful. We have two time periods: in period 1 companies make investments and consumers allocate their funds between consumption C_1, the provision of deposits in banks D, and direct loans to companies $\hat{L}$, while companies invest any funds obtained from bank loans L and these direct loans. Banks take any deposits and lend these to companies. In time period 2, companies repay their bank and direct loans including interest of r_L and r_C, respectively, provided the investment was successful; banks repay the deposits to consumers with interest r_D included if they are able to. Finally, consumers fully consume any funds they obtain in period 2, C_2. Consumers also own banks and companies and as such obtain all the profits they generate in period 2 and can use these to increase their consumption.

The market is perfectly competitive in that all banks take the interest rate on bank loans and deposits as given and these are identical across banks. Furthermore, the interest rate on direct loans is given as well. Companies are also perfectly competitive in that their returns on investment are given and they take all interest rates as given, the same as consumers. We now investigate each market participant in turn before deriving the resulting equilibrium.

Consumers Consumers are endowed with an initial wealth W and decide between consumption in periods 1 and 2 as well as the allocation of any non-consumed wealth from period 1 into bank deposits and direct loans, which are repaid in period 2. Consumers are also the owners of the companies and banks, and as such will receive any profits they make at the end of period 2.

Thus consumers have the following budget constraints:

$$C_1 + \hat{L} + D = W, \tag{2.1}$$

$$C_2 = \Pi_C + \Pi_B + \pi (1 + r_C) \hat{L} + \pi (1 + r_D) D, \tag{2.2}$$

where C_i denotes the consumption in period i, $\hat{L}$ the amount of direct lending, D the amount of deposits, Π_C and Π_B the profits of the companies and banks, respectively, and r_C (r_D) the interest paid on direct loans (deposits). We note that direct loans are only repaid if the investment by companies are successful, which happens with probability π. Deposits are also only repaid if the investment of the companies are successful as we will show below when discussing banks.

Consumers will now maximize their utility, subject to the constraints in equations (2.1) and (2.2). We clearly notice from equations (2.1) and (2.2) that bonds and deposits are perfect substitutes for the consumer and hence in equilibrium, we require that

$$r_D = r_C. \tag{2.3}$$

If the interest on deposits would be higher than the interest on direct loans, then all consumers would allocate any funds not consumed into deposits rather than direct loans. The reverse is true if the interest on direct loans is higher than on deposits. Therefore, the interest on these must be equal in equilibrium to ensure direct lending and deposits can co-exist.

Companies Companies seek to maximize their profits over the optimal investment level I. In the absence of equity, they need to finance this investment using debt, either from bank loans (L) or direct lending ($\hat{L}$).

The profits of companies are then given by

$$\Pi_F = \pi \left((1 + R)\, I - (1 + r_C)\, \hat{L} - (1 + r_L)\, L \right) \tag{2.4}$$

and the investment available is

$$I = L + \hat{L}. \tag{2.5}$$

The second equation follows from the assumption that companies do not have any funds of their own to finance investments and hence are restricted to investing the amount they raise as bank and direct loans. Assuming that companies have limited liability, the loan only needs to be repaid if the investment is successful, which happens with probability π. We implicitly assume that the return on investment R is sufficiently large to cover the repayment of the total loan amount if the investment is successful.

Bank loans and direct loans are perfect substitutes for companies in financing their investments. Hence, in equilibrium we need

$$r_C = r_L. \tag{2.6}$$

If the interest on bank loans would be higher than the interest on direct loans, then all companies would prefer to choose direct loans to finance their investments. The reverse is true if the interest on direct loans is higher than on bank loans. Therefore, in equilibrium these two rates have to be equal for bank loans and direct loans to co-exist.

Equilibrium with banks Banks finance their loans through deposits if we assume that they have no equity and do not need to hold any other assets. Hence

$$L = D \tag{2.7}$$

and their profits, neglecting operating costs, are given by

$$\Pi_B = \pi \left((1 + r_L) L - (1 + r_D) D \right). \tag{2.8}$$

We note that bank loans are repaid by the companies with probability π and hence as the bank lacks any equity will only be able to repay its depositors if the bank loans are repaid; this happens with probability π. The objective function of banks is to maximize these profits subject to the constraint $L = D$.

The equilibrium in this economy is easily characterized by equations (2.3) and (2.6) which imply

$$r_C = r_L = r_D \tag{2.9}$$

and thus from equation (2.8) we obtain when using $L = D$ that

$$\Pi_B = 0. \tag{2.10}$$

With consumers being indifferent between deposits and direct loans, they are unaffected by the existence of banks as well as their size. Similarly, companies are indifferent between bank and direct loans, hence the presence and size of banks does not affect them either. Therefore, the existence of banks and their size in terms of deposits and bank loans are irrelevant in our economy.

This result of banks being redundant in our economy depends crucially on banks offering no reduction in transaction costs when using deposits and bank loans instead of direct loans, as well as providing no additional services to consumers or companies. Banks merely hand through the deposits they receive from consumers and use these to provide bank loans, there is no change in the maturity of loans compared to deposits or other modifications induced by banks. When lifting these assumptions, more sophisticated models will show how banks can increase the welfare in an economy as we will see in chapter 3 when introducing transaction costs and chapter 4 when considering that banks transform short-term deposits into long-term loans.

Summary This model shows that banks as pure intermediaries have no impact in a perfectly competitive economy without any transaction costs and where banks have no inherent advantage over consumers in providing loans to companies. Of course, we could introduce some friction into our model, for example by adding search costs to match borrowers and lenders in direct lending. With banks able to reduce these transaction costs, they can become imperfectly competitive and thus make profits, such that their existence will affect the economy by increasing or decreasing the optimal amount of loans provided to companies. It is, however, that this result is induced by the introduction of such frictions and not by the very nature of banks.

Reading Freixas & Rochet (2008a, ch. 1.7)

2.2 Banks with market power

Loans are used to finance consumption or investment if their own funds are not sufficient. These loans are then repaid from future income, in the case of investments this income can be derived from the investment itself and in the case of consumption this will generally be other income that is obtained in a later time period. Hence loans allow to bring forward expenditure, which is then repaid from future income. Such loans can be arranged directly between those market participants that are seeking to bring forward expenditure and those that seek to delay their expenditure and are therefore not in immediate use of their funds. Alternatively, excess funds can be deposited with a bank who then provides a loan to those seeking additional funds from these deposits. We will compare these two possibilities to assess the implications banks have for the optimality of borrowing and lending decisions.

Direct lending Let us consider a situation in which consumers need to decide their consumption allocation in two time periods; a similar argument can be made for investments by companies. Consumer i has an income of W_i in each of these two time periods. In time period 1 he can decide to postpone same consumption by a granting loan or making a deposit D_i that bears interest r_D and consume their proceeds in time period 2, when they are repaid to him with interest r_D. Alternatively, he can bring forward consumption to time period 1 by taking out a loan L_i; this loan is repaid in time period 2 with interest rate r_L, by reducing consumption. Hence consumption in time period 1 and 2, respectively, are given by

$$\begin{aligned} C_i^1 &= W_i - D_i + L_i, \\ C_i^2 &= W_i + (1 + r_D)D_i - (1 + r_L)L_i. \end{aligned} \tag{2.11}$$

The utility function of consumer i is given by

$$U_i\left(C_i^1, C_i^2\right) = u(C_i^1) + u(C_i^2), \tag{2.12}$$

where we ignore discounting between the two time periods. Consumers choose the optimal amounts of deposits and loans, respectively, and we obtain the first order conditions

$$\begin{aligned} \frac{\partial U_i\left(C_i^1, C_i^2\right)}{\partial D_i} &= -\frac{\partial u(C_i^1)}{\partial C_i^1} + (1 + r_D)\frac{\partial u(C_i^2)}{\partial C_i^2} = 0, \\ \frac{\partial U_i\left(C_i^1, C_i^2\right)}{\partial L_i} &= \frac{\partial u(C_i^1)}{\partial C_i^1} - (1 + r_L)\frac{\partial u(C_i^2)}{\partial C_i^2} = 0. \end{aligned} \tag{2.13}$$

From these conditions we easily get that

$$\frac{\partial u(C_i^1)}{\partial u(C_i^2)} = 1 + r_D, \tag{2.14}$$
$$\frac{\partial u(C_i^1)}{\partial u(C_i^2)} = 1 + r_L,$$

for those depositing or granting loans and those taking loans, respectively. We assume that due to perfect competition between consumers loan and deposit rates are taken as given. We see that the marginal rate of substitution between consumption in time periods 1 and 2 must equal the deposit and loan rate, respectively. For a viable solution of these equations, we of course require $D_i \leq W_i$ and $(1 + r_L)L_i \leq W_i$, which for simplicity we assume to be fulfilled.

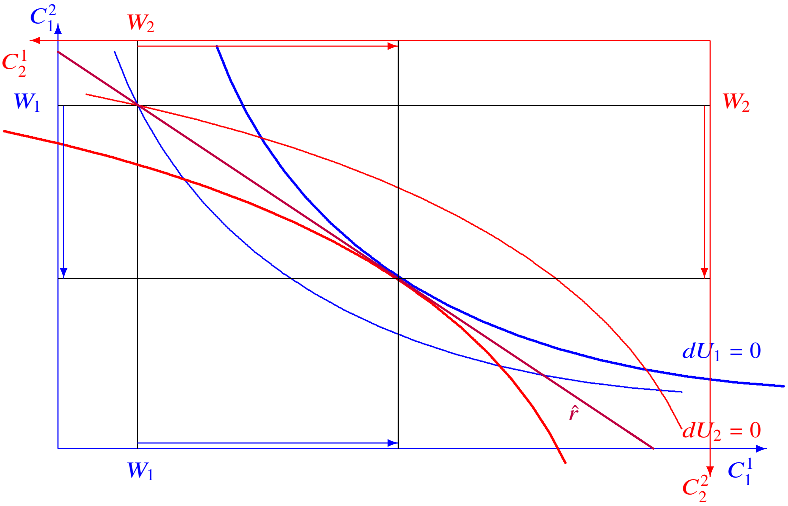
If $r_D > r_L$, consumers could take out a loan and instantly deposit/lend out the proceeds again. This would not affect consumption in time period 1, but increase consumption in time period 2 as the interest earned on deposits/loans exceeds that paid on loans. Hence demand for borrowing and providing loans/deposits would be infinite, thus we need $r_D \leq r_L$. In this case a consumer would not take out a loan and grant a loan or make a deposit at the same time as with the same arguments, consumption in time period 2 would be reduced.

With the usual assumption of concave utility functions $u(\cdot)$, we see from equation (2.14) that consumers with a low income, W_i, are more likely to take out a loan. Their high marginal utility in time period 1 would reduce due to increased consumption and increase in time period 2, reducing the marginal rate of substitution. For depositors the marginal rate of substitution would be increased even more, making it less likely that the first order condition in equation (2.14) is fulfilled as the deposit rate is smaller than the loan rate.

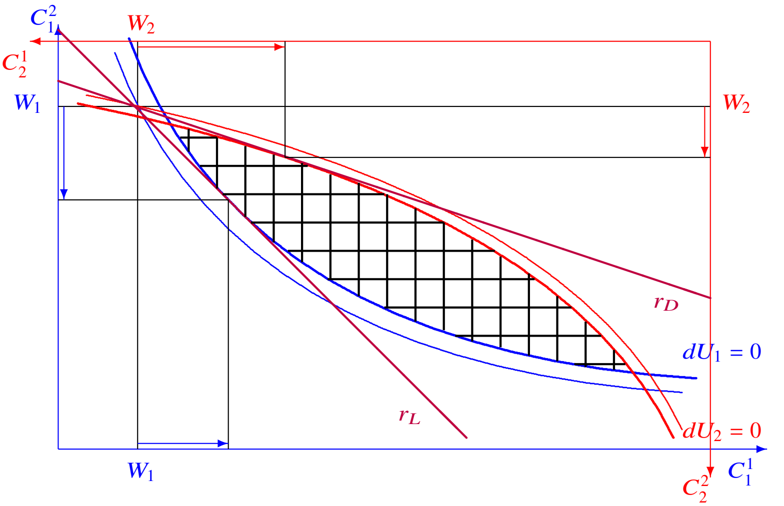
In the absence of an intermediary, consumers would interact directly and thus $r_L = r_D = \hat{r}$ as the deposit of one consumer is the loan of another consumer. Hence from equation (2.14) we require that both, depositors and borrowers, fulfill

$$\frac{\partial u(C_i^1)}{\partial u(C_i^2)} = 1 + \hat{r}. \tag{2.15}$$

In equilibrium the marginal rates of substitution would be identical for borrowers and lenders. Figure 2.1a illustrates this equilibrium for two consumers with incomes W_1 and W_2 for each time period, respectively, using an Edgeworth box, where consumer 1 is shown in blue and consumer 2 in red. The equilibrium is indicated by the point where the bold indifference curves of the depositor and borrower are tangential, indicating they have the same marginal rate of substitution as required from equation (2.15). In addition, its slope is identical to the budget cnstraint implied by the interest rate charged, coloured purple. As we can easily see, the resulting equilibrium is Pareto-efficient. Having established the equilibrium with direct lending, we can now introduce a bank that takes deposits and provides the loan.



(a) Equilibrium with direct lending



(b) Equilibrium with bank lending

Fig. 2.1: Pareto-efficiency with direct and bank lending

Bank lending Let us assume that consumers can only deposit their excess funds with a bank and banks are the only source of loans. Banks seek to maximize their profits, which have two sources; any deposits not lent out are invested at the risk free rate, r, and secondly banks seek profits from interest on loans, which are reduced by paying interest on deposits. Thus bank profits are given by

$$\Pi_B = (1 + r)(D - L) + (1 + r_L) L - (1 + r_D) D. \tag{2.16}$$

We obviously require $D \leq L$ as the bank cannot lend out more in loans than it receives in deposits, neglecting equity and other funding sources here. Furthermore, we need that $(1 + r_D) D \leq (1 + r_L) L$ to ensure all deposits can be returned with interest. We assume for the remainder that these constraints are not binding without changing results. Our direct implication of these constraints is that $r_L \geq r_D$ as required to prevent consumers from demanding infinite amounts of deposits and loans.

Banks set their interest rates in order to maximise their own profits and for a monopolistic bank the first order conditions become

$$\begin{aligned} \frac{\partial \Pi_B}{\partial r_D} &= (1 + r) \frac{\partial D}{\partial r_D} - D - (1 + r_D) \frac{\partial D}{\partial r_D} = 0, \\ \frac{\partial \Pi_B}{\partial r_L} &= -(1 + r) \frac{\partial L}{\partial r_L} + L + (1 + r_L) \frac{\partial L}{\partial r_L} = 0. \end{aligned} \tag{2.17}$$

We here note that the interest rate will have an impact on the demand for deposits and loans, with $\frac{\partial D}{\partial r_D} \geq 0$ and $\frac{\partial L}{\partial r_L} \leq 0$. We can now solve for the deposit and loan rate, respectively, to obtain

$$\begin{aligned} 1 + r_D &= 1 + r - \frac{D}{\frac{\partial D}{\partial r_D}} \leq 1 + r, \\ 1 + r_L &= 1 + r - \frac{L}{\frac{\partial L}{\partial r_L}} \geq 1 + r, \end{aligned} \tag{2.18}$$

where the inequality arises from the sign of the marginal impact of the deposit (loan) rate on the demand for deposits (loans). We thus easily see that $1+r_L \geq 1+r \geq 1+r_D$. The incentives for consumers are unchanged from the case of direct lending, thus their optimal demand for deposits and loans will be determined by equation (2.14), taking the loan and deposit rates set by the bank as given. We now see that the marginal rates of substitution for consumption in periods 1 and 2 are different for depositors and borrowers. This indicates that due to the presence of banks, the resulting equilibrium is not Pareto-efficient anymore.

Figure 2.1b illustrates this result using an Edgeworth box for the same two consumers as in direct lending. We see that the cause for the different marginal rates of substitution is the change in the slope of the budget constraint at $\left(C_1^1, C_1^2\right) = (W_1, W_1)$ and $\left(C_2^1, C_2^2\right) = (W_2, W_2)$, where for lower consumption in time period 1 the consumer would be a depositor and for higher consumption he would be a borrower, which have different interest rates. The resulting equilibrium has to fulfill the budget constraint, exhibiting this change of slope and the indifference curve being tangential to it. This precludes the marginal rates of substitution to be identical for borrowers and depositors. The resulting equilibrium shows a smaller adjustment of consumption in time periods 1 and 2, compared to the initial allocation of consuming W_i in each time period. The reason is that with deposit rates lower than with direct lending, the incentives to save are reduced and the higher loan rate reduces borrowing. The hatched area in the figure shows the area that shows an allocation that is a Pareto-improvement, but which is unattainable in the presence of banks.

This result emerges from the fact that banks can affect deposit and loan rates and will set them optimally. Even if we assume that banks are competitive and make zero profits or that consumers own the banks and obtain their profits in time period 2 for consumption, the result will be unchanged in principle and it is therefore not the result of banks generating profits or bank profits being extracted from consumers. In both cases banks will set loan and deposit rates that are different from each other and hence the allocation is not Pareto-optimal for consumers. Therefore, the introduction of banks with market power reduces the welfare of consumers and consumers would be better off, if banks were not present.

Summary Banks introduce friction into the market of lending and borrowing. Due to banks maximizing their profits, deposit and loan rates are not identical and hence marginal rates of substitution between consumption in different time periods are different for borrowers and lenders, leaving room for an improvement in their welfare. It is therefore that banks affect the allocation of consumption across time and thus affecting economic outcomes.

Reading Spulber (1999, Ch. 2.2)

2.3 The effect of bank monitoring

Banks specialise in the provision of loans and as such it is reasonable to assume that they have accumulated a level of expertise that individuals providing loans directly would not have. Thus bank lending has an advantage over direct lending in that this knowledge can be used by banks to provide more advantageous conditions to borrowers.

Let us assume that the investment of a company succeeds with probability π_i, giving a return of R_i. If the investment is not successful, no funds are available to the company. There are two possible states, a high state H and a low state L, that occur with probability p and $1 - p$, respectively. We surmise that the probability of the investment succeeding is higher in the high state, $\pi_H > \pi_L$, but the return of a successful investment is lower in the high state, $R_H < R_L$, and the expected returns of the investment satisfies $\pi_H (1 + R_H) > 1 + r_D > \pi_L (1 + R_L)$, where r_D denotes the interest paid on deposits. This implies that the high state H is less risky, but also yields a lower return if successful, but also that the expected return in this state H covers the costs of providing funds through deposits $(1 + r_D)$, while in the low state L this is not the case. This reflects a positive risk-return relationship of investments, commonly found in the finance literature. Furthermore we find that providing the loan in state H is desirable as it covers its funding costs, while in state L the loan would not cover its costs.

Assuming the company has no own resources to fund the investment, they rely on additional funds I, that can either be raised in the form of equity, direct lending from individual lenders, or bank lending. We will now consider each of these funding sources in turn and compare their desirability for companies.

Equity issue The company can issue equity to fund the investment and provide new shareholders with a fraction ν of the company. The expected return to shareholders is $p\pi_H(1+R_H)+(1-p)\pi_L(1+R_L)$. If the high state H occurs, which happens with probability p, the investment is successful with probability π_H, giving a return of R_H and no return otherwise. Alternatively, with probability $1-p$, the low state L occurs and the investment is successful with probability π_L yielding a return of R_L and no return if the investment is not successful. The new shareholders receive a fraction ν of these expected returns. Shareholders could use the alternative investment of investing their monies into bank deposits and obtaining a return of r_D; the surplus over this alternative investment generated is thus given by

$$\Pi_D^E = \nu\left(p\pi_H(1+R_H)+(1-p)\pi_L(1+R_L)\right)I-(1+r_D)I. \tag{2.19}$$

Assuming that the market for shares is competitive, we require that this surplus is zero, hence $\Pi_D^E=0$, from which we obtain that new shareholders obtain a fraction

$$\nu = \frac{1+r_D}{p\pi_H(1+R_H)+(1-p)\pi_L(1+R_L)} \tag{2.20}$$

of the company for their equity of I.

For the initial shareholders, holding a fraction $1-\nu$ of the company, the expected profits are then given by

$$\begin{aligned}\Pi_C^E &= (1-\nu)\left(p\pi_H(1+R_H)+(1-p)\pi_L(1+R_L)\right)I \qquad (2.21)\\ &= \left(p\pi_H(1+R_H)+(1-p)\pi_L(1+R_L)-(1+r_D)\right)I,\end{aligned}$$

where the last equation has been obtained after inserting for ν from equation (2.20).

These profits to the initial shareholders after issuing equity can now be compared with the profits they would obtain when raising loans, either directly or using a bank.

Direct lending If the company borrows directly from individuals, we assume that in case of a default, the lender has no mechanism to enforce the repayment of any funds. If the low state L occurs, the company will not be able to repay the full amount due, as by assumption $\pi_L(1+R_L)<1+r_D$. If we assume that the lender cannot distinguish the reason for default, that is whether it is the result of a failed investment or the occurrence of the low state L, the company can easily claim that the investment has failed even if it was successful but the low state L had been realised. Thus, in this situation, the company will not make any repayments to the lender. Charging a loan rate of r_C for this loan of size L, the lender only obtains repayment in the high state H and the expected surplus over the investment into bank deposits are given by

$$\Pi_D^C = p\pi_H\left(1 + r_C\right) L - \left(1 + r_D\right) L. \tag{2.22}$$

Again, competition in the market for direct lending implies that this surplus is zero, thus $\Pi_D^C = 0$, and hence we obtain that the interest charged on the direct loan is given by

$$1 + r_C = \frac{1 + r_D}{p\pi_H}. \tag{2.23}$$

The company repays the loan only in the high state H and retains the return made in the low state L without making payments to the lender. Hence, its expected profits are given by

$$\begin{aligned}\Pi_C^C &= p\pi_H\left(\left(1 + R_H\right) L - \left(1 + r_C\right) L\right) + (1 - p)\,\pi_L\left(1 + R_L\right) L \qquad (2.24)\\ &= \left(p\pi_H\left(1 + R_H\right) + (1 - p)\,\pi_L\left(1 + R_L\right) - \left(1 + r_D\right)\right) L\\ &= \Pi_C^E,\end{aligned}$$

where the second equality emerges when inserting equation (2.23) for the loan rate r_C and we assumed that the entire investment is funded by a loan, thus $I = L$. Comparing this expression with the profits the existing shareholders obtain when issuing equity, equation (2.21), shows that these profits are identical. As in both cases the new shareholders and the direct lenders make zero profits, they would also be indifferent between these two finance forms.

Bank loan We assume that a bank has the ability to monitor companies and distinguish between a failure of investment and the low state L occurring. Thus, it can enforce that in case the loan is not repaid in full, all resources of the company are given to the bank. This means that in the low state L, banks obtain $\pi_L\left(1 + R_L\right) L$, the resources available to the company. The profits of the bank using loan rate r_L are then given by

$$\Pi_B = \left(p\pi_H\left(1 + r_L\right) + (1 - p)\,\pi_L\left(1 + R_L\right)\right) L - \left(1 + r_D\right) D, \tag{2.25}$$

where D denotes the deposits the bank holds. We assume here that the bank does not hold any other assets than the loan provided to the company or has any other sources of finance for the loan, such that $D = L$. In the high state H the loan is repaid in full with probability π_H and in the low state L the successful investment is seized by the bank. Competitive banking markets imply that banks make zero profits, thus $\Pi_B = 0$, and we hence obtain that the loan rate is given by

$$1 + r_L = \frac{\left(1 + r_D\right) - (1 - p)\,\pi_L\left(1 + R_L\right)}{p\pi_H}. \tag{2.26}$$

The profits of the company are now such that they repay the loan in the high state H, if successful, and in the low state L lose the revenue they obtained from their investment. Thus we can use equation (2.26) for the loan rate and obtain the company profits as

$$\begin{aligned} \Pi_C^B &= p\pi_H \left((1 + R_H) L - (1 + r_L) L \right) \\ &= \left(p\pi_H (1 + R_H) + (1 - p) \pi_L (1 + R_L) - (1 + r_D) \right) L \\ &= \Pi_C^E. \end{aligned} \tag{2.27}$$

Hence, we see that bank loans provide the same profits to companies as raising equity or direct lending. The ability of the bank, through monitoring distinguishing between failed investments and the low state L and then to extract any surplus from the company if it cannot repay its loan in the low state L, is compensated for by the lower interest rate to be paid on the loan that is repaid in the high state H. The competition between banks in our model takes into account the ability of the bank to recover funds in the low state L through a lower loan rate, which is paid in the high state H. These two effects exactly offset each other and companies lose the same amount in the low state L, when they have to give up their successful investment, as they gain from lower loan rates that are paid on successful investments in the high state H. The consequence is that companies are indifferent between any of the three potential financing sources. As lenders are in all cases also making the same surplus of zero, all market participants are indifferent between equity issue, direct lending and bank loans to finance the investment. A consequence is that whether banks exist in an economy is irrelevant, they neither increase nor decrease the welfare of any market participant. Making zero profits due to the assumed perfect competition between providers of funds to companies, will also negate any effect that profits generated by banks might have on the economy.

Summary Even if banks are uniquely able to monitor borrowers and extract additional surplus from them in cases they cannot repay their loans in full, they do not gain an inherent advantage from this ability. While borrowers are more under scrutiny in this case, and therefore have to repay the loan in more situations, the increased repayments reduce the losses of the bank from default and therefore in competitive markets the loan rate will be lower. This reduction in the loan rate, which applies if the company fully repays its loan, is exactly offset with the increased amount that is actually repaid to the bank. This makes the company equally well off than when it had taken out a direct loan or issued equity. Depositors in banks, direct lenders and equity investors are also equally well off in all scenarios, all making zero profits due to our assumption of perfect competition. It is therefore that banks do not add any value in the economy, nor do they reduce welfare. Banks are thus irrelevant and their absence would not affect the welfare in the economy. However, going beyond the view of banks acting merely as intermediaries between borrowers and lenders, we will see how banks can increase the welfare in an economy, be it through the reduction of transaction costs as we will explore in chapter 3 or in chapter 4 when considering that banks transform short-term deposits into long-term loans.

Reading Bolton & Freixas (2000)

Conclusions

If markets are perfect in the sense of having no transaction costs, all market participants having the same information, and being competitive, banks do not affect the outcome for companies or lenders; the welfare of both market participants are equal whether banks exist or are absent. This result holds even in cases where banks have an advantage over direct lending in that they can ensure that any surplus of the company is extracted more efficiently if the loan cannot be repaid in full. The higher loan repayments borrowers make to banks are compensated for through lower loan rates, offsetting these additional costs. It is thus that the existence of banks is not improving welfare, nor is it reducing welfare. It can be claimed that banks are irrelevant. If banks, however, have market power an the ability to determine loan and deposit rates, the resulting equilibrium will reduce welfare to consumers and banks introduce some friction into the economy.

Using such results as presented here, is often used to justify that banks are ignored in (macroeconomic) models. If they do not affect the outcomes in an economy, it should be possible to ignore banks for the sake of simplicity, without affecting results in a meaningful way. Of course, markets are not frictionless and therefore it is often that such models then introduce some friction arising from the existence of banks, such as imperfect competition between banks or banks facing costs in their provision of loans, allowing banks to be profitable, or giving rise to different loan and deposit rates. Other modifications might include the need for banks to retain a certain amount of cash and thus not being able to lend out all deposits, which would then again give rise to different loan and deposit rates. These frictions, however, do not consider the role of banks in the actual economy adequately.

The following chapters will show that the relevance of banks does not arise due to the existence of frictions, that is deviations from the ideal market conditions considered here. Instead banks can offer a number of benefits to an economy that range from reducing transaction costs and providing a more effective monitoring of borrowers in chapter 3 to the transformation of short-term deposits into long-term loans in chapter 4. It is these benefits of banks that make their existence beneficial to an economy and the view of banks as simple intermediaries is very incomplete and does not address the main role they play in an economy. While here we focus on the benefits of banks, it is worth remarking already at this point that the existence of banks does not only provide benefits, there might well be costs that arise if banks fail. Especially if such bank failures spread, called systemic risk, can these costs be substantial. These aspects are discussed further in part VI.

Chapter 3

Reducing transaction costs

Having considered the case of frictionless markets in chapter 2, we will now introduce transaction costs into the borrowing and lending process. A key result thus far was that the absence of transaction costs does not give bank lending any inherent advantage over direct lending and therefore the existence of banks is irrelevant. Whether a bank exists in an economy or not, does not affect the welfare in that economy.

Transaction costs can take several forms. One transaction cost would be the negotiation of the loan contract itself, which needs to agree not only the amount of the loan, the interest to be paid and the maturity of the loan, but also the use of the loan and any other safeguards the bank or a direct lender might want to seek to ensure the loan is repaid whenever possible. In order to inform this negotiation, any lender must have sufficient information to assess the borrowers' prospects of being able to repay the loan. The collection of such information and the subsequent negotiation will be time-consuming and costly. We consider these negotiation costs in chapter 3.1 and establish that the use of banks is (mostly) beneficial to borrowers and lenders.

The involvement of lenders does not end when a loan is given to a borrower as throughout the life-time of the loan, lenders will continue to monitor the borrower. This continued monitoring by a lender will ensure that the loan is used for the purpose it was originally given and on which the assessment of the creditworthiness of the borrower was based. In chapter 3.2, we will see how delegating such monitoring to banks can increase welfare. Delegating monitoring to banks will reduce the costs of lenders as duplicate efforts can be avoided if a borrower will obtain loans from multiple lenders; these reduced costs will be passed on to borrowers if markets are sufficiently competitive and benefits therefore lenders and borrowers alike. However, in some situation such duplication of effort might be beneficial. If only banks are able to monitor lenders, we will also see that providing deposits rather than direct lending does not introduce additional risks as banks can structure their lending such that deposits are safe. Finally, we will also see in chapter 3.3 how banks can provide direct benefits to lenders by reducing their costs of borrowing through the pooling of loans in banks, without the need to have advantages in the monitoring of borrowers.

It is the ability of banks to reduce transaction costs that make them intermediaries benefitting an economy. In this sense, banks can be seen as institutions that reduce the costs of borrowing and lending, while retaining their position as intermediaries bringing together these two market participants.

3.1 Negotiation costs

Let us take the view that banks act merely as intermediaries by facilitating the matching of borrowers and lenders. We assume that borrowers and lenders can negotiate a contract directly among themselves at a cost of C for the borrower and lender each, or use a bank as an intermediary where no additional costs are incurred. These costs would include finding a borrower (lender) that matches the lender's (borrower's) preferences in terms of risks, but also size, time to maturity of the loan, and other conditions. They would also include the negotiation of these conditions themselves. Given the set procedures of banks, we assume that no such negotiation is required when choosing bank loan and matching does not involve meaningful costs as banks can pool the funds of many depositors and distribute them onto multiple borrowers, making dealing with banks cost-free. Similarly, a lender making a deposit will also not incur any costs as deposits are standard form contracts that allows monies to be withdrawn at any time. Thus using banks reduces the negotiation costs of borrowers and lenders.

We will evaluate the situation where only direct lending between borrowers and lender occurs, i. e. no banks exist, then continue to explore the case where all borrowing and lending is conducted via banks, and finally look at the case where both direct lending and bank lending might co-exist. For simplification, we only model the negotiation of the interest rates and assume all other conditions to be either fixed or to be negotiated at no costs.

Direct lending only A company has an investment opportunity with a return of R if successful, which happens with probability π, and if it is not successful no funds are generated. We assume the company has no own funds and relies fully on a loan L for its investment. Then, with a loan rate of r_C and negotiation costs of C, their profits from direct lending are given by

$$\hat{\Pi}_C = \pi\left((1+R)\,L - (1+r_C)\,L\right) - C, \tag{3.1}$$

where we assume that companies have limited liability and only repay the loan if their investment is successful. For the lender (which in anticipation of introducing a bank, we call 'depositor') we find that they are repaid the loan, including interest r_C, with probability π. They have an initial outlay of the loan amount L and face negotiation costs of C, and hence their profits are given by

$$\hat{\Pi}_D = \pi\,(1+r_C)\,L - L - C. \tag{3.2}$$

The two parties, borrower and lender (depositor), engage in Nash bargaining to determine the optimal interest rate r_C. The outside option for both parties is to walk away from the negotiations and not enter any contract, having incurred negotiation costs C. Thus we maximize

$$\mathcal{L} = \left(\hat{\Pi}_C + C\right)\left(\hat{\Pi}_D + C\right), \tag{3.3}$$

which gives us

$$\frac{\partial \mathcal{L}}{\partial (1 + r_C)} = \pi L \left(\hat{\Pi}_C + C\right) - \pi L \left(\hat{\Pi}_D + C\right) = 0 \tag{3.4}$$

or $\hat{\Pi}_D = \hat{\Pi}_C$. Solving this relationship after inserting from equations (3.1) and (3.2), we get the expected repayment from the loan as

$$\pi (1 + r_C) L = \frac{1}{2} (\pi (1 + R) + 1) L \tag{3.5}$$

and the expected profits of the borrower and lender are given by

$$\hat{\Pi}_C = \hat{\Pi}_D = \frac{1}{2} (\pi (1 + R) - 1) L - C. \tag{3.6}$$

The participation constraint requires that this arrangement is profitable for both parties, hence we need $\hat{\Pi}_C = \hat{\Pi}_D \geq 0$, which solves for

$$C \leq C^* = \frac{1}{2} (\pi (1 + R) - 1) L. \tag{3.7}$$

Direct lending is feasible only if $C \leq C^*$. In situations where the negotiation costs are higher than C^*, direct lending will not be profitable and will hence not be observed. This threshold C^* is increasing the more likely the investment is succeeding (π) as the loan as more likely to be repaid and therefore higher costs can be incurred without eroding profits fully. A higher return on investment R also leads to a higher threshold because in this case a higher loan rate can be negotiated that allows both parties to be profitable at higher negotiation costs. A larger loan L allows for a wider spread of the costs C and makes lending more profitable.

Bank lending only Assume now that lending is only conducted through banks and lenders become depositors in the bank. Using a bank imposes no negotiation costs on any of the participants. Any party, depositor, borrower, and bank, can walk away from negotiations for free at any time and not enter any contract. With loan rates r_L and deposit rates r_D, we then have the profits of companies, depositors, and the bank given by

$$\begin{aligned}\Pi_C &= \pi\left(\left(1+R\right)L-\left(1+r_L\right)L\right), \\ \Pi_D &= \pi\left(1+r_D\right)L-L, \\ \Pi_B &= \pi\left(\left(1+r_L\right)L-\left(1+r_D\right)L\right).\end{aligned} \tag{3.8}$$

Companies and banks both have limited liability. Therefore companies will repay their loans only if the investment is successful, and banks will be able to repay deposits only if they have been repaid their loans.

The bank and depositor negotiate the deposit rate using Nash bargaining, which gives us the objective function $\mathcal{L} = \Pi_B \Pi_D$ as both can walk away from the negotiations without having incurred any costs. The first order condition for a maximum is given by

$$\frac{\partial \mathcal{L}}{\partial\left(1+r_D\right)} = \pi L \Pi_B - \pi L \Pi_D = 0, \tag{3.9}$$

and hence $\Pi_B = \Pi_D$, which easily solves for

$$\pi\left(1+r_D\right)L = \frac{1}{2}\left(\pi\left(1+r_L\right)+1\right)L. \tag{3.10}$$

Similarly, for the negotiation of the bank and company on the loan rate, the objective function is $\mathcal{L} = \Pi_B \Pi_C$ and we get

$$\frac{\partial \mathcal{L}}{\partial\left(1+r_L\right)} = \pi L \Pi_C - \pi L \Pi_B = 0 \tag{3.11}$$

and the resulting $\Pi_B = \Pi_C$ solves for

$$\pi\left(1+r_D\right)L = 2\pi\left(1+r_L\right)L - \pi\left(1+R\right)L. \tag{3.12}$$

Combining equations (3.10) and (3.12), we solve these two equations for the expected repayments of the loan to the bank and the deposits to the depositors, respectively, to become

$$\begin{aligned}\pi\left(1+r_L\right)L &= \frac{2}{3}\pi\left(1+R\right)+\frac{1}{3}, \\ \pi\left(1+r_D\right)L &= \frac{1}{3}\pi\left(1+R\right)+\frac{2}{3},\end{aligned} \tag{3.13}$$

from which we easily obtain the expected profits by using equation (3.8) to be

$$\Pi_B = \Pi_C = \Pi_D = \frac{1}{3}\left(\pi\left(1+R\right)-1\right)L. \tag{3.14}$$

A participation constraint is that profits are positive, thus $\Pi_B = \Pi_C = \Pi_D \geq 0$, which easily solves for $\pi\left(1+R\right)L \geq L$. If this condition is fulfilled, implying that the expected outcome of the investment is at least covering its initial outlay, bank lending will be profitable.

Direct and bank lending The more realistic case is that direct lending and bank lending co-exist. A borrower might negotiate with a bank, but if this fails, it might well enter negotiation using direct lending; the same applies to a depositor. The process might also work in the opposite way that a borrower might negotiate direct lending and on failing to reach an agreement, seeks a loan from a bank, likewise for the depositor. Thus, borrowers and lenders have outside options in their negotiation, apart from not entering any contract at all. The bank still has only the option to enter a contract with the depositor and lender, thus has no outside option.

When negotiating with a bank, the outside options would be to revert to direct lending, giving profits of $\hat{\Pi}_C$ for a borrower and $\hat{\Pi}_D$ for a lender as determined in equation (3.6). The profits when engaging in bank lending are given by equation (3.8). Hence the objective function for the negotiation between the bank and depositor is $\mathcal{L} = \Pi_B \left(\Pi_D - \hat{\Pi}_D\right)$, as the bank still has no outside option. The first order condition

$$\frac{\partial \mathcal{L}}{\partial \left(1 + r_D\right)} = \pi L \Pi_B - \pi L \left(\Pi_D - \hat{\Pi}_D\right) = 0 \tag{3.15}$$

implies $\Pi_B = \Pi_D - \hat{\Pi}_D$. Inserting from equations (3.2) and (3.8), this can be solved for

$$2\pi \left(1 + r_D\right) L = \pi \left(1 + r_L\right) L + \pi \left(1 + r_C\right) L - C. \tag{3.16}$$

Similarly, the negotiation between the bank and company maximizes $\mathcal{L} = \Pi_B \left(\Pi_C - \hat{\Pi}_C\right)$, and following the same steps as above, we obtain that

$$2\pi \left(1 + r_L\right) L = \pi \left(1 + r_D\right) + \pi \left(1 + r_C\right) L + C. \tag{3.17}$$

Finally, the company and depositor negotiating directly would require the maximization of $\mathcal{L} = \left(\Pi_C - \hat{\Pi}_C\right)\left(\Pi_D - \hat{\Pi}_D\right)$ as the objective function. Here the outside options for both, lender and borrower, are to use deposits and bank lending, respectively. As $\Pi_C - \hat{\Pi}_C = \Pi_D - \hat{\Pi}_D = \Pi_B$ from the first order conditions of the negotiation with the bank, we have $\mathcal{L} = \Pi_B^2$. As Π_B is independent of $1 + r_C$, the first order condition $\frac{\partial \mathcal{L}}{\partial (1+r_C)} = 0$ is fulfilled for all values of r_C. We thus have one free parameter and assume we set the deposit rate independently. Solving equations (3.16) and (3.17), we get the expected repayments of the bank and direct loan, respectively, as

$$\begin{aligned} \pi \left(1 + r_L\right) L &= \pi \left(1 + r_D\right) L + \frac{2}{3} C, \\ \pi \left(1 + r_C\right) L &= \pi \left(1 + r_D\right) L + \frac{1}{3} C. \end{aligned} \tag{3.18}$$

We note that a bank loan attracts a higher interest rate than direct lending. This is to cover the profits of the bank, but the company might still benefit from bank loans as no negotiation costs are incurred, reducing the overall costs of the loan.

Inserting these results into the profits of borrowers, lenders (depositors), and the bank, we easily get from equations (3.1), (3.2), and (3.8) that

$$\begin{aligned}
\Pi_B &= \frac{2}{3}C > 0, \\
\Pi_D &= \pi\,(1 + r_D)\,L - L, \\
\Pi_C &= \pi\,(1 + R)\,L - \pi\,(1 + r_D)\,L - \frac{2}{3}C, \\
\hat{\Pi}_D &= \pi\,(1 + r_D)\,L - L - \frac{2}{3}C = \Pi_D - \frac{2}{3}C < \Pi_D, \\
\hat{\Pi}_C &= \pi\,(1 + R)\,L - \pi\,(1 + r_D)\,L - \frac{4}{3}C = \Pi_C - \frac{2}{3}C < \Pi_C.
\end{aligned} \tag{3.19}$$

From the final two results we thus see that using the bank is preferred by companies and depositors, with the bank also being profitable. Therefore, if a bank is available, the absence of negotiation costs with banks makes its use preferable to direct lending. The reason is that the bank would not take full advantage of the lower costs but the total cost savings of $2C$ are distributed equally between all market participants, making everyone better off using the bank.

In order for depositors to use the bank we need $\Pi_D \geq 0$, which implies

$$\pi\,(1 + r_D)\,L \geq L, \tag{3.20}$$

and thus depositors would participate as long as the exogenously set deposit rate is sufficiently high. Similarly, for companies to participate, we require $\Pi_C \geq 0$, implying

$$C \leq C^{**} = \frac{3}{2}\left(\pi\,(1 + R)\,L - \pi\,(1 + r_D)\,L\right). \tag{3.21}$$

Thus, if $C \leq C^{**}$ the company would obtain a loan from the bank in the situation where banks and direct lending co-exist. Higher negotiation costs would imply that companies would not seek a loan at all as the negotiation costs are so high that the loan rate by banks is too high to make borrowing profitable. The company would also not seek a loan directly from a lender as equation (3.19) shows that the profits from this are even lower than from bank lending.

Market structure Having established the conditions for the viability of direct lending only, bank lending only, and the co-existence of both forms of lending, we can now proceed to establish which market structure is preferred by borrowers and lenders. Fist we compare the profits from direct lending only in equation (3.6) and bank lending only in equation (3.14) and we see that banks are preferred by companies and depositors if $\Pi_C = \Pi_D \geq \hat{\Pi}_C = \hat{\Pi}_D$, which gives us

$$C \leq C^{***} = \frac{1}{6}\left(\pi\,(1 + R) - 1\right)L. \tag{3.22}$$

Therefore if $C \leq C^{***}$ bank lending is preferred to direct lending by both depositors and companies, and for $C > C^{***}$ direct lending is preferred.

Similarly, we can now compare the profits of a market with direct lending only and a market in which direct and bank lending co-exist. Comparing the profits of depositors from equations (3.6) and (3.19), we find that direct lending is preferred to the co-existence of direct and bank lending if

$$\pi (1 + r_D) L \leq \frac{1}{2} (\pi (1 + R) + 1) L - C. \tag{3.23}$$

Similarly, we see that companies prefer direct lending over the co-existence of direct and bank lending if

$$\pi (1 + r_D) L \geq \frac{1}{2}\pi (1 + R) L + \frac{1}{2}L + \frac{1}{3}C. \tag{3.24}$$

These two conditions are not compatible with each other as we can easily verify and hence direct lending is not generally preferred over the co-existence of direct and bank lending, leaving a conflict of interests between companies and depositors on the best market structure.

Comparing the profits of a market with bank lending only and a market with direct and bank lending, we see when comparing equations (3.14) and (3.19) that depositors prefer bank lending if

$$\pi (1 + r_D) L \leq \frac{1}{3}\pi (1 + R) L + \frac{2}{3}L. \tag{3.25}$$

Companies would prefer bank lending if

$$\pi (1 + r_D) L \geq \frac{2}{3}\pi (1 + R) L + \frac{1}{3}L - \frac{2}{3}C. \tag{3.26}$$

These two conditions are compatible if $C \geq C^*$, thus in this case companies and depositors prefer bank lending only over the co-existence of direct and bank lending. In the case that $C < C^*$ we find a conflict of interest on the optimal market structure between companies and depositors.

Figure 3.1 combines our results on the optimal market structure. We see that for higher negotiation costs unsurprisingly bank lending will dominate as they can offer better conditions to companies and depositors due to the absence of negotiation costs. If negotiation costs are lower, direct lending becomes more attractive as the profits of the bank do not have to be extracted from depositors and companies. As negotiation costs are reducing even further, bank lending becomes attractive again as the ability of banks to extract surplus will be limited due to the small benefit they have over direct lending, while still reducing the negotiation costs.

As the expected returns of the investment of the company, $\pi (1 + R)$, increases, the surplus that potentially can be extracted, makes the co-existence of bank and direct lending attractive. The reason is that the threat of companies and depositors engaging directly with each other, will limit the amount of profits that banks can extract. This

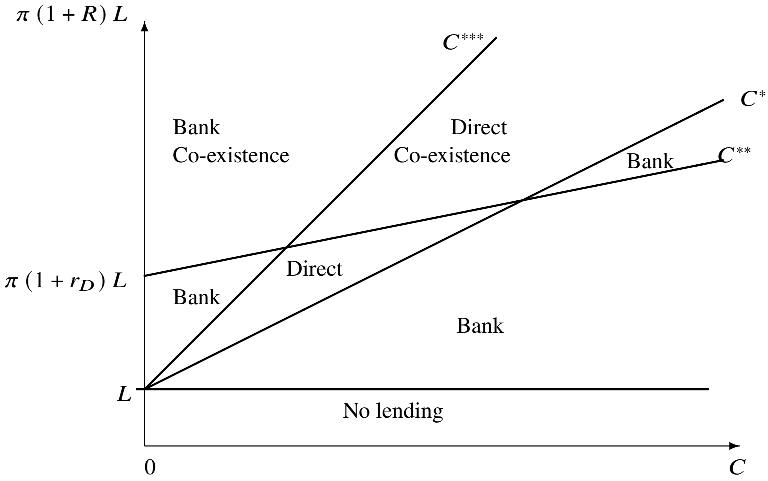


Fig. 3.1: Equilibrium market structures with negotiation costs for direct lending

makes the co-existence of direct and bank lending feasible to companies as long as the outside option of direct lending is attractive enough to be a credible threat. With lower expected returns on investment, this threat is not credible as the surplus available to companies from which banks can generate profits is not sufficient.

While the market structure will allow for the co-existence of direct and bank lending, we know from equation (3.19) that in this case we will only observe bank lending, direct lending is only used an outside option to obtain more attractive loan and deposit conditions from banks. Looking at the observed source of lending in figure 3.2, we see that bank lending dominates for high and very low negotiation costs, while direct lending can be observed for intermediate ranges, although companies with high expected investment returns will prefer to take out bank loans, although direct lending is more attractive for depositors as in this case they obtain a higher fraction of these returns.

Summary Banks emerge as the result of reduced costs of negotiations compared to direct lending. Knowing that their customers will be able to resort to direct lending if the conditions offered are not sufficiently competitive to both, depositors and companies, banks will share the benefits of these lower negotiation costs. This leads to a situation where, in most cases, bank lending is chosen, even though the interest rate on bank loans is higher than on direct loans and the interest paid to depositors is less than the loan rate in direct lending. The saved costs, which are partially retained by borrowers and depositors, allow for this result. At the same time, banks appropriate some of the cost savings to generate a profit. It is only for intermediate negotiation costs that direct lending would be the optimal solution for all concerned. If the investment returns of companies are high, the resulting surplus that could be

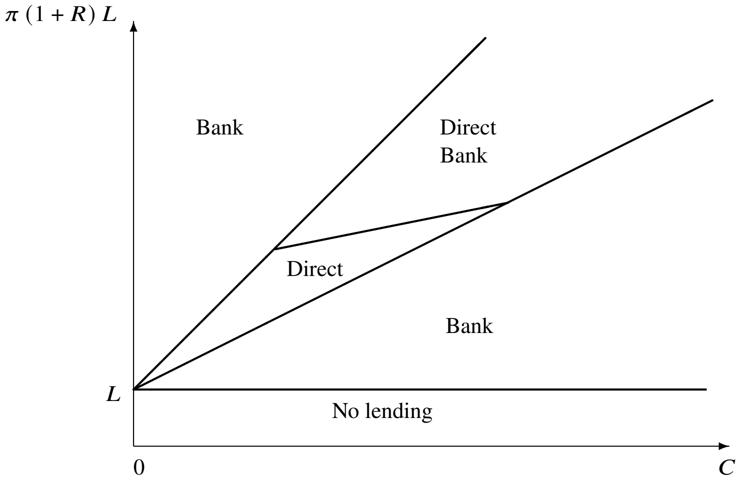


Fig. 3.2: Equilibrium observed lending with negotiation costs for direct lending

extracted by banks is substantial and competition in form of direct lending being available, will limit the profits banks make and subsequently benefit companies and depositors when choosing bank loans and deposits, respectively.

Reading Bester (1995)

3.2 Delegated monitoring

Lending funds to a company is risky in that the loan might not be repaid by the borrower. To mitigate this risk, a lender will monitor borrowers. Monitoring by a lender encompasses a range of actions to safeguard the repayment of a loan and can take a wide variety of forms. It typically includes the initial assessment of the risk of the loan to determine whether the loan is given in the first place, but may well continue to ensure the proceeds of the loan are invested as initially agreed and not used in a way to increase the risk to the lender. Finally, it may also include the auditing and sanctioning of any borrower who does not meet the scheduled repayments. The way the monitoring is conducted during the life-time of the loan may vary considerably depending on the specific conditions of the borrower. It may include the requirement to present accounts regularly, ongoing monitoring of payments made and received by the borrower, meetings with relevant senior staff of the company, amongst many other measures.

It would be easy to justify the existence of banks in this context by proposing that they can conduct this monitoring activity at lower costs than general members of the public, for example arising from superior skills, experience, or economies of scale due to their large-scale exposure to loans. In such an interpretation, banks would act as a means to conduct monitoring efficiently. As the following models show, even in the absence of such considerations, banks can be beneficial. Banks avoid the duplication of monitoring effort without introducing a new risk to depositors, namely that the bank might fail and not repay its deposits, see chapter 3.2.1, with the effect of an overall welfare increase. However, as chapter 3.2.2 will show, duplicating monitoring from multiple banks can be beneficial and reduce the risk of loans.

Picking up on the idea that using banks is more efficient as it lowers monitoring costs, chapter 3.2.3 does go beyond the argument that this induces a welfare gain by showing that if only banks, but not individual lenders, are able to monitor borrowers, banks allow for companies to obtain loans that would otherwise have no access to them, even if they are socially desirable. Thereby, the existence of banks increases welfare through granting loans to a wider range of companies, especially more risky companies that may act as innovators in the economy and in the long-run stimulate economic growth.

3.2.1 Avoiding monitoring duplication

We compare loans given directly to borrowers by individual lenders with lending through banks, focussing on monitoring costs. We have N companies (borrowers) seeking loans from M potential lenders (depositor) and each lender seeks to diversify their lending by splitting their funds such that they give a small loan to each borrower. Each lender has funds of L available. An alternative for lenders is to deposit their whole funds in a bank, who then lends to the companies. Each lender as well as the bank face monitoring costs of C for each loan they provide, covering the initial assessment of the borrower as well as subsequent monitoring of the activities of the company. This monitoring might well ensure that the company does not choose an investment that is riskier than accepted by the lender. Hence there are no cost benefits of monitoring for banks, banks and direct lenders face the same costs of C for each loan they provide.

Direct lending With lenders seeking to diversify by providing loans to all N companies, the total funds available by a lender for each borrower is $\frac{L}{N}$. The interest rate charged on the loan is r_C and if the company fails in its investments, it does not repay the loan at all. The company repays the loan with probability π and thus the expected repayment from the loan is $\pi\,(1 + r_C)\,\frac{L}{N}$. The expected return is assumed such that even considering monitoring costs C, it generates profits to the lender for each loan, making direct lending profitable. We thus have

$$\pi\,(1 + r_C)\,\frac{L}{N} - C \geq \frac{L}{N}. \tag{3.27}$$

With M lenders and N borrowers the aggregate monitoring costs in the economy will be NMC.

Bank lending Instead of lending the money directly to companies, lenders could deposit their whole funds with a bank, who in turn provides a loan and monitors companies. In this case lenders (depositors) need to consider the cost arising from the bankruptcy of the bank if loans are not repaid and the bank cannot meet its obligation to repay the deposits with interest.

The bank will fail if it cannot repay the deposits given to it, including any interest r_D. With the bank lending to all N companies, and incurring the resulting monitoring costs, each being lent the amount of ML which has also been received as deposit, this condition becomes

$$\pi (1 + r_L) ML - NC < (1 + r_D) ML. \tag{3.28}$$

Depositors will either receive back their deposit L with interest at maturity, or in the case of default of the bank, the value of the loans that are repaid net of the monitoring costs of the bank that have already been spent. In order to attract depositors, the amount they obtain from the bank must be at least as much as they obtain from direct lending as shown on the left-hand side of equation (3.27), representing the expected payoff for each of the N loans. This then gives us

$$\begin{aligned} &\frac{1}{M} \min \{\pi (1 + r_L) ML - NC, (1 + r_D) ML\} \\ &\quad \geq \pi (1 + r_C) L - NC. \end{aligned} \tag{3.29}$$

The first term indicates the total amount of loans, net of monitoring costs, repaid to the bank and the second term the amount due to depositors. Depositors will at most be rapid their deposits with interest, or if the bank does not have sufficient funds, the funds the bank has, which are the loans repaid to it. These funds are available to the M depositors, each receiving an equal share. The right hand side represents the expected repayments from direct lending for a lender from all N loans it has given.

Let us now assume that the bank does not fail and hence equation (3.28) is not fulfilled. In this case equation (3.29) reduces to $(1 + r_D) L \geq \pi (1 + r_C) L - NC$ as depositors will be repaid their deposits with interest. If banks have market power and can extract all surplus from competitive depositors, then they would not pay deposit rates higher than necessary to fulfill this condition, which would hold with equality. We thus obtain that

$$1 + r_D = \pi (1 + r_C) - \frac{NC}{L}. \tag{3.30}$$

Let us now assume that direct and bank lending are competitive such that the loan rates are identical, $r_C = r_L$. Inserting this solution into equation (3.28), we obtain that for a bank to fail we require $\pi (1 + r_L) ML - NC < \pi (1 + r_L) ML - NMC$, which cannot be fulfilled if $M > 1$. This implies that the bank can never fail as the condition for a bank failure in equation (3.28) is never fulfilled. The total monitoring costs with bank lending are NC arising from the N loans given by the bank, one for each borrower. These costs are lower than the monitoring costs from direct lending, which were NMC. Thus, bank lending is more efficient and there are no additional costs to depositors as a bank can never fail.

Summary If a bank exists, competitive depositors receive the same outcome as if they were to lend directly to the companies. The bank would set deposit rates such that it would never fail. Given that not all M lenders need to monitor each of the N companies, but only the bank, the monitoring costs are reduced from NMC to NC and this difference would be the bank profit if we assume that companies are also competing for loans and would thus not benefit from lower loan rates. The reduced monitoring costs increase the social welfare of the economy and would allow the bank to make a profit without adversely affecting companies or depositors. Alternatively, these benefits may be shared with depositors by banks paying higher deposit rates as long as equation (3.28) is violated. We thus require that $\pi(1+r_L) - \frac{NC}{ML} > 1 + r_D \geq \pi(1+r_L) - \frac{NC}{L}$, i. e. the deposit rate is not too high. Similarly companies could obtain a share of the benefits through lower loan rates, again as long as the previous condition is fulfilled.

There are no additional costs of losses from a bank failing or monitoring by depositors of the bank that would reduce these benefits, because the bank cannot fail and this makes monitoring unnecessary. Thus banks reduce the monitoring costs without increasing costs from its possible bankruptcy and the overall welfare is increased. The origin of this result is that monitoring efforts are not duplicated across M individual lenders for each loan but only incurred once by the bank. If banks face lower costs of monitoring, arising from expertise, experience or economies of scale, this advantage of banks is even more pronounced and the range of deposit rates supported would be wider.

Reading Diamond (1984)

3.2.2 Optimal monitoring duplication

Through monitoring their borrowers, banks can increase the quality of the loans they provide. Banks can not only gather additional information on their borrowers, but also through appropriate intervention and advice improve the chances of such loans being repaid. They might achieve this by providing advice to the management of the company or by reducing the risk of company funds being invested different from the agreement made when taking out the loan. Of course, monitoring will normally be imperfect and the reduction in the risk to the bank will not always materialise. Having multiple banks conducting such monitoring will increase the probability of monitoring being successful in reducing loan risk. However, this duplication of monitoring efforts is costly and the additional costs of monitoring by multiple banks will have to be balanced against the benefits of the reduced credit risk.

Companies make investments using loans L that are succeeding with probability π_i, allowing the loan and its interest r_L to be repaid. This probability of success can take two values, either π_H or π_L, where $\pi_H > \pi_L$. Banks can induce companies to increase their efforts such that the probability of success is high, π_H, is realised more often. Let us assume that the probability of the investment with high success chances being realised, p, can be influenced by monitoring efforts of banks. Such monitoring by banks is, however, costly to them and the costs c increase in the size of the loan L and marginal costs are increasing in the probability p.

Companies are only able to repay their loan if the investment is successful and hence the expected probability of repayment is given by $p\pi_H + (1-p)\,\pi_L$. Thus bank profits if there is only a single bank monitoring, are given by

$$\Pi_B = (p\pi_H + (1-p)\,\pi_L)\,(1+r_L)\,L - (1-r_D)\,D - \frac{1}{2}cLp^2, \tag{3.31}$$

with r_D denoting the deposit rate and loans fully financed by deposits such that deposits D equal the amount of loans given, $D = L$. The optimal monitoring level we get from maximizing bank profits such that $\frac{\partial \Pi_B}{\partial p} = 0$, which solves for

$$p^* = \frac{(\pi_H - \pi_L)\,(1+r_L)}{c}. \tag{3.32}$$

Having established the optimal amount of monitoring, as measured by the probability of the investment with the high success rate being realised, if a single bank provides a loan, we can now continue to compare this result with a case in which multiple banks provide a loan and each bank monitors the company.

Let us now assume there are N banks providing loans and hence monitoring the company. These monitoring efforts are substitutes as it is sufficient for one bank to successfully monitor the company and thus ensure the investment with the high success rate is chosen. For monitoring to fail and companies using the low success rate π_L, we need to assume that all banks fail in their monitoring efforts. Each bank fails with probability $1-p_j$, thus all banks fail with probability $\prod_{j=1}^{N}\left(1-p_j\right)$, assuming that the success of monitoring is independent across banks. Thus the probability of success in monitoring is given by

$$\hat{p} = 1 - \prod_{j=1}^{N}\left(1-p_j\right). \tag{3.33}$$

We can now rewrite equation (3.31) for the profits of a single bank as

$$\hat{\Pi}_B = (\hat{p}\pi_H + (1-\hat{p})\,\pi_L)\,(1+r_L)\,\frac{L}{N} - (1+r_D)\,\frac{L}{N} - \frac{1}{c}c\frac{L}{N}p_i^2, \tag{3.34}$$

where we assume that the aggregate lending is identical to the case of a single bank providing the loan and each bank lends the same amount, in this case $\frac{L}{N}$. We now get the optimal monitoring effort from evaluating

$$\frac{\partial \hat{\Pi}_B}{\partial p_i} = \prod_{j=1, j\neq i}^{N}\left(1-p_j\right)(\pi_H - \pi_L)\,(1+r_L)\,\frac{L}{N} - \frac{cL}{N}p_i = 0, \tag{3.35}$$

where we have used equation (3.33) to replace $\hat{p}$. This expression easily solves for

$$(1 - p_i)^{N-1} (\pi_H - \pi_L) (1 + r_L) - cp_i = 0. \tag{3.36}$$

Here we used that $p_j = p_i$ as all banks are identical and we only consider symmetric equilibria. If p_i^*, the solution to equation (3.36) were equal to p^* from equation (3.32), then inserting for cp^* from equation (3.32), we would require that $(\pi_H - \pi_L) (1 + r_L) \left(\left(1 - p_i^*\right)^{N-1} - 1 \right) = 0$. However, the left-hand side is clearly negative unless $p_i^* = p^* = 0$, which due to equation (3.32) can be excluded and hence we find that $p_i^* < p^*$ for the optimal monitoring efforts of each individual bank. This implies that each bank monitors less and hence faces less monitoring costs, compared to a situation where there is only a single lender. Each bank has a positive external effect on the profits of other banks through the increase in the likelihood of the high success rate π_H being realised, while fully internalizing their monitoring costs. This causes their monitoring efforts to be reduced and seeking to benefit from the monitoring of other banks, a classical moral hazard situation for banks.

Solving (3.36) for $\left(1 - p_i^*\right)^{N-1}$ and multiplying by $1 - p_i^*$, we easily get the aggregate monitoring effort of all banks jointly as

$$\hat{p}^* = 1 - \left(1 - p_i^*\right)^N = 1 - \frac{cp_i^* \left(1 - p_i^*\right)}{(\pi_H - \pi_L) (1 + r_L)}. \tag{3.37}$$

Comparing the expressions for $\hat{p}^*$ and p^* in equations (3.37) and (3.32), we find that $\hat{p}^* \geq p^*$ if

$$\left((\pi_H - \pi_L) (1 + r_L) - \frac{1}{2} c \right)^2 \leq c^2 \left(\frac{1}{4} - p_i^* \left(1 - p_i^*\right) \right). \tag{3.38}$$

If this inequality is fulfilled, the aggregate monitoring of multiple banks will exceed the monitoring of a single bank. We thus observe that the aggregate monitoring efforts with N banks providing loans, is higher if on the one hand monitoring costs c are sufficiently high and p_i^* is sufficiently far away from $\frac{1}{2}$ to ensure the final expression on the right-hand side is not too small. High monitoring costs make monitoring by a single bank costly and the bank will only provide a low level of monitoring due to high marginal costs. In this case multiple banks monitoring the company will reduce monitoring by each bank and reduce their monitoring costs substantially; however, as all banks are monitoring, the reduction in individual monitoring is not sufficient to reduce the overall monitoring effort. If monitoring costs c are low, the lower marginal costs of monitoring will lead to a more pronounced reduction of monitoring if this is shared by multiple banks compared to a single bank. On the other hand, if the benefits of monitoring are small, thus π_H is close to π_L, monitoring by multiple banks will reduce the overall monitoring effort as banks will rely more on each other's monitoring efforts and will contribute little of their own monitoring efforts due to the limited benefits of doing so.

Summary Duplication of monitoring effort can be beneficial and increase the quality of loans if monitoring costs for each bank are high or the benefits in reducing lending risks through monitoring are substantial. By multiple banks monitoring a company, each bank can reduce their individual monitoring activity and save costs, but at the same time the reduction in monitoring is not such that the overall monitoring effort of all banks jointly reduces. Hence multiple lenders are beneficial for companies that are either difficult to monitor and hence monitoring is costly, for example due to the complexity of their business or a lack of expertise by the bank in the business of the company. Similarly, multiple lenders are beneficial where benefits of monitoring are substantial, such as companies that have significant discretion on the use of their funds or where the impact of bank providing advice to the company is particularly strong. Lending by multiple banks can be beneficial to reduce the loan risk banks face due to increased aggregate monitoring.

3.2.3 Monitoring advantage by banks

The existence of banks does not only reduce monitoring costs, it also allows to expand lending to a wider range of companies compared to direct lending. Banks achieve this by monitoring companies and thereby ensuring that more suitable investments are conducted. Let us assume a company can choose freely between two investments, one which succeeds with probability π_H and yields a return of R_H, if successful, and the other succeeds with probability π_L and yields a return of R_L, if successful. In both cases an unsuccessful project yields no revenue and companies finance their investments fully with a loan of size L. We assume that $\pi_H > \pi_L$, such that investment H is less risky, but with $R_L > R_H$ it yields a lower return in case of success. We furthermore assume that $\pi_H (1 + R_H) > 1 + r_L > \pi_L (1 + R_L)$, with r_L denoting the loan rate, implying that the loan cannot be repaid on average for the more risky investment, while it is possible to do so for the safer investment. Therefore, a lender would not provide a loan to a company that chooses the risky investment L as the expected repayment is lower than the agreed repayment, while the safe investment H generates sufficient expected revenues that would allow the company to repay the loan.

Direct lending Let us firstly assume that the company obtains a loan directly from lenders who cannot monitor the company. With the interest on the loan denoted r_C, the company will choose the low-risk project H if its expected profits, Π_C^H, after repaying the loan and interest, is higher than the profits from the high-risk investment L, Π_C^L. The company having limited liability and no other assets, will only obtain these profits if the investment is successful, which happens with probabilities of π_H and π_L, respectively. Thus their profits need to fulfill the condition that

$$\begin{aligned} \Pi_C^H &= \pi_H \left((1 + R_H) L - (1 + r_C) L \right) \\ &\geq \pi_L \left((1 + R_L) L - (1 + r_C) L \right) = \Pi_C^L, \end{aligned} \tag{3.39}$$

which solves for

$$1 + r_C \leq 1 + r_C^* = \frac{\pi_H \left(1 + R_H\right) - \pi_L \left(1 + R_L\right)}{\pi_H - \pi_L}. \tag{3.40}$$

As lenders do not know which project the company chooses, they will only lend if this condition is fulfilled to ensure that the low-risk project is chosen. This is because with the high-risk project lenders cannot expect to make a profit as we assumed that $\pi_L \left(1 + R_L\right) < 1 + r_C$. Alternatively to lending directly to the company, lenders could instead invest into deposits at a bank that pay a deposit rate of r_D. In order to lend directly, lenders need to obtain an at least equal return to that offered by deposits, hence we require that

$$\pi_H \left(1 + r_C\right) L \geq \left(1 + r_D\right) L \tag{3.41}$$

which we can combine with equation (3.40) to obtain

$$\pi_H \left(1 + r_C^*\right) \geq \pi_H \left(1 + r_C\right) \geq 1 + r_D. \tag{3.42}$$

Using the first and last relationship, we obtain that the probability of success for the low-risk investment has to fulfill

$$\pi_H \geq \pi_H^* = \frac{1 + r_D}{1 + r_C^*}. \tag{3.43}$$

Hence only companies whose low-risk investment is sufficiently safe will obtain a loan from direct lenders.

Bank lending Assume now that a bank can monitor companies at cost C to ensure they always choose the low-risk project H, which direct lenders are unable to do. The bank's profits are then given by

$$\Pi_B^H = \pi_H \left(1 + r_L\right) L - \left(1 + r_D\right) L - C, \tag{3.44}$$

taking into account the interest paid on deposits used to fund the loan. We also assume here that banks cannot fail and will always repay their deposits, which implies they have unlimited liability. This allows us to avoid the complication of depositors facing the risk of not being repaid their deposits. Banks need to be profitable, thus $\Pi_B^H \geq 0$ and solving this equation, we have

$$1 + r_L \geq 1 + r_L^* = \frac{\left(1 + r_D\right) L + C}{\pi_H L} \tag{3.45}$$

with the constraint that $\Pi_C^H \geq 0$, or $1 + r_L \leq 1 + R_H$ from equation (3.39). Using this constraint in equation (3.45), we easily obtain that this implies that

$$\pi_H \geq \pi_H^{**} = \frac{\left(1 + r_D\right) L + C}{\left(1 + R_H\right) L}. \tag{3.46}$$

If the low-risk investment H is sufficiently safe, the bank will provide a loan to the company. The higher the monitoring costs and deposit rates are, the higher this threshold is.

Comparing direct and bank lending The threshold for providing a loan in direct and bank lending are given by equations (3.43) and (3.46), respectively. We can now see that the threshold in the success rate of investments for direct lending is higher, $\pi_H^* > \pi_H^{**}$, if $1 + r_C^* < \frac{(1+r_D)(1+R_H)L}{(1+r_D)L+C}$. Inserting for $1 + r_C^*$ from equation (3.40), this becomes

$$C < \frac{\pi_L (R_H - R_L)(1 + r_D)}{\pi_H (1 + R_H) - \pi_L (1 + R_L)} L. \tag{3.47}$$

Therefore, assuming that the monitoring costs are not too high, bank lending will allow the provision of loans to more risky companies. If the difference in expected returns between low-risk investments, $\pi_H (1 + R_H)$, and high risk investments, $\pi_L (1 + R_L)$, is large, monitoring costs have to be small for bank lending to extend the range of loans given. Similarly if the difference in the actual returns between the low-risk and high-risk investments, R_H and R_L, are low, only small monitoring costs can be accommodated, as is the case if the high risk investment is unlikely to succeed (low π_L). In all of these cases the reason is that the difference for the company between the low-risk and the high-risk investment does make the choice of the low-risk investment more attractive, even in the absence of monitoring the threshold $1 + r_C^*$ will be quite high and hence the benefits of monitoring are small, allowing only for low costs to be beneficial to bank lending, which incurs the additional costs C from monitoring that direct lending does not have to bear.

If banks and direct lenders set the highest possible loans rate $r_L = R_H$ and the highest loan rate for direct lending is chosen, r_C^*, we can easily see that $r_C \geq r_L$ and direct borrowing is always cheaper for companies. Therefore, companies that are low risk with $\pi \in \left[\pi_H^*; 1\right]$ will choose direct lending, while more risky companies with $\pi \in \left[\pi_H^{**}; \pi_H^*\right]$ will seek bank loans, because direct lending is not feasible. If the risk is even higher such that $\pi < \pi_H^{**}$, no loan can be obtained.

Summary If banks can monitor companies with sufficiently low costs to ensure they select low-risk investments, they are able to provide loans to companies whose low-risk investments have a higher risk than would be possible in direct lending. While there are additional costs to the bank, that through the loan rate are charged to companies, banks do not have to rely on incentive constraints in the loan rate to ensure companies choose low-risk investments. This allows banks to charge loan rates that provide incentives to choose high-risk investments, but companies are prevented from doing so through the monitoring of banks. Hence in these cases, direct lending would seize as the expected returns of the high-risk investments are not sufficient to provide the return lender requires, but banks would still find the loan profitable as it ensures the low-risk investment is chosen.

Banks allow the financing of companies that have higher risks compared to direct lending if their monitoring costs are sufficiently small. For higher monitoring costs, the need to recover these through the loan rate may well make direct lending, and a reliance on incentives to choose the low-risk investment, more attractive, e.g. in industries that are unfamiliar to banks or where high-risk investments could easily be misrepresented as low-risk. Alternatively, banks might forego monitoring in these cases and provide loans by relying on the same incentive constraints as direct lenders. Overall, bank lending extends the scope of lending and allows to finance investments that are too risky for direct lending. This will allow economies to provide funding for more innovative investments that will ultimately benefit economic growth in the economy.

Reading Keiding (2016a, ch. 1.3.2) or Freixas & Rochet (2008a, Ch. 2.5.1)

Résumé

Comparing the outcomes from direct lending and bank lending has shown that if monitoring is required to ensure they are choosing low-risk investments for a positive expected return to the lender, bank lending has some distinct advantages. Firstly, it reduces the amount of monitoring necessary if the loan required by the company is not provided by a single direct lender but requires multiple smaller loans. This duplication of monitoring efforts by direct lenders, ignoring any moral hazard of direct lending from relying on other borrowers and their efforts, compared to the monitoring by a single bank, gives the bank a distinct advantage. This is before even considering the free-riding problem in monitoring as direct lenders might not conduct monitoring adequately and instead seek to rely on the efforts of other direct lenders in order to save on costs, which might result in less than optimal monitoring and could therefore increase the risks to direct lenders. If bank lending is chosen instead, the direct lender would deposit their funds with the bank who then provides the loan. This might expose the now depositors to the risk of the bank failing and them losing their deposits. However, such a scenario was ruled out as the bank will set interest rates such that the failures of individual loans are covered by the profits of those loans that are repaid and overall the bank will always repay the deposits and depositors face no additional risks. This leaves the overall costs from monitoring to be reduced, but no additional costs being imposed.

While banks might have advantages in monitoring companies, it nevertheless imposes costs on banks. Having a loan funded by multiple banks and each bank monitoring the company, allows the banks to reduce their individual monitoring efforts and thereby reducing monitoring costs, while the overall monitoring efforts of all banks aggregated may increase, reducing the risks of the loan. This effect can be observed if monitoring is very costs and the benefits of monitoring in terms of reducing the lending risk are substantial. Despite monitoring having positive external effects on other banks, the reduction in individual monitoring effort is not sufficient in these cases to eliminate the benefits of multiple monitors.

It is difficult to argue that direct lenders would always be able and willing to monitor loans, but it is reasonable to propose that banks can do so in a wider range of scenarios. Companies are often faced with a wide variety of investment opportunities, some more risky but also more profitable than others, and it seems reasonable to suggest that once the loan has been provided, companies can change the investments they actually make, compared to what had initially been proposed. The incentive might well be to choose a more risky investment which provides a high profit if successful, while the losses in case of it being unsuccessful, are restricted due to limited liability. As the lender does not share the higher profits but only receives a fixed interest, this more risky investment only increases the risks to the lender, which might make the loan in these circumstances not profitable. By choosing loan rates that are not too high, the lender might provide incentives that will ensure the company chooses a low-risk investment, even without monitoring. If we introduce a bank that can monitor lenders and whose monitoring ensures that the company will always choose the low-risk investment rather than rely on incentives, they might be able to provide loans to companies that are overall riskier than a direct lender relying on incentives would find profitable. The reason is that while a bank would have to recover their costs of monitoring from the company through a higher loan rate, it does not rely on incentives to ensure the low-risk investment is chosen. The higher loan rate might give an incentive to a company to choose high-risk investments, but due to monitoring it cannot make this choice. The result is that as long as the low-risk investment is profitable, a loan will be provided, even if the incentives to choose the high-risk investment would in direct lending cause a switch of investments such that the loan would no longer be profitable to the lender.

Overall, banks can be beneficial in that on the one hand they reduce monitoring costs without introducing additional costs in the form of possible bank failures and they can extend the scope of loans provided to more risky companies if we assume that direct lenders are not able to monitor loans effectively. The latter benefit may well allow for more innovative companies to be financed, whose success often is less certain but who might benefit future economic growth through the innovations they introduce.

3.3 Diversification

Banks can be viewed as intermediaries to facilitate the provision of loans by making the matching and lenders and borrowers more efficient. In a different view, using banks can make the monitoring of borrowers more effective. While the first aspect would be beneficial to borrowers and lenders alike, the second aspect would mainly benefit lenders (depositors). Alternatively, we might view banks as an association of borrowers that use banks to reduce their loan costs. Some banks have originally been set up as a means of providing loans to specific groups of borrowers, such as house buyers, small local businesses, farmers, or members of disadvantaged groups. Here we will assess the benefits banks might give to lenders, relative to direct lending, from their ability to pool deposits and provide loans to a large number of companies.

Direct lending Let us consider an individual lender providing a loan to a borrower (company) directly. Investments made by the company succeed with probability π

and it is only then that the company is able to repay the loan, and if the investments are not successful, they provide no funds such that the company cannot repay the loan. The loan rate is r_C on a loan of size L. The variance of the outcomes to the lender is given by $\pi(1-\pi)$ for each unit of final outcome. If we assume the lender to be risk averse with absolute risk aversion z, the expected utility of this lender is given by $u\left(\pi(1+r_C)L - \frac{1}{2}z\pi(1-\pi)(1+r_C)^2 L^2\right)$, where $u(\cdot)$ denotes the utility function. If the lender would not provide the loan it would retain its investment L in the form of cash, giving utility $u(L)$. To provide the loan, the direct lender's utility from lending must exceed that of not providing the loan and comparing coefficients from the utility function, gives us the condition that

$$1 + r_C \geq \frac{1}{\pi} + \frac{z(1-\pi)(1+r_C)^2 L}{2}. \tag{3.48}$$

This solves for a minimum direct loan rate of

$$1 + r_C \geq 1 + r_C^* = \frac{1}{z(1-\pi)L} - \sqrt{\frac{1}{z^2(1-\pi)^2 L^2} - \frac{2}{z\pi(1-\pi)L}}. \tag{3.49}$$

Hence direct lending would occur at loan rates that are at least r_C^*. Competition between direct lenders would ensure the loan rate to equal this value.

Bank lending Banks collect deposits and lend these to multiple borrowers. Let us assume there are N depositors, each depositing the amount of L, which is then provided as loans to N borrowers. If we further assume that the repayments of the loans are independent of each other, the variance of the portfolio of N loans is $\frac{\pi(1-\pi)}{N^2}$, which arises from taking the variance of a portfolio of N independent assets of equal weight, i. e. $\frac{1}{N}$, each having variance $\pi(1-\pi)$. Then the utility from lending N loans is $u\left(\pi(1+r_L)NL - (1+r_D)NL - \frac{1}{2N^2}z\pi(1-\pi)(1+r_L)^2 N^2 L^2\right)$, where r_D denotes the deposit rate that is paid to depositors. If not lending, the bank conducts no business and if it takes on deposits it has to repay these with interest while not obtaining any interest from lending, making such a business unprofitable; this gives them then a utility of $u(0)$. We thus obtain that bank lending will be profitable for the bank if the utility of lending exceeds that of not lending, which results in

$$1 + r_L \geq \frac{1+r_D}{\pi} + \frac{z(1-\pi)(1+r_L)^2 L}{2N}, \tag{3.50}$$

solving for the minimum bank loan rate

$$1 + r_L \geq 1 + r_L^* = \frac{N}{z(1-\pi)L} - \sqrt{\frac{N^2}{z^2(1-\pi)^2 L^2} - \frac{2N(1+r_D)}{z\pi(1-\pi)L}}. \tag{3.51}$$

Therefore, bank lending would occur at loan rates that are at least r_L^*. If banks are fully competitive, the loan rate will be equal to r_L^*. We can now compare the costs to companies of borrowing between direct lending and bank lending.

Comparing direct and bank lending Let us assume that competition between direct lenders or between banks requires either to charge the minimum loan rates, i. e. r_C^* and r_L^*, respectively. We can easily show that $\frac{\partial(1+r_L^*)}{\partial N} < 0$ and hence the more loans a bank provides, the lower the loan rate of banks will become. Furthermore for $N = 1$ we have $1+r_L^* > 1+r_C^*$ if $r_D > 0$ and the increased costs from paying interest on deposits makes bank lending more expensive. As the number of loans N increase, the loan rate for banks reduces and as we have $\lim_{N\to\infty} 1 + r_L^* = \frac{1+r_D}{z\pi(1-\pi)L} \approx 0$, there will be a N^* at which $r_C^* = r_L^*$ for a sufficiently large L. A further increase in N beyond N^* implies that bank loans are offered at a lower rate than direct loans. Therefore, if banks are big enough, that is give a sufficient number of loans, they can offer better loan rates than direct lenders, while paying interest on deposits. The reason for this result is that despite the higher costs of paying interest on deposits, banks are able to benefit from diversification. By providing a large number of loans, the risk of their loan portfolio reduces sufficiently to compensate for these increased costs. The risk aversion of direct and bank lenders increases the loan rate required to compensate them for taking on the default risk. With the risk reducing due to diversification from larger loan portfolios, this compensation becomes ever smaller, until it has reduced so far that it outweighs the increased costs from the interest on deposits.

Of course, for banks to be successfully introduced, they need to attract deposits that they can be lent out. Therefore, lenders need to deposit their funds L with the bank rather than lending directly. With a bank deposit, assuming the bank cannot fail, they obtain a utility of $u\left(\left(1+r_D\right)L\right)$ and comparing this with the utility of direct lending, $u\left(\pi\left(1+r_C\right)L - \frac{1}{2}z\pi\left(1-\pi\right)\left(1+r_C\right)^2 L^2\right)$, we get from comparing coefficients that

$$1 + r_D \geq \pi\left(1+r_C\right) - \frac{1}{2}z\pi\left(1-\pi\right)\left(1+r_C\right)^2 L. \tag{3.52}$$

Thus deposit rates need to be sufficiently high to compensate the direct lender for the lost revenue from providing a loan; due to the risk involve din lending, the deposit rate will be lower than the expected return from direct lending. From equation (3.50), the condition for bank lending to occur can be rewritten as

$$1 + r_D \leq \pi\left(1+r_L\right) - \frac{1}{2N}z\pi\left(1-\pi\right)\left(1+r_L\right)^2 L. \tag{3.53}$$

Combing these two inequalities we easily see that we require a deposit rate that fulfills

$$\pi\left(1+r_C\right)-\frac{1}{2}z\pi\left(1-\pi\right)\left(1+r_C\right)^2L\leq 1+r_D \qquad (3.54)$$
$$\leq\pi\left(1+r_L\right)-\frac{1}{2N}z\pi\left(1-\pi\right)\left(1+r_L\right)^2L.$$

If we now insert for r_C the value of r_C^* of equation (3.49) and for r_L the value of r_L^* from equation (3.51) by assuming that both bank and direct lending are competitive, we see that the first and last term in this equality are identical and hence it must be fulfilled with equality and r_D can be easily determined. Hence with an adequately determined deposit rate, banks can be sustained. In this case borrowers benefit from a lower lending rate and depositors (lenders) from a deposit rate that provides them with a utility level identical that what they would have obtained from direct lending.

Summary Bank lending has the advantage that the provision of a large number of loans reduces the risk of lending in this portfolio of loans due to diversification. This requires a smaller risk premium on the loan rate for risk-averse banks, from which borrowers benefit as the bank can charge a lower loan rate. Even if the bank has higher costs than individual lenders due to having to pay interest on deposits, the diversification benefits from a sufficiently large number of bank loans would outweigh these costs. Depositors will be equally well off compared to direct lending in a competitive environment, as they are compensated by the interest paid on deposits for the returns they are not obtaining from direct lending. It is therefore, that depositors are equally well off and borrowers are better off with bank lending compared to direct lending.

Reading Leland & Pyle (1977)

Conclusions

Banks can increase the efficiency of lending activities. They reduce transaction costs of providing loans in various ways. Loans need to be monitored to ensure the borrower does adhere to the terms of the loan contract and in particular does not engage in activities that jeopardise the repayment of the loan, for example the investment into more risky projects. By bundling the many smaller loans individual lenders could give, banks can reduce these monitoring costs significantly as fewer monitoring activities by required by banks than the larger number of individual lenders. With monitoring costs commonly not dependent on the size of the loan, this would result in a significant reduction of transaction costs when using banks compared to direct lending, leading to an increase in welfare. For small loans, monitoring would not be cost-effective and would potentially not be undertaken at all by direct lenders, while the larger size of bank loans would often result in monitoring that is beneficial.

The effect of monitoring is that borrowers cannot easily seek out more risky, and for them more profitable, investment opportunities that increase the risk to their lenders. Without monitoring, lenders have to rely on incentives to ensure that borrowers do not increase the risk of their investments, while banks can ensure this through the monitoring process. The consequence is, that banks do not have to rely on incentives for companies to limit risks, increasing the range of companies that can be given loans. Companies whose incentives are such that they would choose higher risk investments cannot obtain direct loans, but might still be able to obtain banks loans given the ability of banks to monitor them. Increasing the range of companies being granted loans would increase investment and hence lead to a higher growth rate of the economy, especially if the affected companies are innovative and operating in more risky high-growth industries. Despite the higher costs banks face due to the monitoring, this would still be beneficial to companies who otherwise would not be able to pursue their investments.

If banks have the advantage of having lower costs of negotiation of the loan in the first place, then the deposit rate banks pay will be lower than the loan rate in direct lending and the bank loan rate will be higher than the loan rate in direct lending. This, looking only at the costs of borrowing or lending makes banks less attractive than direct lending. However, the reduced costs of negotiating loans must be balanced against these higher costs. If banks are negotiating fairly with their customers, depositors and borrowers, they would share the saved costs and overall all participants are better off using a bank. Only in rare instances would direct lending be preferred, mainly if the negotiation costs are low and the return on investment to the company, and thereby the upper ceiling of the loan rate, sufficiently high; in this case the lower costs cannot accommodate sufficient profits for all market participants, lender (depositor), borrower, and bank, making relying on direct lending necessary.

By providing a large number of loans to many companies, banks diversify their risks, while direct lenders will be limited in this diversification. If banks and direct lenders are risk averse, banks can charge a lower loan rate than direct lenders could due to the lower risks banks face. This is despite banks having higher costs as they have to pay interest on deposits such that direct lenders are attracted to providing their funds to banks. These higher costs can be offset with the reduced lending risk arising from diversification. If the loan portfolio is sufficiently diversified, then banks can provide loans at lower interest rates than direct lenders, making borrowers better off without negatively affecting direct lenders.

Overall, banks offer a mechanism to reduce transaction costs and the effect is either a direct benefit in the form of lower loan rates and/or higher deposit rates, arising from reduced monitoring costs. There are also other positive effects, such as the extension of the range of companies that can obtain loans. In both cases, banks will increase the welfare in the economy, making their existence desirable.

Chapter 4

Liquidity provision

Investments by companies are in many cases, if not most, long-term and if loans need to be repaid early, this might cause significant disruption to the company. However, those providing the loans for such investments, would in most cases prefer to be able to withdraw funding if the need arises, i. e. they prefer liquidity. A solution to these incompatible interests of borrowers and lenders would be to establish a mechanism that would allow lenders to withdraw their funds while at the same time allowing lenders to retain loans on long-term basis. This transformation of short-term deposits into long-term loans is a key benefit of banks and often referred to as 'liquidity insurance'. How banks achieve this transformation, and how it is superior to other mechanisms, is discussed in chapter 4.1, with alternative banking specifications explored in chapter 4.2. It is not only that banks provide this liquidity for the benefit of their depositors, but, as we will see in chapter 4.3, bank are willing to accept such short-term deposits, is may even be cheaper for banks to do so.

But banks do not only provide liquidity to depositors, they also allow borrowers (companies) access to liquidity by standing ready to provide short-term loans if they face an unexpected requirement for additional funds. The existence of credit lines to companies on which they can draw is another feature of banks and explored in chapter 4.4.

It is thus that banks increase the welfare of their customers not only by reducing transaction costs as discussed in chapter 3, but in addition allow them access to funds if and as they need, while at the same time giving borrowers the stability of finance they seek. In this sense, banks are uniquely placed to breach the gap between the preferences of borrowers and depositors.

4.1 Maturity transformation of deposits

In general, depositors do not know in advance when they need access to cash and therefore would like the ability to withdraw from any investments they made. Whether they seek to withdraw funds from an investment might depend on a number of exogenous factors, such as consumption possibilities, alternative investment opportunities, or liquidity shocks. We assume that total deposits D can be withdrawn either in time period 1 or in time period 2, but not in both. Let us assume that depositors withdraw in time period 1 with probability p and obtain interest r_D^1, and otherwise obtain interest r_D^2 in time period 2. Thus, a fraction p of deposits D, worth $p\left(1+r_D^2\right)D$ is withdrawn in time period 1 and a fraction $1-p$ of the deposits D, worth $(1-p)\left(1+r_D^2\right)D$, remain invested and are repaid in time period 2. In time period 0, banks or individuals can invest an amount $0 \leq L \leq D$ into a loan with loan rate r_L, that is repaid in time period 2 with probability π and is not repaid with probability $1-\pi$. Holding cash does not attract any interest. The loan can be liquidated at some cost in time period 1, such that only a fraction $0 \leq \lambda < 1$ of the initial loan L is realized.

With utility function $u(\cdot)$, the expected utility of depositors is then given by

$$E\left[U(D)\right] = pu\left(\left(1+r_D^1\right)D\right) + (1-p)\,u\left(\left(1+r_D^2\right)D\right), \tag{4.1}$$

neglecting discounting between time periods.

We can now compare the utility depositors obtain from different arrangements. We consider direct lending by the 'depositors', with and without the possibility to trade loans made in time period 1, and bank lending. Comparing these cases with the social optimum, we can establish which arrangement is the best alternative.

Social optimum The expected withdrawals of the depositors in time period 1, $p\left(1+r_D^1\right)D$, would share the available cash, $D-L$, and thereby avoid the costly liquidation of any loans while not leaving any cash unused. The expected deposits of the remaining depositors, $(1-p)\left(1+r_D^2\right)D$, would obtain the proceeds of the loan to be distributed in time period 2, $\pi(1+r_L)L$. Thus we find that

$$\begin{aligned} p\left(1+r_D^1\right)D &= D-L, \\ (1-p)\left(1+r_D^2\right)D &= \pi(1+r_L)L, \end{aligned} \tag{4.2}$$

which can be combined by eliminating L and dividing by D as

$$p\left(1+r_D^1\right) + (1-p)\frac{\left(1+r_D^2\right)}{\pi(1+r_L)} = 1. \tag{4.3}$$

Depositors will maximize their expected utility in equation (4.1), subject to constraint (4.3), which can be solved for $1+r_D^2 = \frac{\pi(1+r_L)\left(1-p\left(1+r_D^1\right)\right)}{1-p}$ and inserted into equation (4.1). This allows us to determine the optimal deposit rate by maximizing the amount paid out to depositors in time period 1, giving rise to the first order condition

$$\begin{aligned}\frac{E\left[U\left(D\right)\right]}{\partial\left(1+r_D^1\right)D} &= p\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} \\ &\quad + \left(1-p\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}\left(-\frac{p}{1-p}\pi\left(1+r_L\right)\right) \\ &= 0,\end{aligned} \tag{4.4}$$

which solves for

$$\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} = \pi\left(1+r_L\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}, \tag{4.5}$$

implying that the marginal rate of substitution equals the expected return on the loan. Knowing the utility function $u\left(\cdot\right)$, we could the solve explicitly for the optimal deposit rates. This result on the social optimum serves as a benchmark to analyse the subsequent cases of direct and bank lending.

Direct lending If an individual provides a loan directly and seeks to withdraw its funds in time period 1, he will have to liquidate his loan at a loss, thus the deposits returned in time period 1 are

$$\left(1+r_D^1\right)D = D - L + \lambda L = D - \left(1-\lambda\right)L \leq D \tag{4.6}$$

where $D - L$ represents the cash held and λL the realization of the liquidated loan. The total amount available to such an individual will be less than the initial deposit unless $L = 0$ or $\lambda = 1$. We can interpret r_D^1 as the return on investment for these individuals.

Those individuals not liquidating their loan, realize the return it generates in time period 2 and hence

$$\begin{aligned}\left(1+r_D^2\right)D &= D - L + \pi\left(1+r_L\right)L \\ &= D - \left(1-\pi\left(1+r_L\right)\right)L \\ &\leq \pi\left(1+r_L\right)D,\end{aligned} \tag{4.7}$$

where the inequality arises for all $L \leq D$, as can easily be verified. The payment in time period 2 consists of the cash retained, $D - L$ and the repayment of the loan given, $\pi\left(1+r_L\right)L$. As before, r_D^2 can be interpreted as the return on investment for those not withdrawing funds.

The objective function of maximizing expected utility as defined in equation (4.1) remains unchanged. Using the inequalities in equations (4.6) and (4.7), we can now obtain the constraint to our optimization as

$$p\left(1+r_D^1\right) + \left(1-p\right)\frac{\left(1+r_D^2\right)}{\pi\left(1+r_L\right)} \leq 1 \tag{4.8}$$

which is more stringent than (4.3) in the social optimum. The inequality here is strict if either $\lambda < 1$ or $D < L$ and as we assumed $\lambda < 1$ to impose a cost of liquidating loans, this inequality will be strict. With a binding and more restrictive constraint, the resulting optimal solution in the case of direct lending will in egneral be inferior to that of the social optimum.

Direct lending with trading Rather than liquidating their loans in the event of withdrawing funds, individuals could sell their loans to those not wanting to withdraw their funds. The price obtained, P, will be quoted relatively to the expected value of the loan in time period 2, which is $\pi(1 + r_L)$. Hence, an individual withdrawing funds would obtain

$$\left(1 + r_D^1\right) D = D - L + \pi \left(1 + r_L\right) LP = D - \left(1 - \pi \left(1 + r_L\right) P\right) L. \tag{4.9}$$

Individuals hold readily available cash to the amount of $D - L$ and loans to the future value if $\pi \left(1 + r_L\right) L$, which are then sold at price P.

Those not withdrawing funds buy these loans using their cash reserves, $D - L$, and obtain $\frac{D-L}{P}$ loans from this purchase. Including their original purchase of loans, $\pi \left(1 + r_l\right) L$, this gives rise to total funds in time period 2 of

$$\begin{aligned} \left(1 + r_D^2\right) D &= \frac{D - L}{P} + \pi \left(1 + r_L\right) L \\ &= \frac{1}{P} \left(D - L + P\pi \left(1 + r_L\right) L\right) \\ &= \frac{\left(1 + r_D^1\right) D}{P}. \end{aligned} \tag{4.10}$$

The price P must be set such that the market clears. If $P > \frac{1}{\pi(1+r_L)}$, all individuals would invest all their deposits into loans because it increases $\left(1 + r_D^1\right) D$, as equation (4.9) shows. However, the level of loans is not affecting $\left(1 + r_D^2\right) D$, as we see from the last equality in equation (4.10), which implies that there is no potential buyer of the investment, given no cash reserves would be held. In the case of $P < \frac{1}{\pi(1+r_L)}$, the reverse situation occurs and no loans are provided in the first place. Hence we need

$$P = \frac{1}{\pi \left(1 + r_L\right)}. \tag{4.11}$$

Inserting this into equations (4.9) and (4.10), we obtain

$$\begin{aligned} \left(1 + r_D^1\right) D &= D, \\ \left(1 + r_D^2\right) D &= \pi \left(1 + r_L\right) D \end{aligned} \tag{4.12}$$

and see that the deposit rates are independent of the amount of loans provided.

In order to achieve market clearing in the sale of the loan, the proceeds received by those selling in period 1 have to equal the cash amounts from individuals not withdrawing, i.e.

$$pP\pi\left(1+r_L\right)L=(1-p)\left(D-L\right), \tag{4.13}$$

which solves, when inserting from equation (4.11), for

$$L=(1-p)\,D<D. \tag{4.14}$$

Using equation (4.12) we obtain the constraint on optimization as

$$p\left(1+r_D^1\right)+(1-p)\,\frac{\left(1+r_D^2\right)}{\pi\left(1+r_L\right)}=1, \tag{4.15}$$

identical to the social optimum constraint in equation (4.3). However the first order condition for an optimum, identical to equation (4.5), would only be fulfilled if the utility function is such that

$$\frac{\partial u\left(D\right)}{\partial\left(1+r_D^1\right)}=\pi\left(1+r_L\right)\frac{\partial u\left(\pi\left(1+r_L\right)D\right)}{\partial\left(1+r_D^2\right)}. \tag{4.16}$$

This arises from the fact that the deposit rates r_D^t do not depend on the amount of loans provided and market clearing in the sale and purchase of loans requires a fixed relationship between deposits and loans.

Even if this condition for and optimum were fulfilled, a superior solution can be found if the consumers have a sufficiently large relative risk aversion, i. e. $-D\frac{\frac{\partial^2 u(D)}{\partial D^2}}{\frac{\partial u(D)}{\partial D}}>1$, as in this case $D\frac{\partial u(D)}{\partial D}$ is decreasing in D and therefore $\pi\left(1+r_L\right)\frac{\partial u(\pi(1+r_L)D)}{\partial D}<\frac{\partial u(D)}{\partial D}$, implying a better allocation can be found when increasing r_D^1 and decreasing r_D^2.

Providing the possibility of trading would increase the welfare of depositors, though. If the above outcome provides depositors with a lower utility than selling their loans at a fraction λ of its face value, then depositors would choose this option instead. Hence, while not reaching the social optimum, allowing for trading would weakly increase welfare in the economy compared to direct lending without the ability to trade loans.

Bank lending If all consumers deposit their wealth into a bank and the bank retains $p\left(1+r_D^1\right)D$ as cash to be paid out to those depositors withdrawing in time period 1 and providing loans with the remaining deposits, the optimal allocation as implied by equations (4.2) and (4.5) can be achieved. The constraints in equation (4.2) are trivially fulfilled and we can set deposit rates in line with the requirements of equation (4.5). Hence banks would be able to implement the social optimum.

This allocation is an equilibrium as no depositor individually has an incentive to withdraw their deposits if they do not require cash. This is because the optimal allocation requires $r_D^1<r_D^2$ and thus withdrawing deposits without the need for cash reduces the utility of the depositor due to him receiving a lower return. To see this requirement, consider the optimality criterion in equation (4.5); with the marginal

utility decreasing we would have in the case of $r_D^1 \geq r_D^2$ that $\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)} \leq \frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)}$ and hence the equality in equation (4.5) can never be fulfilled as this would require $\pi(1+r_L) < 1$, but for a loan to be viable we need that $\pi(1+r_L) \geq 1$ as otherwise the repayment of the loan cannot be guaranteed and depositors or banks would make a loss, making them better off not lending at all.

If we allow for the selling of loans, depositors not withdrawing deposits could reclaim their deposit, obtaining an amount of $\frac{D-L}{p}$ if we solve equation (4.2) for $\left(1+r_D^1\right)D$, and use these proceeds to buy the loan at price $P = \frac{1}{\pi(1+r_L)}$, resulting in funds $\frac{D-L}{pP} = \pi(1+r_L)\frac{D-L}{p}$ in time period 2. The amount $\frac{D-L}{p}$ would buy up all the loans which are worth $P\pi(1+r_L)L = L$ if using equation (4.11) for the price P. These two expressions need to be equal in market clearing, which solves for $D = (1+p)L$. This gives us total revenue of $\frac{D-L}{pP} = \pi(1+r_L)\frac{D-L}{p} = \pi(1+r_L)L$. When retaining the deposit, depositors obtain $\frac{\pi(1+r_L)L}{1-p}$ as derived from equation (4.2), which is higher if $p > 0$. Therefore, not only do banks allow us to achieve the socially optimal allocation, there are no incentives for depositors to withdraw early and buy loans from banks. It is that banks are robust to trading arrangements in loans.

Another equilibrium exists though, in which all depositors withdraw their funds in period 1, whether they require the cash or not. The argument for this equilibrium is that if all depositors withdraw their deposits, the bank will liquidate the loans, receiving λL and distribute these proceedings, with the retained cash, to those depositors seeking to withdraw. A depositor not seeking repayment in these circumstances will not receive any funds as all loans have been sold, leaving the bank with no means to repay his deposit in time period 2. If withdrawing deposits in time period 1, he would, in contrast, obtain a share of the liquidated funds and not be left empty-handed. It is therefore rational to withdraw deposits if everyone else does so. Such a 'bank run' will lead to the failure of the bank as will be discussed in more detail chapter 15.

Summary Banks take deposits and lend these out with long maturities. In contrast, deposits, which are used to finance these loans, are short-term in nature and can be withdrawn at any time. This apparent mismatch between the maturity of loans and deposits is managed by banks in that they retain a certain amount of cash to satisfy those depositors who withdraw their funds prior to the maturity of the loans they give. While this limits the amount of loans that can be given, and consequently the deposit rate that can be paid, it has the benefit of allowing depositors access to their deposits at any time. This benefit to depositors outweighs the lower deposit rate they obtain and banks are able to achieve the social optimum of balancing loan provision and cash holdings. This social optimum cannot be achieved by direct lending, even if loans can be traded, thus a market for loans does not achieve the same level of social welfare. As chapter 15 will show, the possibility of a second equilibrium in which depositors withdraw all deposits early and cause the bank to fail, impose a cost on the existence of banks that need to be weighed against these and other benefits.

Reading Diamond & Dybvig (1983)

4.2 Alternative banking structures

Commonly a bank provides long-term loans and finances this with short-term deposits, providing liquidity to depositors as was shown in chapter 4.1 to be optimal. Alternative banking models have been proposed that would in particular avoid the possibility of a bank run, i. e. a situation in which deposits are withdrawn, even if they are not needed as discussed in more detail in chapter 15. These proposals are often freshly discussed after or during a banking crisis as an alternative model of banking, but it is shown here that these provide inferior solutions to the established banking practices, at least during times in which no bank runs occur. Whether the suggested alternatives are preferable overall, would have to be decided by weighing the welfare gains from operating the established banking systems against the losses arising from bank runs, taking into account the frequency of such events. Other mechanisms to reduce the costs of bank runs, either those established by banks themselves (see chapter 16) or through regulation and government bailouts (see chapter 18 and part VI), have also to be considered for a complete assessment.

4.2.1 Narrow banking

Let us assume that depositors can withdraw their deposits D either in time period 1 or in time period 2, but not in both. The bank invests the proceeds from deposits into loans $L \leq D$ that are only repaid in period 2 with probability π, including interest r_L, and not repaid with probability $1 - \pi$. In narrow banking, banks are required to hold cash reserves such that they can meet all possible obligations to depositors, i. e. all depositors withdrawing. For all depositors withdrawing in time period 1, we have

$$\left(1 + r_D^1\right) D = D - L, \tag{4.17}$$

where r_D^1 represents the deposit rate and $D - L$ the amount retained as cash. Thus depositors withdrawing in time period 1 are obtaining the cash the bank holds. If $L > 0$, then $r_D^1 < 0$ and depositors will make a loss to withdraw. Those depositors that do not withdraw in time period 1, obtain the proceeds of the loan This gives then for late withdrawal that

$$\left(1 + r_D^2\right) D = \pi \left(1 + r_L\right) L + (D - L), \tag{4.18}$$

where r_D^2 denotes the deposit rate for these two time periods. We can now compare these two requirements with that of direct lending in equation (4.2) of chapter 4.1, which we reproduce here:

$$\left(1 + r_D^1\right) D = (D - L) + \lambda L, \tag{4.19}$$
$$\left(1 + r_D^2\right) D = \pi (1 + r_L) L + (D - L).$$

We immediately see, that for depositors withdrawing in time period 1, the repayments will be lower and for those withdrawing in time period 2, the repayments will be identical. It is therefore obvious that the welfare in narrow banking is not only lower than with traditional banks (which has a higher welfare than direct lending), but even lower than an economy without banks relying on direct lending.

One argument against traditional banks is the possibility of bank runs, i. e. all depositors withdrawing in time period 1, which is a second equilibrium in conventional banking systems. However, no such possibility exists with direct lending, and given that narrow banking is inferior even to this market structure, it cannot be optimal in any case. Hence, the proposal of narrow banking, which effectively requires banks to hold all deposits as cash reserves to avoid any losses to depositors ($L = 0$), is economically not desirable. Depositors would be better off or equally well off retaining their funds as cash without depositing it in a bank. Furthermore, no lending would be possible in this case, reducing welfare even further.

Reading Wallace (1996)

4.2.2 Market-valued deposits

Let us assume that depositors invest into the bank as shareholders and they do not have a certain repayment amount if they want to withdraw their deposits, but they will have to sell their shares at the prevailing market price. The bank takes deposits D that can be withdrawn either in time period 1 or in time period 2, but not in both. The deposits received are invested into loans $L \leq D$ that are repaid, including interest at a loan rate of r_L, in time period 2 with probability π or not repaid with probability $1 - \pi$.

If depositors are withdrawing in time period 1, they receive their share of the cash the bank holds, $D - L$, and the value S of the shares of the bank they are owed, which they can sell in the market. If the fraction of depositors withdrawing in time period 1 is p, we get with r_D^1 denoting the implied deposit rate of those withdrawing in time period 1 that

$$p \left(1 + r_D^1\right) D = p (D - L) + pS. \tag{4.20}$$

Those depositors that are not withdrawing, will buy these shares using their share of the cash the bank holds as this is not needed. They will buy an additional $\frac{(1-p)(D-L)}{S}$ shares. The payments these depositors obtain are their original claim on the fraction $1 - p$ of the loan repayments $\pi (1 + r_L) L$ and the shares they bought, each providing them with a payment of $\pi (1 + r_L) L$, giving them a fraction $\frac{(1-p)(D-L)}{S}$ of shares, such that

$$\left(1-p\right)\left(1+r_D^2\right)D=\left(1+\frac{D-L}{S}\right)\left(1-p\right)\pi\left(1+r_L\right)L, \tag{4.21}$$

where r_D^2 denotes the deposit rate of those retaining their deposits until time period 2.

In equilibrium, the market for shares clears and with the withdrawing depositors selling shares worth pS and the non-withdrawing depositors buying to the amount of $(1-p)(D-L)$. We get from equalling these two expressions that

$$S=\frac{1-p}{p}\left(D-L\right). \tag{4.22}$$

Inserting this into equations (4.20) and (4.21), we get the repayments in time periods 1 and 2, respectively as

$$\begin{aligned}\left(1+r_D^1\right)D&=\frac{D-L}{p},\\ \left(1+r_D^2\right)D&=\frac{\pi\left(1+r_L\right)}{1-p}L.\end{aligned} \tag{4.23}$$

We see that the amounts depositors obtain will depend on the fraction p of depositors withdrawing in time period 1. A higher withdrawal rate p will correspond to a lower repayment in time period 1 and a higher repayment in time period 2. The reason is that we see from equation (4.22) that in this case the value of the shares will be lower, given many are seeking to sell, and hence funds available to pay to depositors in time period 1 are lower. As the remaining depositors buy shares in the bank at a low price, the repayment of their deposits in time period 2 will increase.

We find that

$$p\left(1+r_D^1\right)+\left(1-p\right)\frac{\left(1+r_D^2\right)}{\pi\left(1+r_L\right)}=1, \tag{4.24}$$

which is identical to the social optimum we obtained in in equation (4.3) of chapter 4.1. Hence equity contracts should be optimal. Furthermore, a bank run where all depositors withdraw their deposit in time period 1 (see chapter 15) cannot occur as this reduces the amount repaid to depositors to $D-L$. A depositor which does not require cash, could retain his deposits in the bank and buy up all shares of the bank at a very low price, giving him a higher repayment than withdrawing his deposits. By adding to those who withdraw deposits, each depositor will reduce the amount he receives in time period 1 and therefore there is no incentive for a bank run. Even if all the other depositors have withdrawn, the final remaining depositor will be able to buy up the shares of the bank at very low cost, giving him a high return for retaining his deposit.

However, as the amount the bank repays in any time period is uncertain and will depend on the actual withdrawal rate, it is obvious that with risk-averse depositors, such an arrangement will be inferior to that of a fixed repayment as offered by traditional banks. We therefore see that as long as depositors are risk neutral, a bank offering deposits whose repayment is driven by market forces from early withdrawal, would provide the same socially optimal allocation as a conventional bank would achieve, but it has the additional advantage that a bank run should not occur. The uncertain outcomes if withdrawal rates p are not certain, though, make this proposal less attractive to depositors than conventional banks.

Another requirement is that a market for deposits (shares) has to exist with prices reflecting true values. In a realistic setting this will be difficult to achieve, given there are many uncertainties about the valuation of the loans, such as the value of π, the probability with which it will be repaid. If withdrawal rates fluctuate over time, the amount a depositor receives from withdrawal, will depend on the timing of his withdrawal decision. Therefore, while the proposal is attractive from a theoretical perspective, its implementation would be much more challenging and the certainty of the deposit repayments in conventional banks, makes them more attractive. This is especially the case if the likelihood of bank runs is low and other measures to mitigate their effects can be employed. We will discuss some of these aspects in chapters 16 and 18 as well as part VI.

Reading Jacklin (1987)

Résumé

We have seen two alternative ways banks could operate. In narrow banking, banks hold reserves that are sufficient to meet all possible withdrawal scenarios of depositors. This can be shown to be not only worse than the standard banking system, but even direct lending would provide a better solution. While narrow banking eliminates the risk of bank runs, this benefit has to be weighed against the loss in welfare arising from the inferior welfare during the much more common time periods in which bank runs do not occur. Another feature of banks is that they guarantee the return of the deposits, including any interest. It would be possible to modify deposits such that the payment for early withdrawal consists of the cash the bank has, and then provide them with shares in the bank that can be sold to generate additional cash for depositors. While bank runs cannot occur and the constraints on deposit repayments are compatible with the social optimum, the value of deposits is not longer guaranteed. This additional uncertainty would reduce the utility of any depositor exhibiting risk aversion. There are also practical difficulties in establishing a trading system for deposits that are fair to all depositors, making this proposal less attractive than conventional banks.

Different banking systems have been proposed that avoid the possibility of a bank run, the second and 'bad' equilibrium of conventional banks for the withdrawal of deposits, in addition to the 'good' equilibrium in which only those withdraw deposits who need to do so, for example for consumption. While it is possible to make bank runs irrational in such alternative banking systems, they have not been implemented on a wider scale, if at all. This is most likely the result of the simplicity of the conventional banks, combined with the rarity of banking runs on the one hand and the drawbacks of the alternative proposals on the other hand. Instead economies have regulated banks heavily to reduce the risk of bank runs (amongst other risks).

4.3 Deposit maturity

Banks allow deposits to be withdrawn at any time, while bank loans are in most cases not repayable for many time periods. In chapter 4.1 we have seen that this is socially optimal, but in order to implement this solution, it must also be optimal for banks to offer such short term deposits while lending out long-term. Let us assume that besides short-term deposits, banks would also be able to offer long-term deposits D that finance a loan L for two time periods. This loan is repaid with interest r_L with probability π, where $\pi(1+r_L) > 1$ making the loan viable. We will now ascertain how the provision of long-term and short-term deposits affects bank profits.

Long-term deposits After two time periods the deposits are due to be repaid. Hence we have the expected profits of depositors as

$$\Pi_D = \pi(1+r_D)D - D, \tag{4.25}$$

where r_D denotes the deposit rate paid for two time periods. Depositors are competitive and break even, hence we need $\Pi_D = 0$, giving us a deposit rate of

$$1 + r_D = \frac{1}{\pi}. \tag{4.26}$$

The profits of the bank are then given by

$$\Pi_B = \pi((1+r_L)L - (1+r_D)D) = (\pi(1+r_L) - 1)D. \tag{4.27}$$

We assume here that $L = D$ such that banks do not hold any excess cash and the final equality emerges when inserting equation (4.26). In addition we assume that banks have limited liability and only repay deposits if the loan is repaid.

Short-term deposits Alternatively, banks may only accept short-term deposits for a single time period, that subsequently need to be rolled over. If they are rolled over, the deposit rate $\hat{r}_D^1$ is paid in time period 1 and a new deposit raised at interest rate $\hat{r}_D^2$ for the second time period. As the loan is only due to be repaid at the end of time period 2, we assume that the interest in time period 1 is paid with certainty, giving the depositor a certain profit of $\left(1+\hat{r}_D^1\right)D - D$. In time period 2, the deposit will be repaid with probability π. i. e. only if the loan is repaid. Therefore, the profits of depositors across the two time periods, neglecting any discounting of future income, is given by

$$\hat{\Pi}_D = \left(1+\hat{r}_D^1\right)D - D + \pi\left(1+\hat{r}_D^2\right)D - D, \tag{4.28}$$

with the first two terms representing the profits in time period 1 as the deposits is repaid with certainty and the final two terms the expected profits in time period 2, where deposits are only repaid if the loan is repaid. Assuming again that depositors are competitive such that $\hat{\Pi}_D = 0$, we easily obtain that

$$1 + \hat{r}_D^2 = \frac{1 - \hat{r}_D^1}{\pi}. \tag{4.29}$$

The bank's profits are such that it pays interest in time period 1 and then obtains its profits from repaid loans during time period 2, such that

$$\hat{\Pi}_B = -\hat{r}_D^1 D + \pi \left(r_L - \hat{r}_D^2 \right) D = (\pi (1 + r_L) - 1) D, \tag{4.30}$$

where the final equality arises by inserting from equation (4.29). We can now compare the profitability f long-term and short-term deposits.

Comparing long-term and short-term deposits Comparing the bank profits with long-term and short-term deposits in equations (4.27) and (4.30), respectively, we see that $\hat{\Pi}_B = \Pi_B$ and thus banks are indifferent between long-term and short-term deposits. Similarly, depositors are competitive and in both instances obtain zero profits, would thus also be indifferent between either form. Banks would therefore be willing to provide short-term deposits.

The total interest costs to the bank with short-term deposits are $\hat{r}_D^1 + \hat{r}_D^2$ and r_D for long term deposits. If we set $\hat{r}_D^1 < 1 + \frac{1}{1-\pi}$, we can easily see that $\hat{r}_D^1 + \hat{r}_D^2 < r_D$. In general, this condition is fulfilled if the deposit rates are positive. It is therefore that interest costs of short-term deposits are lower than for long-term deposits. The reason is that as the bank rolls over deposits, they are repaid with certainty in time period 1, this reduces the risk to depositors as only the risk in time period 2 needs to be compensated. This risk is the same as for the long-term deposits and given they already have obtained some payments, are content with a lower deposit rate for time period 2. Taking into account this risk, however, reduces the expected total payments to depositors, and ensures the aggregate returns are identical for long-term and short-term deposits. Similarly, for banks, deposits in time period 1 are always repaid, making the costs higher to banks, but this is compensated with lower costs in period 2 as a reduced deposit rate is payable.

Summary Banks and depositors should be indifferent between long-term and short-term deposits as both provide the same expected profits. If we include, however, preferences for depositors to be able to withdraw deposits early, they would prefer short-term deposits. Being able to give a slightly lower deposit rate on short-term deposits due to these preferences, would increase the profits of banks and they would also prefer short-term deposits. It is thus not only socially desirable to have short-term deposits financing long-term loans, as shown in chapter 4.1, but it would also be the most profitable form of deposits to banks and the preferred deposit form to depositors.

Reading Brunnermeier & Oehmke (2013) and Cao (2022, Ch. 11.2)

4.4 Liquidity provision to borrowers

Banks do not only provide loans L at a loan rate r_L to companies over two time periods, but let us assume they also give credit lines $\hat{L}$, such as arranged overdrafts, for a fee $\hat{r}_L$ that companies can draw down as needed. Such credit lines can be used in the second time period if the company requires additional funding. We assume that with probability γ these credit lines are taken up and then charged at the loan rate r_L. Banks initially also hold cash C on which they earn no interest. They can raise additional funds M in the second time period to cover the loans demanded from the credit lines on which they are charged an interest rate r_M. These funds might be raised from the central bank, the interbank market or by approaching large institutional investors for additional deposits. However, raising such funds at short notice causes additional costs of $\frac{1}{2}cM^2$; these costs might reflect higher interest rates that need to be paid in the market if raising larger amounts.

In addition, we assume that in time period 1, depositors can withdraw their deposits. This happens with probability λ and might reflect their desire to consume. Thus, the bank faces two uncertainties, the demand for loans arising from the credit lines and the possible withdrawal of deposits. We will investigate how these uncertainties affect the holding of cash in time period 1 and the amount of liquidity that is provided to companies in the form of credit lines.

Bank balance sheets The expected bank profits in this situation are given by

$$\Pi_B = E\left[r_L L + \hat{r}_L \hat{L} + \mathbf{1}_{\hat{L}} r_L \hat{L} - r_M M - \frac{1}{2}cM^2 - (1 - \mathbf{1}_D)\, r_D D\right], \tag{4.31}$$

where $E[\cdot]$ denotes the expected value. These expected profits consist of the revenue from the loan, $r_L L$, and the credit line, $\hat{r}_L \hat{L}$, the interest on the additional loan, if taken up, $\mathbf{1}_{\hat{L}} r_L \hat{L}$, less the costs of raising the additional funds $r_M M + \frac{1}{2}cM^2$ and the payments to depositors, if not withdrawn $(1 - \mathbf{1}_D)\, r_D D$. The term $\mathbf{1}_{\hat{L}} \in \{0; 1\}$ is 1 if the credit line is used and zero otherwise; we have in addition that $E\left[\mathbf{1}_{\hat{L}}\right] = \gamma$. Similarly we have that $\mathbf{1}_D \in \{0; 1\}$ is 1 if the deposits are withdrawn and zero otherwise, where $E[\mathbf{1}_D] = \lambda$. Initial cash holdings will be positive to account for the uncertainty around the take up of credit lines and, as we will see below, the possible redemption of deposits. In the second time period there is no uncertainty about the need of funds and therefore cash holding will be zero.

Depositors may withdraw their funding D with probability λ and the credit lines are taken up with probability γ. Hence, the bank faces uncertainty in demand for its cash, arising from the need to repay depositors and pay out on credit lines. We assume that these events are positively correlated such that depositors and companies both demanding the same option, i.e. demand liquidity or not demand liquidity, has probability ρ. Thus, banks are exposed to liquidity shocks from depositors and borrowers, which are positively correlated. This can easily be justified that in times of high demand for goods (such as high consumption), depositors will withdraw deposits to consume while at the same time companies increase their investments to meet the increasing demand. Thus we can define

$$Prob\left(\mathbf{1}_D = \mathbf{1}_{\hat{L}} = 1\right) = Prob\left(\mathbf{1}_D = \mathbf{1}_{\hat{L}} = 0\right) = \frac{1}{2}\rho, \tag{4.32}$$
$$Prob\left(\mathbf{1}_D = 1, \mathbf{1}_{\hat{L}} = 0\right) = Prob\left(\mathbf{1}_D = 0, \mathbf{1}_{\hat{L}} = 1\right) = \frac{1}{2}(1-\rho),$$

with $Prob(\cdot)$ denoting the probability of an event.

For the balance sheet, we find at the start of the first and second time period, respectively, that

$$L + C = D, \tag{4.33}$$
$$L + \mathbf{1}_{\hat{L}}\hat{L} = (1 - \mathbf{1}_D)\, D + M,$$

where $\mathbf{1}_D \in \{0;1\}$ is 1 if deposits are withdrawn and zero otherwise. The first equation denotes that initially the deposits banks have obtained are invested into loans and cash. In the second time period no more cash is held, the credit lines may have been called on. This will be financed by the remaining (not withdrawn) deposits and any newly raised funds. For simplicity we ignore here the possibility that the take up of credit lines is so low that cash remains with the banks; we also do not consider that banks not requiring the cash in this situation might lend out the funds in the interbank market. Hence the additional funds would come from an external source, such as the central bank or large institutional investors.

From equation (4.33) we easily get by eliminating L that

$$M = \mathbf{1}_{\hat{L}}\hat{L} + \mathbf{1}_D D - C, \tag{4.34}$$

which implies that the additional funding raised, covers the additional loans given, deposits lost, less the cash reserves held initially.

Optimal credit lines After inserting equation (4.34) into equation (4.31) and differentiating, we get the following first order conditions for a profit maximum of the bank with respect to the credit lines and cash holdings:

$$\frac{\partial \Pi_B}{\partial \hat{L}} = \hat{r}_L + \gamma\,(r_L - r_M) - \frac{1}{2}c\frac{\partial E\left[M^2\right]}{\partial \hat{L}} = 0, \tag{4.35}$$
$$\frac{\partial \Pi_B}{\partial C} = -r_M - \frac{1}{2}c\frac{\partial E\left[M^2\right]}{\partial C} = 0.$$

When using equation (4.32), we obtain that

$$E\left[M^2\right] = \frac{1}{2}\rho\left(\hat{L} + D - C\right)^2 + \frac{1}{2}(1-\rho)\left(\left(\hat{L} - C\right)^2 + (D - C)^2\right), \tag{4.36}$$

having used that for $\gamma = \lambda = 1$ we have $M = \hat{L} + D - C > 0$, for $\lambda = 0$ and $\gamma = 1$ it is $M = \hat{L} - C$, for $\lambda = 1$ and $\gamma = 0$ we find $M = D - C > 0$ and for $\gamma = \lambda = 0$

$M = -C < 0$ and we assume in this case no additional funding is raised and we set $M = 0$ as negative funds cannot be raised and we ignore the possibility of lending out excess cash.

From equation (4.36) we then get

$$\frac{\partial E\left[M^2\right]}{\partial \hat{L}} = \hat{L} - C + \rho D, \tag{4.37}$$
$$\frac{\partial E\left[M^2\right]}{\partial C} = (2-\rho)\,C - \hat{L} - D.$$

Inserting the second equation in (4.37) into the final equation in (4.35), we get the optimal cash holdings as

$$C = \frac{\hat{L} + D - \frac{2r_M}{c}}{2-\rho}. \tag{4.38}$$

Using equation (4.38) in the first equation of (4.37), we get from the first equation in (4.35) that the optimal amount of credit lines is given by

$$\hat{L} = (1-\rho)\,D + 2\frac{(2-\rho)\left(\hat{r}_L + \gamma\left(r_L - r_M\right)\right) - r_M}{c\,(1-\rho)}. \tag{4.39}$$

We instantly see from equation (4.39) that

$$\frac{\partial \hat{L}}{\partial D} = 1 - \rho < 1, \tag{4.40}$$
$$\frac{\partial \hat{L}}{\partial \rho} = -D + 2\frac{\left(\hat{r}_L - r_M\right) + \gamma\left(r_L - r_M\right)}{c\,(1-\rho)^2}.$$

Hence, as deposits increase, banks will increase their credit lines and that way provide a liquidity cushion to companies. It is worth noting, however, that as $\frac{\partial \hat{L}}{\partial D} < 1$, the credit lines increase less the larger the bank becomes. The reason is that additional funding M for larger banks also increases and with costs increasing in M, larger banks are less willing to provide credit lines to customers due to these increased costs.

The impact of the correlation ρ between using the credit line and deposit withdrawals are not unambiguous. However, for low take ups of credit lines (γ) and the reasonable assumption that $\hat{r}_L < r_M$, i. e. the interest on additional funding is higher than the fee charged on credit lines, the second term in the second equation of (4.40) will be negative. Only for higher take up rates of credit lines, γ, and with loan rates exceeding the additional funding rates substantially will this term turn positive. If either the additional funding costs c are low or the correlation between liquidity demands by depositors and companies ρ is high, might the second term be larger than D and turn the entire expression positive. We can therefore state that in most cases the second derivative in equation (4.40) will be negative. This implies that as the correlation in liquidity demands by depositors and companies increases, the provision of credit lines reduces to account for the increased strain on cash resources from deposit withdrawals and the take-up of credit lines.

Summary Banks will provide companies with credit lines they can call on if needed and thereby provide liquidity not only to depositors but also companies. Giving companies access to credit lines provides an additional source of revenue for the bank, but the possibility of having to raise additional funds if credit lines are taken up, at potentially significant costs, will limit how much such liquidity banks are willing to provide. This willingness is reduced the more highly correlated the demand by companies to use such credit lines is with that of depositors withdrawing funds. Both, companies and depositors, may require banks to raise additional costly funds and a high correlation will incentivise banks to reduce their exposure to such risks by reducing credit lines, especially as the exposure to deposit withdrawal cannot be reduced. As the costs of raising additional funds is increasing in the amount demanded, large banks commonly will find that they have higher costs than smaller banks, unless due to their size they can obtain more favourable conditions for additional funding. This will lead to large banks providing less generous credit lines to their customers, relative to the bank size.

Reading Kashyap, Rajan, & Stein (2002)

Conclusions

We have seen that banks allow depositors to withdraw their deposits at any time and doing so is socially optimal. By holding a small amount of cash to repay those depositors wishing to withdraw deposits and investing the remainder into loans, they can provide the long-term loans borrowers desire without having to bind their depositors to the same length of time. This liquidity insurance to depositor is a key benefit that banks provide and that other banking structures are unable to achieve. It is for this reason that the traditional form of banking is often referred to as *fractional reserve banking*, because a fraction of the deposits banks receive, and that can be withdrawn at any time, are retained as cash, while the remainder is lent out long-term.

It is, however, not only socially desirable that banks provide this liquidity. If neither banks nor depositors have preferences for the maturity of deposits, banks and depositors are indifferent between long-term and short-term deposits as the interest rates are adjusted such that the expected outcome is the same in both cases, even though interest costs are lower for short-term deposits. These lower costs are the result of the roll-over of deposits where interest is paid out in each time period and the risk of bank failure arising from the default of the loan only affects the deposits in the final time period, while for long-term deposits all funds are affected. If there is a small preference by depositors for short-term over long-term deposits, it could be beneficial for both, banks and depositors, to depend on short-term deposits as depositors would accept slightly lower deposit rates, which would increase the profits to banks.

However, banks do not only provide liquidity insurance to depositors, they are also providing liquidity to borrowers by agreeing credit lines, such as overdrafts, on which companies can rely to cover any unexpected liquidity needs. In a similar way to ensuring depositors can withdraw their deposits, this adds flexibility to companies which do not need seek to take out excessive loans, but instead can rely on much cheaper credit lines. Any costs that arise from banks having to raise additional funds, such as from the institutional investors or central banks, will naturally limit the size of credit lines provided.

It is therefore banks that allow access to funds for either consumption or investment, while at the same time ensuring the amounts needed to held back in cash for this purpose are minimal. This allows for more loans to be provided and thus a much larger amount of investment in an economy, which ultimately benefits economic growth through the most efficient use of financial resources.

Chapter 5

Investment risks

Rather than assuming that the risks of an investment are given, we might surmise that borrowers can affect this risk. They might do so by increasing the effort level of managers or changing their strategy such that the interests of the lender are taken into account better. This might impose costs on the issuer, like increased efforts imposed on their management or reduced benefits to the management due to the changed strategy of the issuer. Both types of costs will affect the management's incentives to reduce the risks of investments. If a bank is able to better identify the risks of investments, this will affect the loan rate they charge, which in turn will determine the incentives to adjust the risk of the investment.

We assume that companies investing the proceeds of a loan have one of two possible outcomes where investments will succeed with probabilities π_H and $\pi_L < \pi_H$, respectively, resulting in returns of R_H and R_L. If an investment fails, no proceeds are available to the company. Through exerting effort, the company can affect the likelihood of realising investment H, succeeding with probability π_H. We can thus interpret p as a measure for the risk taken by the company as investment H is more likely to succeed. The higher p is, the lower the risk as the expected success rate is increasing in p.

The lender obtains a signal s on the likelihood of the company investing into investment H that has precision p_j, where $j = D$ for direct loans from 'depositors' and $j = B$ for the bank providing the loan. We have $p_B > p_D > \frac{1}{2}$. With p_j indicating the probability that the signal received is correct, we set

$$Prob\,(s = H|H) = Prob\,(s = L|L) = p_j. \tag{5.1}$$

The signal the bank receives, is more precise than that of the depositor, reflecting the assumption that banks have superior information.

With initial beliefs p on the likelihood of the investment H being realised, the signal allows the direct lender and the bank to update their beliefs such that for signals of realising investment H and investment, L, respectively, we have when using Bayesian learning that

$$Prob\,(H|s=H) = p_j^H = \frac{pp_j}{pp_j + (1-p)\left(1-p_j\right)}, \tag{5.2}$$
$$Prob\,(L|s=L) = p_j^L = \frac{p\left(1-p_j\right)}{p\left(1-p_j\right) + (1-p)\,p_j},$$

with $p_B^H > p_D^H > p > p_D^L > p_B^L$. Depending on the type of investment, i, the returns are given by $\pi_i\,(1+R_i)$, where we assume that $\pi_H\,(1+R_H) > 1 > \pi_L\,(1+R_L)$, such that investments of type H have a higher value than investments of type L and only the investment of type H is able to repay the loan amount fully.

Within this framework, we can now analyze the optimal risk of investments, as measured by p, and indirectly the optimal risk of loans provided, firstly in the absence of a bank by companies relying on direct lending only and then in its presence.

Direct loans Let us first assume that borrowers obtain a loan directly from a lender. This lender will be repaid the loan only if the investment of the company is successful. Depending on the signal he receives, the expected profits for signals H and L, respectively, are then given by

$$\Pi_D^H = p_D^H \pi_H \left(1+r_L^H\right) L + \left(1-p_D^H\right) \pi_L\,(1+R_L)\,L - L, \tag{5.3}$$
$$\Pi_D^L = p_D^L \pi_L\,(1+R_L)\,L + \left(1-p_D^L\right) \pi_H \left(1+r_L^H\right) L - L.$$

The first term in each expression denotes that the observed signal represents the type of investment realised truthfully, while the second term that this signal is wrong. If signal H is observed and this is correct or the signal L is observed and this is wrong, the loan is repaid in full, including interest r_L^H, given our assumption that $\pi_H\,(1+R_H) > 1$. If signal L is observed and this is correct or signal H is observed and this is wrong, the loan cannot be repaid fully due to our assumption that $\pi_L\,(1+R_L) < 1$ and the lender will seize the full value that is generated by the investment. For simplicity we assume that $\pi_L = 0$, and hence investments of type L are never successful. This simplifies equation (5.3) to

$$\Pi_D^H = p_D^H \pi_H \left(1+r_L^H\right) L - L, \tag{5.4}$$
$$\Pi_D^L = \left(1-p_D^L\right) \pi_H \left(1+r_L^H\right) L - L.$$

Perfect competition between lenders implies $\Pi_D^s = 0$ and hence

$$1 + r_L^H = \frac{1}{p_D^H \pi_H}, \tag{5.5}$$
$$1 + r_L^L = \frac{1}{\left(1-p_D^L\right)\pi_H}.$$

We see from equations (5.2) and (5.5) that

$$\begin{aligned}
\frac{\partial p_j^H}{\partial p} &= \frac{p_j\left(1-p_j\right)}{\left(pp_j+(1-p)\left(1-p_j\right)\right)^2}, \\
\frac{\partial p_j^L}{\partial p} &= \frac{p_j\left(1-p_j\right)}{\left(p\left(1-p_j\right)+(1-p)\,p_j\right)^2}, \\
\frac{\partial\left(1+r_L^H\right)}{\partial p} &= -\frac{\partial p_D^H}{\partial p}\frac{1}{\left(p_D^H\right)^2\pi_H}, \\
\frac{\partial\left(1+r_L^L\right)}{\partial p} &= \frac{\partial p_D^L}{\partial p}\frac{1}{\left(1-p_D^L\right)^2\pi_H}.
\end{aligned} \tag{5.6}$$

The company will now be able to obtain investment H with probability p. This probability will be affected by the efforts the company exerts, which comes at costs C and whose marginal costs are increasing in the probability p. We furthermore assume that $\frac{\partial C}{\partial p}=0$ if $p=0$ and $\frac{\partial C}{\partial p}=+\infty$ if $p=1$. The profits of the company are then given by

$$\Pi_C = p\left(p_D\pi_H\left(R_H-r_L^H\right)+\left(1-p_D\right)\pi_H\left(R_H-r_L^L\right)\right)-C, \tag{5.7}$$

where we note that only if investment H is chosen, does the company generate any funds as we assumed that $\pi_L=0$ and the loan is only repaid if the investment is successful. The loan rate will depend on the signal received by the direct lender for an investment of type H being realised; it is high with probability p_D and low with probability $1-p_D$. The first order condition of maximizing the company's profits by choosing the optimal probability of obtaining investment H, p, is given by $\frac{\partial \Pi_C}{\partial p}=0$, which solves for

$$\begin{aligned}
\frac{\partial C}{\partial p} = {} & p_D\pi_H\left(R_H-r_L^H\right)+\left(1-p_D\right)\pi_H\left(R_H-r_L^H\right) \\
& -p\pi_H p_D\frac{\partial\left(1+r_L^H\right)}{\partial p}-p\pi_H\left(1-p_D\right)\frac{\partial\left(1+r_L^L\right)}{\partial p}.
\end{aligned} \tag{5.8}$$

We can now compare this solution on the optimal probability of obtaining investment H with the situation in which a bank provides the loan.

Bank loans If banks provide loans, direct loans remains available. Given the more precise information of the bank, the signal received is correct with probability $p_B>p_D$, the loan rates they are charging with signals H and L, respectively, are given by

$$1+\hat{r}_L^H = \frac{1}{p_B^H\pi_H}, \tag{5.9}$$

$$1+\hat{r}_L^L = \frac{1}{\left(1-p_B^L\right)\pi_H}. \tag{5.10}$$

This result can be obtained in a similar way to the loan rates for direct lending as we assumed that banks do not face any financing costs by paying no interest on deposits.

We have due to $p_B^H > p_D^H > p > p_D^L > p_B^L$ that $\hat{r}_L^H < r_L^H < r_L^L < \hat{r}_L^L$. A bank would not provide a loan for which they receive a low signal L as the loan rate they would offer, $\hat{r}_L^L$, is higher than that of a direct loan, r_L^L. Hence in order to provide loans to finance investments for which they have received signal L, banks would have to charge a lower loan rate, making a loss from the loan and instead they prefer to not provide the loan. On the other hand, for loans on investments for which signal H has been obtained, the loan rate they can offer, $\hat{r}_L^H$, is lower than what a direct lender would charge, r_L^H. Thus investment banks will only provide loans for which they have received the high signal H. Only if banks receive signal L will direct lenders provide the loan, independent of their own signal. Implicitly we assume that the direct lender does not update its belief by using the information that a company seeking a loan from them, the bank must have obtained signal L.

A company would seek a loan from a bank if the bank obtains signal H as the bank than can offer better conditions than a direct loan. If the investment is of type H, the bank receives a signal H with probability p_B and the company takes a bank loan. If the bank receives signal L, which has probability $1 - p_B$, then the company takes a direct loan, which will be based on signal H of the direct lender with probability p_D, or signal L with probability $1 - p_D$. The case of the company making an investment of type L, which happens with probability $1 - p$, does not need to be considered as in this case the company will not obtain any proceeds and make zero profits. Thus we have the profits of the company given by

$$\begin{aligned}\hat{\Pi}_C = {} & p\left(p_B\pi_H\left(R_H - \hat{r}_L^H\right) + (1 - p_B)\, p_D\pi_H\left(R_H - r_L^H\right)\right. \\ & \left.+ (1 - p_B)(1 - p_D)\,\pi_H\left(R_H - r_L^L\right)\right) - \hat{C}.\end{aligned} \tag{5.11}$$

The first-order condition $\frac{\partial \hat{\Pi}_C}{\partial p} = 0$ for maximizing these profits then solves for

$$\begin{aligned}\frac{\partial \hat{C}}{\partial p} = {} & p_B\pi_H\left(R_H - \hat{r}_L^H\right) + (1 - p_B)\, p_D\pi_H\left(R_H - r_L^H\right) \\ & + (1 - p_B)(1 - p_D)\,\pi_H\left(R_H - r_L^L\right) \\ & - p p_B\pi_H\frac{\partial\left(1 + \hat{r}_L^H\right)}{\partial p} - p\,(1 - p_B)\, p_D\pi_H\frac{\partial\left(1 + r_L^H\right)}{\partial p} \\ & - p\,(1 - p_B)(1 - p_D)\,\pi_H\frac{\partial\left(1 + r_L^L\right)}{\partial p}.\end{aligned} \tag{5.12}$$

We can now compare riskiness of the investments and hence the loan, by evaluating the probability p of obtaining the investment of type H.

Comparison of direct and bank loans Comparing the expressions for the first order conditions of direct and bank loans from equations (5.8) and (5.12), we can rewrite the first order condition in equation (5.12) as

$$\frac{\partial \hat{C}}{\partial p} = \frac{\partial C}{\partial p}\left(1-p_B\right) + p_B\pi_H\left(R_H - \hat{r}_L^H\right) - pp_B\pi_H\frac{\partial\left(1+\hat{r}_L^H\right)}{\partial p}. \tag{5.13}$$

The company is more likely to obtain investment H, thus choose a higher value for p, if $\frac{\partial \hat{C}}{\partial p} > \frac{\partial C}{\partial p}$. The fact that marginal costs are increasing implies that higher marginal costs in the presence of bank loans, results in a higher probability of realising investment H. We can simplify this relationship to become

$$\begin{aligned}\frac{\partial C}{\partial p} &< \pi_H\left(\left(R_H - \hat{r}_L^H\right) - p\frac{\partial\left(1+\hat{r}_L^H\right)}{\partial p}\right) \\ &= \pi_H\left(1+R_H\right) + \frac{1-2p_B}{p_B},\end{aligned} \tag{5.14}$$

where we obtained the second line by inserting for $\hat{r}_L^H$ from equation (5.9) and differentiating this expression accordingly. With marginal costs being positive, this condition for bank loans resulting in lower risk investments is never fulfilled if the right-hand side is negative. This is the case if $p_B > \frac{1}{2-\pi_H(1+R_H)}$. Thus if banks are skilled in analysing company's investments and their signals sufficiently precise, bank loans are actually increasing the risk of investments, relative to a situation where only direct loans are available. The reason for this observation is that with more precise signals, the bank will reduce the loan rate ever more, as we can easily see from equation (5.9). This increases the profits of companies and they can reduce their efforts of reducing the risk of their investment by increasing p and saving on costs. For lower signal quality, p_B, however, the reduction in the loan rate from banks having more precise signals is not sufficient and companies will increase their efforts to increase p as well, because these efforts are recognised better by the bank than by direct lenders and the saving from a lower loan rate exceeds the increase in effort costs.

Even if $\frac{1}{2} < p_B \leq \frac{1}{2-\pi_H(1+R_H)}$ the condition in equation (5.14) can be fulfilled. As we had assumed that for $p = 0$ we have for the marginal costs that $\frac{\partial C}{\partial p} = 0$ and the marginal costs are increasing, there will exist a p^* such that for $p < p^*$ this condition is fulfilled, while for $p \geq p^*$ it will not be fulfilled. Hence if the equilibrium effort in a situation where only direct lending is available, is sufficiently low, bank lending increases efforts levels and thus reduces risks, while if it is sufficiently high, effort levels will reduce and risks increase.

Thus we observe two effects, one effect reduces the risk of investment by incentivising companies to exert more effort. The better informed banks will reduce the loan rate in response to a lower risk and receiving signal H more than a direct lender would and hence increasing efforts is beneficial. On the other hand, however, the more precise signal of banks will reduce the loan rate in any case if they receive signal H, increasing profits to companies, allowing them to reduce efforts and hence costs. For high-risk investments, a low p, or imprecise bank signals, a low p_B, the first effect dominates, reducing investment and loan risk. However, once the risks are sufficiently low and bank signals precise enough, the second effect will dominate, increasing investment and loan risk.

Summary Banks are assumed to have superior skills in identifying the risks of loans correctly. If borrowers, such as companies, can affect the risks of their investments, this increased ability to identify risks provides them with incentives to reduce the investment risk to increase their profits by being offered a lower loan rate. On the other hand, a better ability to identify investment and thereby loan risks will in any case reduce the loan rate if the bank received a positive signal on these risks, thereby reducing incentives to reduce risks. Which effect dominates, will depend on the level of risk taken and the precision of signals the bank receives. If the signals banks receive are not too precise and the investment risks high, bank lending will reduce investment and hence loan risks compared to direct lending; in the other case, with low risk investments and banks obtaining highly precise signals, investment risks will increase.

The introduction of banks will affect the risks taken by companies, either increasing or decreasing them. Thus banks are not only having an effect on transaction costs, making borrowing and lending more efficient, and providing liquidity to borrowers and depositors, but due to their ability to assess risks more precisely, they have an influence on the level of risk companies take. We thus see that the influence of banks exceeds that of an ordinary intermediary that seeks to bring together borrowers and lenders, they have a profound impact on investment decisions that goes well beyond the influence reduced transaction costs would have.

Reading Biglaiser & Li (2018)

Review

If banks were mere intermediaries between savers and borrowers, their impact would rather minimal. Chapter 2 has shown that even if banks have an advantage over individual lenders in providing loans, there is no overall economic benefit. The costs associated with these advantages will passed on to depositors and be balanced against the benefits and the overall effect is that banks have no positive or detrimental effect on the wider economy.

However, banks are more than merely intermediaries and can provide significant gains in efficiency of the lending process. Rather than every potential lender and borrower negotiating directly with each other, it is much more efficient to pool resources and manage the lending process centrally in banks as shown in chapter 3. This avoids duplication of effort, reducing the overall costs and increasing economic welfare. These benefits are, however, also achievable using online platforms that facilitate any matching of borrowers and lenders. The monitoring of borrowers would be more difficult to implement, but non-bank solutions could be sought.

Going beyond the gain in efficiency that banks can achieve, the main contribution of banks is their ability to transform short-term deposits into long-term loans. This liquidity provision or liquidity insurance, as discussed in chapter 4, is a central benefit in which banks can create economic gains that cannot be replicated in another way. The importance of banks here lies in their ability to satisfy simultaneously the needs of depositors for instant and easy access to their deposits, while borrowers want long-term and predictable access to funds that match their investment preferences. Banks achieve this by retaining a small amount of the deposits as reserves to pay out to those depositors requiring access to their money. It introduces a maturity mismatch between the long-term assets of the bank, loans, and their short-term liabilities, deposits, whose management we will discuss in more detail in part V. No other institutional set-up has been proposed that is able to achieve these benefits and thus banks are unique in this role.

Banks are more than intermediaries and they provide more than efficiency gains an economy. While many intermediaries are mainly increasing the efficiency of transactions, such as retailers or brokers in financial markets, banks are unique in that they provide more benefits to the economy. They provide liquidity to depositors, and to some extend also to borrowers. In addition, they may also affect the characteristics of investments by providing incentives to companies to change their risk profile, either reducing risks or increasing them. It is not possible to achieve these benefits without some costs, banks might take too much risks or create instabilities in other areas of the economy. The remaining parts of this book will look at how banks manage these risks, how these risks actually emerge, and how economies have reacted with regulation to mitigate some of the downside of banks, without reducing the benefits of banks substantially.

Part II
The provision of loans

Providing loans is one of the two key roles of banks, along with taking deposits. Such loans are used by companies to finance investments, to modernise existing facilities, expand their business, or ensure the continuation of the current business from liquidity shortages. Individual borrowers use loans to purchase houses, cars, household appliances, or finance education, holidays and weddings. Banks may also finance governments and non-government organisations, although they more commonly finance their expenditure through capital markets. While the motivations for seeking a loan might be very different, the key aspects a bank has to consider are very similar. They need to assess the risks of the loan, the likelihood with which it is repaid, and set a loan rate according to this risk and the general market conditions. In addition, they might want to consider whether the maturity of the loan is such that it finances the investment or purchase for its entire time or whether the loans would need to be rolled over, with the option for the bank to call in the loan pre-maturely and the borrower to switch their loan to another bank. Finally, some loans are only granted if collateral is provided, or the borrower might offer collateral to the bank in order to obtain better loan conditions, and the bank needs to make decisions on the use of such collateral.

A loan contract and decisions surrounding it is, however, much more complex than these decisions seem to imply. Providing a borrower with a loan gives mainly rise to two complications. Firstly, the borrower might not use the loan as anticipated by the bank, they might use the funds for another investment with different characteristics than envisioned by the bank, or they might not exert the effort levels required to ensure the repayment of loans. Such moral hazard will affect the way loan contracts are structured to align the incentives of borrowers with those of the bank. In addition, companies might know better the prospects of them repaying their loans, giving them an informational advantage over the bank when negotiating a loan. Once again, this will affect the loan contract such that the interests of banks and companies are more aligned. Furthermore, differences in information may also persist between banks, with some banks better informed about a company than another bank, which has implications for the competition between banks for granting loans and the loan conditions they will be offering.

The basic contract specifications of loans are discussed in chapter 6. We will see how the typical loan contract is optimal in the presence of moral hazard, but will also look at the optimal maturity of loans and seniority structure. A key concern for banks is the ability of companies to repay their loan, but even if companies might be able to repay a loan, they might decide to default on their obligations. We will discuss such strategic default in chapter 7 and the use of collateral is then explored in more detail in chapter 9 and in chapter VIII we will discuss why companies may not obtain a loan of the size they seek, but are offered only a smaller loan.

Information about companies applying for a loan is essential for the decision whether grant such a loan. Banks who had previous interactions with that company, such as having granted them loans, will hold some information from that interaction that might be useful to another bank in assessing their loan application. Despite being in competition with each other, banks may find it beneficial to share some information about a company with each other through credit reference agencies as chapter 10 will show. Banks having had interactions with a company before will have an informational advantage over other banks that lack such interactions. In chapter 11 we will explore the consequences of such relationship banking. The informational advantage of banks from relationship banking has lead to the option of companies sharing information with other banks, called open banking, and in chapter 12 we will see under which conditions companies choose the share information and reduce the impact of relationship banking

Banks having granted a loan may want to sell this loan off through a process called securitization. As we will see in chapter 13, securitization allows the bank to free up resources to grant additional loans. We will see under which conditions such securitization is optimal.

Chapter 6

Loan contracts

Loan contracts have commonly very distinct characteristics in that the lender will have a to bear the losses if the borrower's investment is not successful and he has not the means to make the agreed payments to the lender. However, if the investment is successful and the borrower can make repayments the amount the lender is due is strictly restricted; commonly the amount that is repaid consists of the initial amount obtained as a loan and the interest agreed. Hence, there is no participation of the lender in any profits the borrower might make using the proceeds of the loan they have provided. Such a property of loans is in strict contrast to equity, which fully participates in any profits the company makes. This property of loan repayments being restricted to the initial loan amount plus interest is assessed in chapter 6.1 and shown to be optimal as long as the outcomes of the investment the borrower makes, is difficult to verify for the bank.

Beyond the specification of the repayment modalities of loans, other contract specifications are also of relevance. It is often assumed that the length of the investment and the time to maturity coincides; the ability of banks to transform short-term deposits, which most depositors prefer, into long-term loans was seen as one of the key benefits emerging from the presence of banks as discussed in chapter 4. While having loans mature prior to the completion of an investment can leave the company exposed to the risk of not being able to roll-over loans and face losses from any required liquidation of such investments, companies might seek loans that are of longer maturity than the investment they conduct and use a single loan for a sequence of such investments. Chapter 6.3 will explore under which conditions such an extended maturity of loans is optimal. Before this discussion, we will discuss in chapter 6.2 the importance of banks acquiring information and how this affects the competition between banks for the provision of loans as well as loan rates.

Finally, many companies seek loans from several banks. While this might arise if banks facing lending restrictions, such as limits on the size of loans they can provide, the need to diversify their lending activity, or the desire of companies to retain relationships with multiple banks, see chapter 11 for a discussion of relationship banking, this is not always the main reason. In chapter 6.4 we will discuss the use of senior and subordinated loans, loans that have different priorities of being repaid if the company defaults on its obligations. Such arrangements can be optimal for companies to increase their profits by allowing for larger loans and a lower reliance on equity finance. we frequently see that loans are renegotiated, especially the amount that needs to be repaid and in chapter 6.5 we see that such renegotiations can be beneficial fir both companies and banks. Finally, chapter 6.6 will investigate the monitoring of loans by banks and how splitting a loan between two banks may increase monitoring by banks.

Looking at the basic loan contract specifications, how the loan is repaid, the time to maturity, and how any loans the company raises are allocated across different types, this chapter provides the foundation to explore further in subsequent chapters specific aspects of the loan contract and how they can be used to affect the provision of loans, but also the behaviour of companies.

6.1 The optimal repayment of loans

With very few exemptions, loans require the borrower to repay a fixed amount, which consists of the initial loan amount and interest as agreed at the outset. If the borrower cannot repay this amount, he is in default and the bank will have the right seize any assets the borrower has, to maximize the amount they are repaid. As such loan contracts are common and, subject to any regulatory constraints, by far the most common type of loan contract found, it suggests that this type of contract is optimal. We will show in chapter 6.1.3 that such a contract is indeed optimal if the outcome of an investment cannot be readily verified by the bank without incurring costs. Such auditing costs to banks are crucial, as chapters 6.1.1 and 6.1.2 will show that if outcomes of investments are common knowledge, different loan contracts are optimal.

6.1.1 An optimal risk-sharing contract

A company takes out a loan L in order to make an investment with an uncertain outcome V. Having obtained the outcome of their investment, the company has to repay their loan to the bank; this repayment will depend on the outcome the company has obtained as the repayment cannot exceed this amount. Hence for an outcome V, the repayment of the loan, $R(V)$, has to fulfill $R(V) \leq V$. We seek to derive the repayment function $R(V)$ that is optimal for the company, subject to the bank willing to provide such a loan.

The company will maximize their expected utility from the value they obtain from the outcome after repaying the loan, $V - R(V)$, subject to the bank willing to lend, i.e. achieving a utility level from the repayment $R(V)$ to compensate for their costs of providing this loan, such as interest on deposits. Hence the company will seek the repayment function $R(V)$ that maximizes $E[U_C(V - R(V))]$, given that $E[U_B(R(V))] \geq U_B^0$.

As marginal utility is positive, the company would seek to obtain the smallest repayment that meets the requirements of the bank. This implies that the constraint on the bank willing to provide the loan will be binding. Therefore with a Lagrange coefficient λ we seek to maximize

$$\mathcal{L} = E\left[U_C\left(V - R(V)\right)\right] - \lambda\left(E\left[U_B\left(R(V)\right)\right] - U_B^0\right). \tag{6.1}$$

Alternatively we could maximize the expected utility of the bank, subject to it being profitable to the company, yielding an optimization problem equivalent to equation (6.1).

The first order condition of our maximization yields

$$\frac{\partial \mathcal{L}}{\partial R(V)} = -\frac{\partial E\left[U_C\left(V - R(V)\right)\right]}{\partial R(V)} + \lambda \frac{\partial E\left[U_B\left(R(V)\right)\right]}{\partial R(V)} = 0. \tag{6.2}$$

Using the implicit function theorem to solve equation (6.2) for $R(V)$, we get

$$\begin{aligned}\frac{\partial R(V)}{\partial V} &= -\frac{\frac{\partial^2 \mathcal{L}}{\partial R(V)\partial V}}{\frac{\partial^2 \mathcal{L}}{\partial R(V)^2}} \\ &= \frac{\frac{\partial^2 U_C(V-R(V))}{\partial R(V)\partial V}}{\frac{\partial^2 U_C(V-R(V))}{\partial R(V)^2} + \lambda \frac{\partial^2 U_B(R(V))}{\partial R(V)^2}}.\end{aligned} \tag{6.3}$$

We have the absolute risk aversion defined as $z_i = -\frac{\frac{\partial^2 E[U_i(V)]}{\partial V^2}}{\frac{\partial E[U_i(V)]}{\partial V}}$ and solve equation (6.2) for λ and insert this expression into (6.3) to obtain

$$\frac{\partial R(V)}{\partial V} = \frac{z_C}{z_C + z_B} > 0. \tag{6.4}$$

Integrating this equation the repayment function becomes

$$R(V) = R_0 + \frac{z_C}{z_C + z_B} V, \tag{6.5}$$

where we need to set R_0 such that $R_0 \leq \left(1 - \frac{z_C}{z_C + z_B}\right) V$ as we require that $R(V) \leq V$. With this requirement having to be fulfilled for $V = 0$, the lowest possible outcome, this implies that $R_0 \leq 0$ and it will be set such that the participation constraint for banks, $E\left[U_B\left(R(V)\right)\right] \geq U_B^0$, is fulfilled with equality.

Hence the optimal repayment function would consist of a fixed payment R_0 from the bank to the company and a fraction $\frac{z_C}{z_C + z_B}$ of the outcome the company achieves being paid to the bank. The more risk averse the bank is the smaller the fraction of the outcome the bank will be paid and, in order to meet the participation constraint of banks, the fixed payment R_0 they make to companies will be reduced as this reduces the overall uncertainty in the repayments the bank obtains.

We clearly see that the contract specification is in no way comparable to that of a debt contract as commonly found; it more resembles a participation in the investment, comparable to equity, subject to an additional payment R_0. The reason for this result is that we assumed implicity that the outcome V can be verified by the bank to ensure it obtains its share of the outcome. If this verification is not possible, the company could report a lower value, such as $V = 0$ and make additional profits at the cost of the bank. In this more realistic scenario of outcomes not being readily identifiable, such a risk sharing loan contract becomes unviable as the bank would not obtain any repayments.

Readings Freixas & Rochet (2008a, Ch. 4.1), Keiding (2016a, Ch. 5.2)

6.1.2 Effort and moral hazard

The outcome of investments will often also depend on the amount of effort the company exerts to ensure its success. However, if the benefits of this effort go to the bank in order to repay the loan taken out to finance the investment, the incentives to exert effort are limited. As this moral hazard affects the ability of the company to repay its loan to the bank, it should be considered in the structuring of the loan contract and thus affect the repayment of the loan.

Let us assume that the probability of success of an investment, π, wholly financed by a loan, will depend on the effort level, e, but that those efforts also impose costs C on the company. In case of success the investment yields an outcome of V and with the repayment of the loan being denoted as $R(V)$, the company profits are the given by

$$\Pi_C = \pi (V - R(V)) - C. \tag{6.6}$$

With limited liability of the company, repayments cannot exceed the value the company obtains from the investment and the bank is willing to lend to the company if they obtain at least their costs for the loan of size L, which are arising from the deposits used to finance this loan r_D. Finally the profits to the company at the optimal effort level must be higher than at other effort levels. This leads to the following restrictions in the optimization of the company profits in equation (6.6):

$$\begin{aligned} R(V) &\leq V \\ \pi R(V) &\geq (1 + r_D) L \\ \Pi_C^* &\geq \Pi_C. \end{aligned} \tag{6.7}$$

The optimality of the effort level, $\Pi_C^* \geq \Pi_C$, can be replaced by its first order condition, which from equation (6.6) is easily obtained as

$$\frac{\partial \Pi_C}{\partial e} = \frac{\partial \pi}{\partial e} (V - R(V)) - \frac{\partial C}{\partial e} = 0. \tag{6.8}$$

Neglecting the requirement that $R(V) \leq V$ for now, we get the Lagrangian equation thus as

$$\begin{aligned}\mathcal{L} &= \pi\left(V - R\left(V\right)\right) - C + \lambda_1\left(\frac{\partial \pi}{\partial e}\left(V - R\left(V\right)\right) - \frac{\partial C}{\partial e}\right) \\ &\quad +\lambda_2\left(\pi R\left(V\right) - \left(1 + r_D\right) L\right) \\ &= \left(V\left(\pi + \lambda_1\frac{\partial \pi}{\partial e}\right) - C - \lambda_1\frac{\partial C}{\partial e}\lambda_2 - \left(1 + r_D\right) L\right) \\ &\quad +R\left(V\right)\left((\lambda_2 - 1)\pi - \lambda_1\frac{\partial \pi}{\partial e}\right),\end{aligned} \tag{6.9}$$

with λ_i denoting the Lagrange multipliers. The first term in the second equality does not depend on $R\left(V\right)$ and can thus be neglected in the optimization process.

If $\frac{\frac{\partial \pi}{\partial e}}{\pi} < \frac{\lambda_2 - 1}{\lambda_1}$, we see from equation (6.9) that the Lagrangian is increasing in the repayment $R\left(V\right)$ and thus it is optimal for the company that the maximal repayment $R\left(V\right) = V$ is chosen. If $\frac{\frac{\partial \pi}{\partial e}}{\pi} > \frac{\lambda_2 - 1}{\lambda_1}$, then the Lagrangian is decreasing in the repayment $R\left(V\right)$ and we should choose the lowest possible repayment, $R\left(V\right) = 0$.

Let us now assume that $\frac{\frac{\partial \pi}{\partial e}}{\pi}$ is increasing in the outcome of the successful investment, V. The marginal impact of effort on the probability of success, $\frac{\partial \pi}{\partial e}$, is increasing faster than the probability of success, π, itself. This can be interpreted that as the outcome of the investment V increases, the success is more and more likely to be attributed to the effort of the company, rather than chance. Using this assumption, we see that there will exist a V^* such that $\frac{\frac{\partial \pi}{\partial e}}{\pi} = \frac{\lambda_2 - 1}{\lambda_1}$ and for outcomes below V^*, we will set the repayment function such that $R\left(V\right) = V$ and for higher outcomes we require no repayment $R\left(V\right) = 0$.

The optimal loan contract should therefore require the company to give the bank their entire revenue if the outcome is small, but if the outcome is high, no repayment is required. Hence for less profitable investments with low outcomes V, all proceeds are retained by the bank, while for highly profitable investments, high outcomes V, the loan is not repaid at all. The threshold above which the loan does not need to be repaid, V^*, is through λ_2 determined such that the bank will recover its funding costs $\left(1 + r_D\right) L$ and is willing to provide this loan.

We have a loan contract in which the most profitable companies are defaulting on their loans, while less profitable companies use all their proceeds of the investment to repay their loan. Such a contract is different from the loan contract typically observed in the market. The implementation of this contract relies on the ability of the bank to verify the outcome V. If this verification is not possible, the company could report a high outcome $V > V^*$, such that $R\left(V\right) = 0$ and generate profits at the cost of the bank whose loan is not repaid. In the more realistic scenario where outcomes not readily be identifiable, such a loan contract as suggested here, becomes unviable.

Reading Innes (1990)

6.1.3 Optimal loan contracts with auditing costs

A company takes out a loan L in order to make an investment with an uncertain outcome V. Having obtained the outcome of their investment, the company has to repay their loan to the bank; this repayment will depend on the outcome the company has obtained as the repayment cannot exceed this amount. Hence for an outcome V, the repayment of the loan, $R(V)$, has to fulfill $R(V) \leq V$. However, in general, the outcome of investments are not easily observable to banks and verifying any declared investment outcomes imposes costs on the bank in the form of auditing. In the absence of such auditing, the company could misrepresent the outcome achieved to ensure they can repay a lower amount.

If auditing is expensive, the bank will seek to minimize the need to conduct such audits and hence only audit investment outcomes if necessary. Let us assume the bank only initiates an audit if the outcome is in a region $\mathfrak{A}$, thus for $V \in \mathfrak{A}$. Suppose there are investment outcomes $V_i \notin \mathfrak{A}$ and $V_j \notin \mathfrak{A}$ that are both not audited. Then, if the repayment function is such that $R(V_j) < R(V_i)$, the company would always declare an outcome V_j to repay the least amount. Thus for any $V_i \notin \mathfrak{A}$ we can only have an effective repayment function that pays the minimum of the value of all repayments of those investment outcomes that are not audited. We hence obtain that for any $V_i \notin \mathfrak{A}$, it is $R(V_i) = \overline{R} = \inf \left\{ R(V_j) | V_j \notin \mathfrak{A} \right\}$.

Furthermore, if $V_i \in \mathfrak{A}$ and $R(V_j) < R(V_i)$, we need that $V_j \in \mathfrak{A}$ as otherwise the company would declare V_j to reduce repayment and this claim would not be audited. Hence for $V_i \in \mathfrak{A}$ we require that $R(V_i) \leq \overline{R}$. This implies that the repayments of audited investment outcomes are lower than those of investment outcomes that are not audited.

These incentive constraints allow for a large range of repayment forms. Assuming banks want to ensure the highest possible repayment to them, they would request $R(V_i) = V_i$ for $V_i \in \mathfrak{A}$. Hence the contract specification will be

$$R(V) = \min\{V, \overline{R}\}, \tag{6.10}$$

where $\overline{R}$ is set such that the company is willing to conduct the investment, while the bank generates sufficient returns to make a profit, including any auditing costs. We have $R(V) = V$ if $V \in \mathfrak{A}$ and $R(V) = \overline{R}$ if $V \notin \mathfrak{A}$. This result implies that the auditing region is defined as $\mathfrak{A} = \left[0; \overline{R}\right]$; extending the auditing region to any point $\overline{V} > \overline{R}$ would allow a company with outcome $\overline{V}$ to claim to have received an outcome $V \notin \mathfrak{A}$ with $V < \overline{V}$ and reducing loan repayments from $\overline{R}$; hence this cannot be an equilibrium.

In line with common conventions, $\overline{R}$ would include the interest r_L on the loan L and hence we can define the repayment required as $\overline{R} = (1 + r_L) L$. We thus observe that the outcome is not audited if the original amount of the loan, including interest, is repaid in full, while the outcome is audited if only a smaller amount of $V < (1 + r_L) L$ is repaid. The former case is commonly referred to as the loan being repaid and the latter case as a default by the company. This auditing regime ensures that where the repayment depends on the outcome, there is no incentive to misrepresent this outcome, as the auditing is assumed to identify any such misrepresentations.

This repayment function recovers the commonly found contract and its specification of the repayment of a loan. A fixed amount of $(1 + r_L) L$ is repaid if this is possible, $V > (1 + r_L) L$, and if this is not possible, the bank is paid all resources the company has available, V. The optimal auditing region $\mathfrak{A}$ will be obtained by balancing the expected auditing costs, Prob $(V \in \mathfrak{A})\, C$, where each auditing costs C. Reducing the size of the auditing region $\mathfrak{A}$, will reduce these expected auditing costs, but it will also reduce the expected repayments as due to this repayment being determined by $\overline{R} = \inf \{R(V)|V \notin \mathfrak{A}\}$, it will be encompass smaller values due to being expanded to lower values of V, which was identified as the repayment amount if $V \in \mathfrak{A}$. Maximizing its profits, the bank will determine the optimal size of its auditing region $\mathfrak{A}$ and hence its loan rate r_L due to $\overline{R}$ being determined by the lower end of the auditing region and $\overline{R} = (1 + r_L) L$ defines the loan rate implicitly.

To show that this repayment function is unique, let us assume there is an alternative repayment function $\hat{R}(V)$ with a different auditing region $\hat{\mathfrak{A}}$ and a different maximal repayment $\hat{\overline{R}}$ if $V \notin \mathfrak{A}$. As the above contract was optimal, this alternative contract needs to be comparable and we need that $E[R(V)] = E\left[\hat{R}(V)\right]$ and $Prob\,(V \in \mathfrak{A}) = Prob\left(V \in \hat{\mathfrak{A}}\right)$ to equalize expected repayments and auditing costs, thus giving the same expected profits to banks.

As obviously we have to assume that $\mathfrak{A} \neq \hat{\mathfrak{A}}$, we can find an outcome V such that $V \in \mathfrak{A}$ and $V \notin \hat{\mathfrak{A}}$. This then implies that $\hat{R}(V) = \hat{\overline{R}} \leq V = R(V) < \overline{R}$, where the first equality arises from $V \notin \hat{\mathfrak{A}}$, the second inequality from the contract specification in equation (6.10), the third equality from $V \in \mathfrak{A}$ and the final inequality from the contract specification in equation (6.10), again. Thus the contract specified in equation (6.10) is optimal only for the auditing region $\mathfrak{A}$ as any other contract implies $\hat{R}(V) \leq R(V)$ and thus $E[R(V)] \geq E\left[\hat{R}(V)\right]$, making such an auditing region inferior. It is therefore that the auditing region $\mathfrak{A}$ in the contract is unique and the optimal contract is thus unique, too.

Another observation we can make from this model is that if auditing costs are reducing, then these lower auditing costs would allow the bank to expand the auditing region, implying a higher loan rate $\overline{R} = (1 + r_L) L$, but also a wider region in which the bank will receive repayments $R(V) = V$. Hence the loan contract has the characteristic of equity over a larger range and making the full repayment of the loan less likely. Thus low auditing costs will make a loan contract more like a risk-sharing contract and in their absence, full risk-sharing would be implemented. However, in reality auditing costs can be substantial given the complexity of businesses, making the conventional loan contract, which in the vast majority of cases is repaid in full, the most common contract form.

Readings Townsend (1979), Gale & Hellwig (1985)

Résumé

If we assume that the outcome of investments are not observable by banks, but only after a costly auditing, we have shown that the commonly observed loan contract is an optimal arrangement for banks to provide funding for companies. As we have seen that different types of contracts would be optimal if the outcome of investments were freely observable to banks, namely a risk-sharing contract and a contract in which highly profitable companies are not repaying loans at all, the importance of the availability of information has be highlighted. In situations where information is readily available, risk-sharing contracts, akin to equity, are favoured, while in situations where information on the outcome of investments cannot be readily verified, the typically found loan contracts are optimal.

It might well be that investment outcomes are not easily verified in the short-run, but would become apparent over longer time periods. It can be that positive investment outcomes can be hidden for some periods of time, for example by applying accounting measures to reduce the profits shown, while over longer time periods, such concealment of profits will be much more difficult to achieve. This would favour the use of loans for relatively short-term financing requirements, while equity is preferred for long-term financing needs.

6.2 Information acquisition and competition

In order to assess the risks of a company, banks need to acquire information. However, acquiring such information will be costly and the benefits arising from the information will have to outweigh the costs incurred to be beneficial. It is often not a case of either having information or not having information on a company, but banks can acquire more and more precise information on a company, most likely at ever increasing costs. Having more precise information will allow the bank to charge loan rates that reflects the risks the company poses to the bank more precisely, but the bank may face competition from other banks that are not as well informed, who, based on their more incomplete information, might provide a more attractive loan rate to companies, thus taking lending away from informed banks, limiting the value of information. We will here look into the interaction between informed and uninformed banks, how this affects the likelihood of providing loans to companies of specific characteristics and the loan rate banks will charge.

Signals and precision of information Assume there are two types of companies in the market, one succeeds with its investment, wholly financed by a loan L, with probability π_H and the other with probability $\pi_L < \pi_H$. In case the investment is successful, the company obtains an outcome of $(1 + R)L$ and in case of failure no proceeds are obtained and the loan cannot be repaid. Banks know that there is a fraction ν of companies with success rate π_H, such that the average success rate is given by

$$\pi = \nu\pi_H + (1 - \nu)\pi_L. \tag{6.11}$$

Banks receive a noisy signal s about the type of company, denoted H or L. This signal s is correctly reflecting the type of company with probability γ such that we have

$$\gamma = \text{Prob}\,(s = H|H) = \text{Prob}\,(s = L|L) > \frac{1}{2}. \tag{6.12}$$

We can interpret γ as the precision of the information they obtain; the more precise information is, the more likely it is to be correct.

Furthermore, the unconditional probabilities of observing signals H and L, respectively, are given from probability theory as

$$\begin{aligned} \text{Prob}(s = H) &= \text{Prob}\,(s = H|H)\,\text{Prob}\,(H) \\ &\quad + \text{Prob}\,(s = H|L)\,\text{Prob}\,(L) \\ &= \gamma\nu + (1-\gamma)(1-\nu), \\ \text{Prob}(s = L) &= \text{Prob}\,(s = L|H)\,\text{Prob}\,(H) \\ &\quad + \text{Prob}\,(s = L|L)\,\text{Prob}\,(L) \\ &= \gamma(1-\nu) + (1-\gamma)\nu. \end{aligned} \tag{6.13}$$

A signal $s = H$ can be observed if it correctly reflects the true outcome H, but also if it the true outcome is L but the observed signal is wrong. Similarly for a signal $s = L$, which can reflect the true outcome L, but might also be observed if the signal is wrong. The second equalities arise from inserting from equation (6.12) and noting that there are a fraction ν $(1-\nu)$ of companies H (L).

Finally, using Bayes' theorem, we obtain the probability of evaluating a company of a certain type, given the signal we have received, as

$$\begin{aligned} \text{Prob}\,(H|s = H) &= \frac{\text{Prob}\,(s = H|H)\,\text{Prob}\,(H)}{\text{Prob}\,(s = H)} \\ &= \frac{\gamma\nu}{\gamma\nu + (1-\gamma)(1-\nu)}, \\ \text{Prob}\,(L|s = H) &= 1 - \text{Prob}\,(H|s = H) \\ &= \frac{(1-\gamma)(1-\nu)}{\gamma\nu + (1-\gamma)(1-\nu)}, \\ \text{Prob}\,(H|s = L) &= \frac{\text{Prob}\,(s = L|H)\,\text{Prob}\,(H)}{\text{Prob}\,(s = L)} \\ &= = \frac{(1-\gamma)\nu}{(1-\gamma)\nu + \gamma(1-\nu)}, \\ \text{Prob}\,(L|s = L) &= 1 - \text{Prob}\,(H|s = L) \\ &= \frac{\gamma(1-\nu)}{(1-\gamma)\nu + \gamma(1-\nu)}, \end{aligned} \tag{6.14}$$

where the second equalities arise from using the definition of γ in equation (6.12) and recognising that the fraction of companies of type H (L) is ν $(1-\nu)$.

We can now define the probability of the company being successful, given the signal H or L was received, respectively, as

$$\begin{aligned}\hat{\pi}_H &= \pi_H \text{Prob}\,(H|s = H) + \pi_L \text{Prob}\,(L|s = H)\,, \\ \hat{\pi}_L &= \pi_H \text{Prob}\,(H|s = L) + \pi_L \text{Prob}\,(L|s = L)\,.\end{aligned} \tag{6.15}$$

A signal H (L) can be received either correctly, in which case the probability of success is actually π_H (π_L), or incorrectly such that the true probability of success is π_L (π_H). It is straightforward to show that $\hat{\pi}_H > \pi > \hat{\pi}_L$ and receiving signal H increases the belief in the probability of success, compared to the average probability of success, while signal L lowers the belief in the probability of success. For this inequality to be strict, we assume that $0 < \nu < 1$ and both types of companies are present in the market.

Combining equations (6.13), (6.14), and (6.15), we can rewrite the probabilities of observing signal H and L, respectively, as

$$\begin{aligned}\text{Prob}\,(s = H) &= \frac{\pi - \hat{\pi}_L}{\hat{\pi}_H - \hat{\pi}_L}, \\ \text{Prob}\,(s = L) &= \frac{\hat{\pi}_H - \pi}{\hat{\pi}_H - \hat{\pi}_L}.\end{aligned} \tag{6.16}$$

Having established the beliefs of banks on the probabilities of companies being successful, depending on the signal the banks receive about them, if any, we can now continue to assess the impact such information has on banks' profits from providing loans to companies and then assess the equilibrium loan rates that emerge with competitive banks.

Pure strategy equilibrium Let us assume there are two types of banks, one is informed by receiving a signal s, while the other type of bank receives no such signal and is hence referred to as uninformed. The interest charged on loans to a company by uninformed banks will be denoted r_L^U, while those of informed banks are r_L^L and r_L^H, depending whether signal L or signal H has been obtained.

Banks lending to companies make profits of

$$\begin{aligned}\hat{\Pi}_B^U &= \pi(1 + r_L^U)L - (1 + r_D)L, \\ \hat{\Pi}_B^H &= \hat{\pi}_H(1 + r_L^H)L - (1 + r_D)L, \\ \hat{\Pi}_B^L &= \hat{\pi}_L(1 + r_L^L)L - (1 + r_D)L\end{aligned} \tag{6.17}$$

where loans are fully financed by deposits and the deposit rate is denoted r_D. These profits reflect the fact that loans are repaid if the investment of the company is successful, for which they have a belief of a probability of π, $\hat{\pi}_H$, $\hat{\pi}_L$, respectively. If the investment of the company is not successful, the loan cannot be repaid. Banks finance their loans fully by deposits, on which they have to pay interest r_D. Uninformed banks use their inference of the average probability of success in the market, π, while informed banks will use the probability of success depending on the signal they have received, $\hat{\pi}_H$ and $\hat{\pi}_L$, respectively. Banks are profitable when providing loans if $\hat{\Pi}_B^i \geq 0$, which requires that

$$1 + r_L^U \geq 1 + \hat{r}_L^U = \frac{1 + r_D}{\pi}, \tag{6.18}$$
$$1 + r_L^H \geq 1 + \hat{r}_L^H = \frac{1 + r_D}{\hat{\pi}_H},$$
$$1 + r_L^L \geq 1 + \hat{r}_L^L = \frac{1 + r_D}{\hat{\pi}_L},$$

where we easily see that $\hat{r}_L^H < \hat{r}_L^U < \hat{r}_L^L$ as $\hat{\pi}_H > \pi > \hat{\pi}_L$. Hence, a company for which the bank has received signal H has the lowest loan rate to break even, reflecting its high probability of success, while companies for which signal L is received are requiring the highest loan rate. Uninformed banks will require an intermediate loan rate to break even due their inability of distinguishing the two types of companies.

We can now evaluate the equilibrium loan rates of the informed and uninformed bank. To this effect we distinguish four cases. Let us initially assume that loan rates are set such that $r_L^U \leq \min\left\{r_L^H, r_L^L\right\}$. In this case the uninformed bank sets the lowest loan rate and will consequently provide loans to all companies, regardless of the signal the informed bank obtains for the company. With the informed bank not providing any loans, it will make no profits. Assuming the uninformed bank sets a loan rate that is profitable such that $r_L^U \geq \hat{r}_L^U$, the informed bank could set a lower loan rate $r_L^H \leq r_L^U$ that is still above its break even threshold of $\hat{r}_L^H$, given that $\hat{r}_L^H < \hat{r}_L^U$ and attract all companies for which it receives the signal H, making a profit. Hence we find that $r_L^H < r_L^U$, violating the requirement that $r_L^U \leq \min\left\{r_L^H, r_L^L\right\}$ and hence this arrangement of loan rates cannot be an equilibrium.

As the second case let us consider an equilibrium with $r_L^U \geq \max\left\{r_L^H, r_L^L\right\}$. In this case the informed bank provides the loan to companies, regardless of the signal they receive about them and the uniformed bank does not provide any loans and make zero profits. With lending to companies for which the informed bank received signal L assumed to be profitable, otherwise the informed bank would not be willing to lend to them at the stated loan rate, the uninformed bank can lower its loan rate to below $\hat{r}_L^L$ and provide loans to these companies. Theinformed bank cannot compete at that level as it would be making a loss, but for the uninformed bank their threshold of being profitable is lower, $\hat{r}_L^U < \hat{r}_L^L$, and they are able to make a profits, preferring such a loan rate strategy. This then implies that $r_L^U < r_L^L$ and hence the second possible arrangement of loan rates cannot be an equilibrium.

For the third case consider that the equilibrium satisfies $r_L^L < r_L^U < r_L^H$. The uninformed bank provides the lower loan rate for companies for which the informed bank has received signal H and the informed bank provides loans to those it receives signal L for as it charges the lower loan rate. If the informed bank makes profits from lending to companies for which they have received signal L, the uninformed bank can reduce its loan rate below r_L^L and still make a profit as $\hat{r}_L^U < \hat{r}_L^L$. The uniformed bank would provide loans to all companies, increasing their profits. Thus we would find that $r_L^U < r_L^L$, and hence the arrangement $r_L^L < r_L^U < r_L^H$ cannot be an equilibrium.

The final possible arrangement considers an equilibrium that requires $r_L^H < r_L^U < r_L^L$. Here the uninformed bank would provide loans to companies for which the informed bank has received signal L, while the informed bank will provide loans for which it has obtained signal H. If the uninformed bank knows that it will lend only to companies for which the informed bank has received signal L, it knows that it will have to charge at least $\hat{r}_L^L$ to be profitable. Competing with the informed bank for these companies would induce the informed bank to also charge $\hat{r}_L^L$ and hance $r_L^U = r_L^L = \hat{r}_L^L$, violating the condition $r_L^H < r_L^U < r_L^L$ in this arrangement and ruling it out as an equilibrium.

We can conclude that no equilibrium loan rate in pure strategies exists and the only possible equilibrium is in mixed strategies, which we will consider next.

Mixed strategy equilibrium In mixed strategies, banks will randomize the loan rates they are setting according to a distribution function, which we define as

$$\begin{aligned} \lambda_U(r) &= \text{Prob}\left(r_L^U < r\right), \\ \lambda_H(r) &= \text{Prob}\left(r_L^H < r\right), \\ \lambda_L(r) &= \text{Prob}\left(r_L^L < r\right), \end{aligned} \tag{6.19}$$

for the uninformed and informed banks with signals H and L, respectively. Thus $\lambda_i(r)$ denotes the probability that the loan rate they offer is below r.

The profits of the uninformed bank are given by the value of the loan after having repaid their depositors, $\left(\hat{\pi}_s\left(1 + r_L^U\right) - (1 + r_D)\right) L$, where the best belief will depend on the signal the informed bank has received. If the informed bank has received signal H (L), which occurs with probability Prob$(s = H)$ (Prob$(s = L)$), they have to offer loan rates that are below those of the informed bank. With the uninformed bank offering r_L^U, the probability of the informed bank offering a lower loan rate is $\lambda_H\left(r_L^U\right)$ $(\lambda_L\left(r_L^U\right))$ and hence the uninformed bank will offer a lower loan rate with probability $1 - \lambda_H\left(r_L^U\right)$ $(1 - \lambda_L\left(r_L^U\right))$.

For informed banks, knowing the signal they have received, their profits will be the value of the loan after having repaid their depositors, $\left(\hat{\pi}_s\left(1 + r_L^U\right) - (1 + r_D)\right) L$, provided they offer the lowest loan rate. The uninformed bank will offer a lower loan rate than r_L^s with probability $\lambda_U\left(r_L^s\right)$, such that the informed bank will offer the lower loan rate with probability $1 - \lambda_U\left(r_L^s\right)$.

The bank profits are thus given by

$$\begin{aligned} \Pi_B^U &= \text{Prob}(s = H)\left(1 - \lambda_H\left(r_L^U\right)\right)\left(\hat{\pi}_H\left(1 + r_L^U\right) - (1 + r_D)\right) L \\ &\quad + \text{Prob}(s = L)\left(1 - \lambda_L\left(r_L^U\right)\right)\left(\hat{\pi}_L\left(1 + r_L^U\right) - (1 + r_D)\right) L \\ \Pi_B^L &= \left(1 - \lambda_U\left(r_L^L\right)\right)\left(\hat{\pi}_L\left(1 + r_L^L\right) - (1 + r_D)\right) L, \\ \Pi_B^H &= \left(1 - \lambda_U\left(r_L^H\right)\right)\left(\hat{\pi}_H\left(1 + r_L^H\right) - (1 + r_D)\right) L. \end{aligned} \tag{6.20}$$

Let us first assess the case where $\hat{\pi}_L\left(1+R\right)L-(1+r_L^L)L\leq 0$ and companies for which the informed bank receives signal L would produce a loss to informed banks as they are not able to repay their loans. In this case informed banks would never offer a loan to this company and therefore $\lambda_L\left(r_L^U\right)=0$ and $\lambda_U\left(r_L^L\right)=1-\lambda_L\left(r_L^U\right)=1$, implying that uninformed bank will provide the loan to this company. As uninformed banks cannot observe the signal L, they cannot refuse to provide a loan, as that would imply they have to refuse to provide a loan to all companies, including those for which signal H was received by the informed bank.

If in contrast $\hat{\pi}_L\left(1+R\right)L-(1+r_L^L)L\geq 0$, companies for which the informed bank receives signal L would be able to repay their loans and informed as well as uninformed banks would be able to make a profit from lending. While the informed bank would have to charge at least $\hat{r}_L^L$ to break even, the uninformed bank will be able to charge a lower loan rate as they will break if they charge at least $\hat{r}_L^U<\hat{r}_L^L$ and thus obtain all loans of this type of company. This gives is again that informed banks do not provide a loan, $\lambda_L\left(r_L^U\right)=0$, while uninformed banks provide all such loans, $\lambda_U\left(r_L^L\right)=1-\lambda_U\left(r_L^L\right)=1$.

In competition with informed banks, uninformed banks are at a disadvantage and informed banks can always extract all profits from their less informed competitors such that $\Pi_B^U=0$. Given we consider mixed strategy equilibria, we know that for all loan rates it may use, the profits of the informed bank must be equal. The informed bank sets the loan rate of the company for which it receives signal H at the loan rate the uninformed bank would set to break even, $1+r_L^H=\hat{r}_L^U=\frac{1+r_D}{\pi}$; this is because any higher loan rate could be undercut by the uninformed bank. Inserting this loan rate into equation (6.17), we get $\Pi_B^H=\left(\frac{\hat{\pi}_H}{\pi}-1\right)(1+r_D)L$. We can neglect the term $1-\lambda_U\left(r_L^H\right)$ from equation (6.17) as the informed bank can marginally undercut the uninformed bank and thus providing all loans to these companies. Hence, inserting for $\lambda_L\left(r_L^U\right)=0$ and $\lambda_U\left(r_L^L\right)=1$ in the case the informed bank receives signal L, the bank profits become

$$
\begin{aligned}
\Pi_B^H &= \left(1-\lambda_U\left(r_L^H\right)\right)\left(\hat{\pi}_H\left(1+r_L^H\right)-(1+r_D)\right)L \qquad (6.21)\\
&= \left(\frac{\hat{\pi}_H}{\pi}-1\right)(1+r_D)L,\\
\Pi_B^L &= 0,\\
\Pi_B^U &= \mathrm{Prob}(s=H)\left(1-\lambda_H\left(r_L^U\right)\right)\left(\hat{\pi}_H\left(1+r_L^U\right)-(1+r_D)\right)L\\
&\quad +\mathrm{Prob}(s=L)\left(\hat{\pi}_L\left(1+r_L^U\right)-(1+r_D)\right)L\\
&= 0.
\end{aligned}
$$

Using equation (6.16), these expressions can be solved for the probability distribution of loan rates, which take the form

$$\lambda_U\left(r_L^H\right)=\frac{\hat{\pi}_H}{\pi}\frac{\pi\left(1+r_L^H\right)-(1+r_D)}{\hat{\pi}_H\left(1+r_L^H\right)-(1+r_D)},\tag{6.22}$$
$$\lambda_H\left(r_L^U\right)=\frac{\hat{\pi}_H-\hat{\pi}_L}{\pi-\hat{\pi}_L}\frac{\pi\left(1+r_L^U\right)-(1+r_D)}{\hat{\pi}_H\left(1+r_L^U\right)-(1+r_D)}.$$

Here $\lambda_U\left(r_L^H\right)$ and $\lambda_H\left(r_L^U\right)$ characterise the probability distribution of the loan rates in equilibrium, noting that the informed bank will never provide a loan to companies for which it receives signal L, and thus this loan rate being neglected in the further analysis.

Having established the equilibrium distribution of the loan rates as offered by informed and uninformed banks, we can continue to analyse the impact the precision of information γ has on this outcome.

The impact of information Inserting from all expressions into equation (6.15), it is easy to see how the belief on the success rate changes as the precision of information changes. We obtain

$$\frac{\partial\hat{\pi}_H}{\partial\gamma}=\frac{\nu(1-\nu)(\pi_H-\pi_L)}{(\gamma\nu+(1-\gamma)(1-\nu))^2}>0,\tag{6.23}$$
$$\frac{\partial\hat{\pi}_L}{\partial\gamma}=\frac{\nu(1-\nu)(\pi_L-\pi_H)}{((1-\gamma)\nu+\gamma(1-\nu))^2}<0,$$

where the inequality arises from our assumption that $\pi_L<\pi_H$.

The more precise the information becomes, γ, the more the belief of the informed bank moves from π towards π_i, thus the belief increases for companies for which signal H has been received and reduces for companies for which signal L has been received. It similar is straight forward to show that $\frac{\partial\lambda_U\left(r_L^H\right)}{\partial\hat{\pi}_H}=-\frac{\pi(1+r_H)-(1+r_D)}{(\hat{\pi}_H(1+r_H)-(1+r_D))^2}\frac{1+r_D}{\pi}<0$ and $\frac{\partial\lambda_H\left(r_L^U\right)}{\partial\hat{\pi}_L}=\frac{\pi\left(1+r_L^H\right)-(1+r_D)}{\hat{\pi}_H\left(1+r_L^H\right)-(1+r_D)}\frac{\hat{\pi}_H-\pi}{(\pi-\hat{\pi}_L)^2}>0$, which when applying the chain rule gives us

$$\frac{\partial\lambda_U\left(r_L^H\right)}{\partial\gamma}=\frac{\partial\lambda_U\left(r_L^H\right)}{\partial\hat{\pi}_H}\frac{\partial\hat{\pi}_H}{\partial\gamma}<0,\tag{6.24}$$
$$\frac{\partial\lambda_H\left(r_L^U\right)}{\partial\gamma}=\frac{\partial\lambda_H\left(r_L^U\right)}{\partial\hat{\pi}_L}\frac{\partial\hat{\pi}_L}{\partial\gamma}<0.$$

Hence, the more precise the signal is, γ, the less likely it is that the uninformed bank will provide the lowest loan rate to the company for which the informed bank has obtained signal H, while the informed bank is more likely to provide the lowest loan rate. The company for which the informed bank received signal L will only ever obtain a loan from the uninformed bank, regardless of the precision of the signal the informed bank obtains. It is thus that with a more precise signal the informed bank is more likely to provide a loan to the company for which it has obtained signal

H, giving rise to an increased market share in lending. Thus a bank obtaining more precise information is providing more loans.

Increasing the market share itself will not necessarily increase the profits of the informed bank, as the loan rate they can obtain may well reduce. We therefore analyse the impact more precise information has on the loan rate an informed bank will obtain.

A company for which the informed bank obtains signal H, will always be offered a loan, either by the informed bank or by the uninformed bank. The loan rate the company will pay is given is the lower of the two loan rates offered by the informed and uninformed bank, We thus have the loan rate actually paid by the company given as

$$\hat{r}_L^H = \min\left\{r_L^H, r_L^U\right\}. \tag{6.25}$$

We know from order statistics that the distribution of $\hat{r}_L^H$ can be obtained as

$$1 - \hat{\lambda}_H\left(\hat{r}_L^H\right) = \left(1 - \lambda_H\left(r_L^U\right)\right)\left(1 - \lambda_U\left(r_L^H\right)\right), \tag{6.26}$$

where $\hat{\lambda}_H\left(\hat{r}_L^H\right)$ denotes the probability that the loan rate the company pays is below $\hat{r}_L^H$, and hence $1 - \hat{\lambda}_H\left(\hat{r}_L^H\right)$ can be interpreted as the probability that the loan rate the company pays is above $\hat{r}_L^H$.

Using the results from equation (6.24), we easily obtain

$$\begin{aligned}\frac{\partial(1 - \hat{\lambda}_H\left(\hat{r}_L^H\right))}{\partial\gamma} &= -\frac{\partial\lambda_H\left(r_L^U\right)}{\partial\gamma}\left(1 - \lambda_U\left(r_L^H\right)\right) \\ &\quad -\frac{\partial\lambda_U\left(r_L^H\right)}{\partial\gamma}\left(1 - \lambda_H\left(r_L^U\right)\right) \\ &> 0.\end{aligned} \tag{6.27}$$

This can be interpreted that with more precise information, γ, the probability of observing a loan rate above $\hat{r}_L^H$ increases. The expected loan rate, the expected value of $\hat{r}_L^H$, is given by

$$E[\hat{r}_L^H] = \int_0^{+\infty} \hat{r}_L^H\left(1 - \hat{\lambda}_H\left(\hat{r}_L^H\right)\right) d\hat{r}_L^H. \tag{6.28}$$

This expression is increasing in the precision of the signal, γ. This is as $1 - \hat{\lambda}_H\left(\hat{r}_L^H\right)$ is increasing in this precision as shown in equation (6.27) and hence its integral has to be increasing. The consequence of this result is that informed banks having more precise information on a company, allows them a bigger informational advantage and this can be exploited by increasing the loan rate.

Hence, more precise information does not only make it more likely that the informed bank will provide a loan to the companies for which it receives signal H, but it will also be able to charge a higher loan rate. This will increase profits to the informed bank not only by providing more loans, but also by providing these at higher loan rates. Thus obtaining more precise information is beneficial to banks.

Summary Banks acquiring information on the companies they are potentially lending to, gain an advantage over banks without such information. This informational advantage allows informed banks not only to obtain a larger share of the market for the more profitable low-risk companies, but they are also able to increase loan rates by extracting more surplus from companies, further improving their profits. The larger market share of informed banks arises from their ability to better identify companies where lending is highly profitable, low-risk companies that have high probabilities of success in their investments and hence a high likelihood of repaying their loans. Having identified such companies, informed banks are able to offer better loan conditions to these companies. Even though loan conditions are more attractive to such companies, the better knowledge about them, allows informed banks to extract more surplus when lending to them, resulting in a higher loan rate, relative to the risks associated with lending to these companies. This will allow informed banks to generate more profits. These higher profits from acquiring information have, of course, to be balanced against the costs of obtaining this information and an optimal level of information precision will be achieved where the marginal benefits, as discussed here, equal the marginal costs of information acquisition.

Gaining access to more precise information allows banks to strengthen their market position by providing loans to more low-risk companies, those with high success rates, and at the same time increase their profitability by increasing the loan rates they are charging. This result shows the pivotal role of information for banks in the loan market to retain and improve their competitiveness. It also gives us insights into the way a bank can defend itself against competition from existing banks or new entrants to the market, it needs to retain its informational advantage. Consequently, those banks who are less informed need to increase the precision of the information they have access to. This might, of course, lead to a never-ending race to acquire ever more precise information in order to remain competitive. It is likely that currently well-informed bank will have to react to the increased information their competitors have obtained by themselves increasing the precision of information they hold. This will then induce the less well-informed banks to increase the precision of their information, leading to a renewed reaction of the better informed banks, until an equilibrium in the level of information precision has been reached. Such an arms race in information acquisition will most likely result in a level of information precision that, while optimal for the competing banks, is socially sub-optimal. The level of information precision will be too high.

Reading Hauswald & Marquez (2006)

6.3 Debt maturity

It is common to assume that investments by companies only yield an outcome over multiple time periods, exceeding the time length that depositors seek to commit themselves to not withdraw any of their funds. One role of banks is to enable loans to be provided whose terms match that of the investment companies make; this liquidity provision of banks to depositors had been discussed in chapter 4. However, there is no requirement for companies to seek such long-term loans, instead they could rely on rolling over short-term loans. Similarly a long-term loan could be sought to finance a sequence of short-term investments, rather than obtaining a new short-term loan for each individual investment. It is this latter scenario that we discuss here.

To analyse the optimal time to maturity of a loan, assume a company makes investments lasting a single time period; identical investments can be conducted in each time period. There are two types of companies, one with a high probability of success of their investments, π_H, and one with a low probability of success of their investments, $\pi_L < \pi_H$. If the investment is successful, it returns an outcome V that will cover the loan repayment, and no revenue is received in the case the investment fails. The bank cannot distinguish between these two types of companies, but the companies know the probabilities of success of their investments.

We can now analyse the profits of banks and companies choosing long-term and short-term loans, respectively, where we only consider an economy with two time periods.

Short-term loans If companies obtain a loan for a single time period, matching the investment length, the profits of a company of type i is given by

$$\begin{aligned} \Pi_C^i &= \pi_i \left((V - (1 + r_L)L) + \pi_i \left(V - (1 + r_L) L \right) \right) \\ &= \pi_i \left(1 + \pi_i \right) \left(V - (1 + r_L) L \right), \end{aligned} \tag{6.29}$$

where r_L denotes the interest charged on the loan $L < V$. The investment is successful in time period 1 with probability π_i and the company has to repay the loan from the revenue V this generates. The company then can continue with a further investment, also succeeding with probability π_i. We assume that if the initial investment is not successful, the company cannot continue with further investments as they have defaulted on their obligations of the first loan and are excluded from further borrowing.

With a deposit rate of r_D, the bank profits in each time period, when lending to a company of type i, are then given as

$$\Pi_B^i = \pi_i \left(1 + r_L \right) L - \left(1 + r_D \right) D. \tag{6.30}$$

where we assume that the deposits D are fully invested into loans, such that $D = L$. If the investment is successful, the loan is repaid and we assume banks have unlimited liability and will always be able to repay depositors. We consider the lending in each time period as competition between banks allows companies to switch banks after the first loan, necessitating banks to break even with a single loan provided. Perfect competition between banks implies that $\Pi_B^i = 0$ and hence

$$1 + r_L = \frac{1 + r_D}{\pi_i}. \tag{6.31}$$

As the bank cannot distinguish between the two types of companies, it will have to set a single loan rate that both types of companies have access to. Assume for now that the company chooses to set the loan rate such that it would break for the company with the low success rate, π_L, and hence it sets $1 + r_L = \frac{1+r_D}{\pi_L}$. In this case, if a company with a high success rate π_H seeks such a loan, the bank would make a profit as we can easily verify from equation (6.30) after inserting all relevant variables.

Inserting equation (6.31) into equation (6.29), company profits are given by

$$\Pi_C^i = \pi_i (1 + \pi_i) V - \frac{\pi_i (1 + \pi_i)}{\pi_L} (1 + r_D) L. \tag{6.32}$$

As an alternative to the short-term loan considered here, banks can also offer a long-term loan that can be used to finance both investments, which we consider next.

Long-term loans Rather than two short-term loans, each for a single time period, the bank could offer a long-term loan covering both time periods. The loan rate of this long-term loan will be denoted by $\hat{r}_L$. We assume that if the investment fails in the first time period, the company has not sufficient funds to finance an investment in the second time period as it has used all funds provided by the loan for the failed investment. Furthermore, if the company fails in the second time period, we assume that it does not have enough funds available to repay the loan fully. It will have retained $V - L$ from receiving the successful outcome in the first time period, V, and having used L of that to re-invest into the investment of the second time period; hence we assume that $V - L < (1 + \hat{r}_L) L$, or $V < (2 + \hat{r}_L) L$, and the company would default on the loan, only repaying $V - L$. We thus here assume that companies have to use any remaining proceeds from the first investment to repay the loan in the second time period; in contrast to that we allowed companies to retain any such surplus when entering a second short-term loan contract by acknowledging that this was a separate contract and could not bind the company to use previously generated funds to repay this loan.

Hence the company can generate profits of

$$\hat{\Pi}_C^i = \pi_i^2 (2V - L - (1 + \hat{r}_L) L). \tag{6.33}$$

The company retains any profits only if both investments are successful and in this case retains $V - L$ from the initial investment, and the second investment generates V again, before the loan is repaid.

With a deposit rate of $1 + \hat{r}_D = (1 + r_D)^2$ to account for the accumulated interest over two time periods, the profits of the bank lending to company of type i are given by

$$\hat{\Pi}_B^i = \pi_i^2 \left(1 + \hat{r}_L\right) L + \pi_i \left(1 - \pi_i\right) (V - L) - \left(1 + \hat{r}_D\right) L. \tag{6.34}$$

The first term denotes the case where the investment is successful in both time periods and the company repays the loan in full, while the second term denotes the case where the investment is successful only in time period 1 and the company is required to repay the retained profits that time period, $V - L$. If the company fails in time period 1, it has no funds to repay the loan as no revenue was generated an a second investment could not be made, given that no funds were left to invest. In addition, due to unlimited liability, the bank has to repay its depositors.

Perfect competition between banks requires $\hat{\Pi}_B^i = 0$ and hence we obtain

$$1 + \hat{r}_L = \frac{1 + \hat{r}_D}{\pi_i^2} - \frac{1 - \pi_i}{\pi_i} \frac{V - L}{L}. \tag{6.35}$$

As the bank cannot distinguish between the two types of companies, it will have to set a single loan rate that both types of companies have access to. Assume for now that the company chooses to set the loan rate such that it would break even for the company with the high success rate, π_H, and thus $1 + \hat{r}_L = \frac{1+\hat{r}_D}{\pi_H^2} - \frac{1-\pi_H}{\pi_H} \frac{V-L}{L}$. Inserting this result into equation (6.33), we can obtain the company profits as

$$\hat{\Pi}_C^i = \frac{\pi_i^2}{\pi_H^2} \left(\pi_H \left(1 + \pi_H\right) V - \left(\left(1 + \hat{r}_D\right) + \pi_H\right) L\right). \tag{6.36}$$

Having established these profits now allows us to compare the profits of the companies from obtaining short-term and long-term loans, respectively.

Optimal loan terms In order for the company with the high success rate, π_H, to prefer long-term loans, we require that $\hat{\Pi}_C^H \geq \Pi_C^H$ and after inserting from equations (6.32) and (6.36), while noting that $1 + \hat{r}_D = (1 + r_D)^2$, this requirement can be written as

$$\pi_L \leq \pi_L^* = \frac{\pi_H \left(1 + \pi_H\right) \left(1 + r_D\right)}{\left(1 + r_D\right)^2 + \pi_H} < \pi_H. \tag{6.37}$$

The long-term loan is more attractive to companies with high success rates if the loan rate for short-term loans is sufficiently high, which requires a low success rate of the other type of companies, as we can easily see from equation (6.31). The reason that long-term loans might be less attractive than short-term loans despite the loan rate being lower, is that with long-term loans the profits from having an initially successful investment are used to repay parts of the loan if the subsequent investment is unsuccessful. This increases the repayment of loans compared to short term loans, where we assumed that banks cannot obtain any previous surplus the company retained. Hence the loan rate for short-term loans must be sufficiently higher to make long-term loans more attractive.

Similarly, for a company with a low success rate, π_L, to prefer short term loans we require that $\hat{\Pi}_C^L \leq \Pi_C^L$ and after inserting from equations (6.32) and (6.36), while noting that $1 + \hat{r}_D = (1 + r_D)^2$, this becomes

$$\begin{aligned} &\pi_L^2 (1 + r_D)^2 + \pi_L^2 \pi_H - \pi_H^2 (1 + \pi_H)(1 + r_D) \\ &\quad + \pi_H \pi_L (\pi_H - \pi_L) \frac{V}{L} \geq 0. \end{aligned} \tag{6.38}$$

This condition is fulfilled for $\pi_H = \pi_L$ and as we reduce π_L, the expression reduces as we can easily verify by differentiating the left-hand side. There will be a value for π_L, where this condition is no longer fulfilled, as for $\pi_L = 0$ it is violated. This result then implies that the difference between high and low success rates cannot be too large. Combining the conditions in equations (6.37) and (6.38), we require that

$$\begin{aligned} &\frac{\pi_H^2}{\pi_L} (1 + \pi_H)(1 + r_D) - \pi_H (\pi_H - \pi_L) \frac{V}{L} \\ &\quad \leq \pi_L \left((1 + r_D)^2 + \pi_H \right) \leq \pi_H (1 + \pi_H)(1 + r_D), \end{aligned}$$

which admits a viable solution only if

$$\pi_L \geq \pi_L^{**} = (1 + \pi_H)(1 + r_D) \frac{L}{V}. \tag{6.39}$$

As we also require that $\pi_L^{**} < \pi_H$, this is only a possible solution if $\pi_H > \pi_H^{**} = \frac{(1+r_D)\frac{L}{V}}{1-(1+r_D)\frac{L}{V}}$. Thus the probability of success must be sufficiently high for companies with low success rates and high success rates to prefer loans with different maturities. Offering short-term and long-term loans will allow banks to separate companies with high and low probabilities of success, thus separating low-risk and high-risk companies, provided these conditions are met.

The probability of success of the low-risk company, π_H, has to be sufficiently high to ensure that the costs from higher loan repayments with long-term loans in case the second investment fails, are unlikely to occur. In addition, the probability of success of the high-risk company, π_L, cannot be too low, as a detailed analysis of equation (6.38) shows, because in that case the advantage of the lower long-term loan rate that is offered to the high-risk company, compared to the short-term loan rate, outweighs the additional repayment that need to be made with long-term loans.

Hence, if the probabilities of success are sufficiently high and the differences between the success rates of companies are not too large, low-risk companies will seek long-term loans, while high-risk companies will seek short-term loans. Both conditions are likely met for a wide range of companies as banks usually only provide loans to companies that have low default rates and differences in the risks between companies in the loan book of banks are in most cases not substantial. Thus by offering loans of different maturities, banks can distinguish between companies with different risks, reducing the adverse selection of not being able to directly assess this property of companies.

Summary Low-risk companies, those high success rates of investments, will seek long-term loans while high-risk companies, those offering lower success rates, will prefer short-term loans. This is the result of banks offering long-term loans that can recover surplus from previous time periods until the maturity of the loan. With companies showing higher failure rates, such loss recoveries are more common and therefore the loan costs to high-risk companies, including such recoveries, are high. This makes short-term loans, where competition makes the recovery of initial losses impossible, more attractive to such high-risk companies.

We thus see that some companies might prefer to obtain long-term loans to finance a sequence of short-term investments. They do so in order to distinguish themselves from companies that have higher risks and take advantage of lower loan rates that are offered to such companies. For banks the advantage is that it reduces adverse selection as they can identify low-risk and high-risk companies from their choice of loan maturity, which might well enable them to expand their lending by offering more favourable conditions to those exhibiting low risks.

Reading Webb (1991)

6.4 Seniority structure of loans

It is common to assume that an investment is fully financed by a loan from a single bank. In reality, however, companies often seek loans from multiple lenders, in addition to equity, and in some instances these loans have a different level of seniority. Loans of a higher seniority (senior loans) have priority when the company fails to repay their loans fully by being able to make a claim on the remaining assets of the company first; claims arising from loans of lower seniority (subordinated loans) are only met once senior loans have been fully repaid.

Let us assume that a company seeks to finance an investment of size I with a combination of loans L and equity E. The cost of equity is given by r_E and there are two possible banks that have funding costs from deposits by having to pay deposit rates r_D^1 and r_D^2, respectively, depending on the type of bank, such that $0 \leq r_D^1 \leq r_D^2 \leq r_E$. Such differences in funding costs might arise if banks have access to different types of depositors, where the depositors of one bank are willing to accept lower deposit rates than that of the other bank. The outcome of the investment by the company is generating a value V with $0 \leq V \leq \overline{V}$, that has a uniform distribution on this interval.

In case that the company cannot repay the loan, it gets audited by the bank to verify the outcome the company claims to have achieved. Banks incur fixed auditing costs C_i with $\overline{V} \geq C_1 \geq C_2 \geq 0$. Thus the bank that has lower funding costs, r_D^1, faces higher auditing costs, C_1. Higher auditing costs might arise if a bank is less familiar with the company, the region, or the industry. In this sense, each bank has its competitive advantage, with one bank having lower funding costs, while the other bank has lower auditing costs.

Using this framework, we can now determine the optimal financing policy of a company, relying on a loan from single bank, or on both banks, either of equal seniority of with one bank providing a senior loan and the other a subordinated loan.

Single lender Let us firstly assume that the company only borrows from a single bank. The bank profits consist of the (partial) repayment of the loan when defaulting, the repayment of the loan including interest r_L^i if repaying in full, less the cost of auditing and the funding costs of the loan L_i, where we assume that the bank fully funds the loan by deposits such that $D_i = L_i$. Thus we obtain

$$\begin{aligned}\Pi_B^i = & \int_0^{\left(1+r_L^i\right)L_i} V dF(V) + \int_{\left(1+r_L^i\right)L_i}^{\overline{V}} \left(1+r_L^i\right) L dF(V) \\ & - \int_0^{\left(1+r_L^i\right)L_i} C_i dF(V) - \left(1+r_D^1\right) L_i \\ = & \frac{\overline{V}-C_i}{\overline{V}}\left(1+r_L^i\right)L_i - \frac{\left(1+r_L^i\right)^2 L_i^2}{2\overline{V}} - \left(1+r_D^1\right)L_i,\end{aligned} \tag{6.40}$$

using our assumption that the outcome of the investment, V, is uniformly distributed in $\left[0;\overline{V}\right]$.

The bank maximizes its profits by choosing an optimal loan rate such that

$$\frac{\partial \Pi_B^i}{\partial\left(1+r_L^i\right)} = \frac{\overline{V}-C_i}{\overline{V}} L_i - \frac{\left(1+r_L^i\right)L_i^2}{\overline{V}} = 0, \tag{6.41}$$

which solves for

$$1 + r_L^i = \frac{\overline{V}-C_i}{L_i}. \tag{6.42}$$

If we assume that banks are competitive, they will make zero profits. Thus with $\Pi_B^i = 0$ and inserting equation (6.42) into equation (6.40), we easily get the loan the bank provides to be of size

$$L_i = \frac{\left(\overline{V}-C_i\right)^2}{2\overline{V}\left(1+r_D^i\right)}. \tag{6.43}$$

The company will retain any surplus after repaying the loan in full, which is reduced by the cost of equity on the part of the investment that is not financed by a loan, thus $E = I - L_i$. The company profits are then obtained from

$$\begin{aligned}\Pi_C^i = & \int_{\left(1+r_L^i\right)L_i}^{\overline{V}} \left(V - \left(1+r_L^i\right)L_i\right) dF(V) - \left(1+r_E\right)\left(I - L_i\right) \\ = & \frac{C_i^2}{2\overline{V}} + \frac{\left(\overline{V}-C_i\right)^2\left(1+r_E\right)}{2\overline{V}\left(1+r_D^i\right)} - \left(1+r_E\right)I,\end{aligned} \tag{6.44}$$

after inserting from equations (6.42) and (6.43) and using our assumption that the outcome of the investment, V, is uniformly distributed in $\left[0;\overline{V}\right]$..

When not taking a loan, the investment is fully financed by equity and the company makes profits of

$$\Pi_C^0 = \int_0^{\overline{V}} V dF(V) - (1 + r_E) I = \frac{1}{2}\overline{V}^2 - (1 + r_E) I. \tag{6.45}$$

Comparing the profits of taking out a loan in equation (6.44) and financing the investment entirely from equity in equation (6.45), we see that the company takes out a loan only if $\Pi_C^i \geq \Pi_C^0$, which simplifies to become

$$1 + r_D^i \leq 1 + r_D^{i*} = \frac{\overline{V} - C_i}{\overline{V} + C_i}(1 + r_E), \tag{6.46}$$

which we assume to be fulfilled for both banks. To decide which bank to approach, the company would compare the profits in equation (6.44) for the two banks. Bank 2 gets approached if $\Pi_C^2 \geq \Pi_C^1$, which easily becomes

$$\begin{aligned} 1 + r_D^2 &\leq 1 + r_D^{2**} \\ &= \frac{\left(\overline{V} - C_1\right)(1 + r_E)}{\left(\overline{V} - C_1\right)(1 + r_E) + \left(C_1^2 - C_2^2\right)\left(1 + r_D^1\right)}\left(1 + r_D^1\right). \end{aligned} \tag{6.47}$$

Having established the choice of lender in the case the company takes out a single loan, we can now turn to the case that it seeks a loan from both banks.

Using both lenders The company might now approach both banks for a loan and we initially assume that both loans will have the same seniority and thus any proceeds from the company not being able to repay the loans in full will be divided between the banks according to the size of the outstanding repayments. The company repays each bank $R_i = \left(1 + r_L^i\right) L_i$ and hence the full repayment required is $R = R_1 + R_2$. The amount of borrowing for the company will be $L = L_1 + L_2$ and we define $\alpha_i = \frac{R_i}{R}$ as the fraction of the repayments going to bank i. The profits of bank i are then given by

$$\begin{aligned} \hat{\Pi}_B^i &= \int_0^R \alpha_i V dF(V) + \int_R^{\overline{V}} R_i dF(V) \\ &\quad - \int_0^R C_i dF(V) - \left(1 + r_D^i\right) L_i \\ &= \alpha_i \frac{2\overline{V}R - R^2}{2\overline{V}} - \frac{RC_i}{\overline{V}} - \left(1 + r_D^i\right) L_i. \end{aligned} \tag{6.48}$$

The first term denotes the fraction of the outcome of the investment the bank obtains if the company does not repay the loans in full, the second term the full loan repayment, the fourth term encompass the auditing costs, and the final term the funding costs of the loan.

The first order condition for the profit maximum of the bank is then given by

$$\frac{\partial \hat{\Pi}_B^i}{\partial R_i} = \frac{\partial \hat{\Pi}_B^i}{\partial R}\frac{\partial R}{\partial R_i} = \alpha_i \frac{\overline{V} - R}{\overline{V}} - \frac{C_i}{\overline{V}} = 0, \tag{6.49}$$

implying

$$R = \frac{\alpha_i \overline{V} - C_i}{\alpha_i}. \tag{6.50}$$

As this relationship holds for both banks, setting it equal for $i = 1$ and $i = 2$, noting that $\alpha_2 = 1 - \alpha_1$, the proportion α_i is given as

$$\alpha_i = \frac{C_i}{C_1 + C_2}. \tag{6.51}$$

Inserting equations (6.50) and (6.51) into equation (6.48), we get the bank profits as

$$\hat{\Pi}_B^i = \frac{C_i}{C_1 + C_2} \frac{\left(\overline{V} - (C_1 + C_2)\right)^2}{2\overline{V}} - \left(1 + r_D^i\right) L_i. \tag{6.52}$$

Again, perfect competition between banks implies $\hat{\Pi}_B^i = 0$ and hence the optimal loan size is given by

$$L_i = \frac{C_i}{1 + r_D^i} \frac{\left(\overline{V} - (C_1 + C_2)\right)^2}{2\overline{V}\left(C_1 + C_2\right)}. \tag{6.53}$$

Company profits are then given from the outcome of the investment that has been retained after repaying the loans, less the costs of equity from the part of the investment not financed by loans. After inserting for all variables from the expressions above, we obtain

$$\begin{aligned}\hat{\Pi}_C &= \int_R^{\overline{V}} (V - R)\, dF(V) - (1 + r_E)\left(I - (L_1 + L_2)\right) \qquad (6.54)\\ &= \frac{(C_1 + C_2)^2}{2\overline{V}} - (1 + r_E) I \\ &\quad + \left(\frac{C_1}{\left(1 + r_D^1\right)} + \frac{C_2}{\left(1 + r_D^2\right)}\right) \frac{\left(\overline{V} - (C_1 + C_2)\right)^2}{2\overline{V}\left(C_1 + C_2\right)} (1 + r_E).\end{aligned}$$

It is tedious but possible to show that with the constraint $C_1 + C_2 \leq \overline{V}$, two banks providing loan of equal seniority is never optimal and will be dominated by using a single lender as we find that $\hat{\Pi}_C \leq \Pi_C^i$. The reason is that with two lenders auditing costs are incurred by both banks, increasing the overall costs as the loans each provides will be smaller. As there are no other benefits from having to otherwise equal lenders, these additional costs provide the banks with no benefits to compensate for these higher costs. Thus loan rates will be higher than when taking out a single loan only.

With splitting the loan between two banks and offering equal seniority to each bank not being beneficial, we will now investigate whether the use of a subordinated loan can increase the profits of companies.

Subordinated loan Rather than treating both banks equal, the company could assign seniority to bank i, i.e. its claims get paid in full before those of bank j are considered. The profits for the senior bank remain unchanged as it is irrelevant for this bank whether subordinate claims get paid out once it has received full payment. Hence the loan rates and loan amounts in equations (6.42) and (6.43) remain unaffected.

The subordinate bank j only gets repaid its loan if the senior bank has been paid in full, hence the profits of the bank granting the subordinated loan is given by

$$\begin{aligned}\hat{\hat{\Pi}}_B^j &= \int_{R_i}^{R} (V - R_i)\, dF(V) + \int_R^{\overline{V}} R_j dF(V) \\ &\quad - \int_0^R C_j dF(V) - \left(1 + r_D^j\right) L_j \\ &= -\frac{R_j^2}{2\overline{V}} + \frac{C_i - C_j}{\overline{V}} R_j - \frac{\left(\overline{V} - C_i\right) C_j}{\overline{V}} - \left(1 + r_D^j\right) L_j,\end{aligned} \tag{6.55}$$

when inserting from equations (6.42) and (6.43) for the results of the senior bank. The first order condition for a profit maximum is then given by

$$\frac{\partial \hat{\hat{\Pi}}_B^j}{\partial \left(1 + r_L^j\right)} = -\frac{\left(1 + r_L^j\right) L_j^2}{\overline{V}} + \frac{C_i - C_j}{\overline{V}} L_j = 0, \tag{6.56}$$

solving for

$$1 + r_L^j = \frac{C_i - C_j}{L_j}, \tag{6.57}$$

where we have inserted for $R_i = \left(1 + r_L^i\right) L_i$, $R_j = \left(1 + r_L^j\right) L_j$ and $R = R_i + R_j$. Perfect competition between banks granting subordinated loans implies zero profits, $\hat{\hat{\Pi}}_B^j = 0$, and after inserting from equation (6.57) we obtain from equation (6.55) that

$$L_j = \frac{C_i^2 + C_j^2 - 2\overline{V} C_j}{2\overline{V}\left(1 + r_D^j\right)}. \tag{6.58}$$

From equation (6.57) we see that only for $C_i > C_j$ a realistic solution emerges such that the loan rate is actually positive. Hence bank 1 will be the senior bank, i. e. the bank with the higher auditing costs is the bank granting the senior loan, and bank 2, the bank with lower auditing costs, the bank granting the subordinated loan. The reason for this result is that on the one hand the bank granting the subordinated loan has to audit for a wider range of outcomes, hence the costs are more frequently incurred and thus would be higher if auditing costs were higher; it is thus preferred if the bank with lower auditing costs provides the subordinated loan to reduce costs. In addition, the lower funding costs of bank 1, as measured by the lower deposit rate this bank has to pay, makes their loan less expensive as they will provide a senior loan than that is larger than the subordinated loan, allowing for lower loan rates on this larger loan. The higher auditing costs are spread over a larger amount and the company benefits from the lower funding costs on this larger loan.

The company profits in this situation are then given by

$$\begin{aligned}\hat{\hat{\Pi}}_C &= \int_R^{\overline{V}} \left(V - \left(1 + r_L^1\right) L_1 - \left(1 + r_L^2\right) L_2\right) dF(V) \\ &\quad - (1 + r_E)(I - (L_1 + L_2)) \\ &= \frac{C_2^2}{2\overline{V}} - (1 + r_E) I \\ &\quad + \left(\frac{(V - C_1)^2}{2\overline{V}\left(1 + r_D^1\right)} + \frac{C_1^2 + C_2^2 - 2\overline{V} C_2}{2\overline{V}\left(1 + r_D^2\right)}\right)(1 + r_E),\end{aligned} \tag{6.59}$$

where the second equation is obtained by inserting from equations (6.42), (6.43), (6.57), and (6.58), noting that the senior loan is granted by bank 1 and the subordinate loans is granted by bank 2.

In order to assess whether a company would borrow from a single lender or two lenders with one providing a senior loan and the other a subordinated loan, we need to compare the company profits in equations (6.44) and (6.59). The the of a senior and subordinated loan is preferred to a single loan from bank 2 if $\hat{\hat{\Pi}}_C \geq \Pi_C^2$, which when using equations (6.44) and (6.59), becomes

$$1 + r_D^2 \geq 1 + r_D^{2***} = \frac{\overline{V} + C_1}{\overline{V} - C_1}\left(1 + r_D^1\right). \tag{6.60}$$

Similarly, the use of senior and subordinated loans id preferred to a single loan obtained from bank 1 if $\hat{\hat{\Pi}}_C \geq \Pi_C^1$, which becomes

$$1 + r_D^2 \leq 1 + r_D^{2****} = \frac{C_1^2 + C_2^2 - 2\overline{V} C_2}{C_1^2 - C_2^2}(1 + r_E). \tag{6.61}$$

Hence using two banks with a loans assigned as senior and subordinated, respectively, is preferable to a single lender if

$$\frac{\overline{V}+C_1}{\overline{V}-C_1}\left(1+r_D^1\right) \leq 1+r_D^2 \leq \frac{C_1^2+C_2^2-2\overline{V}C_2}{C_1^2-C_2^2}\left(1+r_E\right). \tag{6.62}$$

From this combined condition, we see that taking two loans from different banks, one a senior loan and the other a subordinated loan, can be optimal for companies if two conditions are fulfilled. This is the case if the funding costs of the bank providing the subordinated loan, bank 2, are sufficiently high compared to that of the bank providing the senior loan, bank 1, as required from equation (6.60). The lower auditing costs bank 2 do not allow the bank to provide a loan of sufficient size on its own and the company would seek a senior loan from the other bank with lower funding costs and thereby increase its profits.

Secondly, if the funding costs of the bank providing the senior loan, bank 1, are not too much lower than that of the bank providing the subordinated loan, bank 2, as required from equation (6.61), then taking two loans is beneficial to the company. In this case the cost advantage in auditing of the bank providing the subordinated loan is sufficient for this bank to provide a loan which has a lower loan rate than the cost of equity, despite being more expensive than the senior loan. It is however cheaper than bank 1 extending its lending and incurring more frequent auditing due to the larger loan that will more often not be repaid in full.

With the senior loan being of the same size as a single loan obtained from bank 1, the company is able extent its lending and increase profits by relying less on equity to finance their investments. It is thus beneficial for companies to seek senior and subordinated loans to exploit the competitive advantages of both banks, lower funding costs for bank 1 and lower auditing costs for bank 2.

Summary It can be optimal for banks to borrow from two banks with one bank providing a senior loan and the other bank a subordinated loan. Such a seniority structure of their loans allows companies to extend their borrowing and increase profits as long as the loan rates remain below the cost of equity. The bank providing the senior loan is not concerned about the existence of a subordinated loan as it would be repaid first and thus provide the same loan amount at the same conditions as if being the only lender, hence the subordinated loan would expand the lending a company can obtain. The same result cannot be achieved when being provided by two loans of equal seniority; the frequency of auditing for both banks would be high and both banks would have to recover these costs from a smaller loan amount for each bank, increasing the costs to banks and subsequently loan costs. With subordinated loans, the auditing costs of the bank providing the senior loan remains unchanged, and only the bank providing the subordinated loan would face higher auditing costs due to more frequent auditing, but as their costs are lower, the net benefits to companies are positive. Thus companies use the competitive advantage of bank 1 having lower funding costs of the loan they provide and of bank 2 having lower auditing costs.

Reading Gangopadhyay & Mukhopadhyay (2002)

6.5 Loan renegotiation

Companies may seek to renegotiate the repayment of a loan if they are in distress; this would allow them to reduce the loan repayment and thereby avoid bankruptcy. However, if the bank cannot easily verify whether the company is actually in financial distress, companies might exploit this informational asymmetry and renegotiate the loan even if this is not required. We will seek here to explore the incentives for companies to ask for such a reduction in the amount to be repaid and for the bank to accept such a reduction.

Let us assume companies use the proceeds of a loan L to make an investment with an uncertain return R. This return is distributed uniformly on the interval $\left[\underline{R}; \overline{R}\right]$. In order to secure the loan, the company has pledged collateral to the value of C, which the bank can seize if it fails to repay the loan fully. This collateral can be sold by the bank only for a fraction $\lambda < 1$ of its value, for example due to the specialist nature of the collateral and the limited number of buyers. It is thus that the collateral is worth more to the company than it is worth to the bank.

We will consider the incentives for the company to ask for a reduced repayment and then consider the bank's optimal response.

Companies renegotiating We assume that the company knows the return on their investment, R, at the time they seek to renegotiate the loan repayments. If the company does not renegotiate the loan repayment its profits will be given by

$$\Pi_C = \begin{cases} (1+R)\,L - (1+r_L)\,L & \text{if } (1+R)\,L \geq (1+r_L)\,L \\ -C & \text{if } (1+R)\,L < (1+r_L)\,L \end{cases}. \tag{6.63}$$

If the return on investment is such that the company has sufficient funds to repay the loan including interest r_L, it will retain the difference between the return it generated and the repayments for the loan. If the return on the investment is not sufficient to repay the loan the company will file for bankruptcy and lose its collateral. We neglect the partial repayment of the loan, assuming that any administration costs will allow for no additional payment to either the bank or the company.

If, on the other hand, the company renegotiates the repayment of the loan such that the repayment is reduced to L^*, which would include any interest, then the bank profits are given by

$$\hat{\Pi}_C = \begin{cases} -C & \text{if } (1+R)\,L < L^* \\ (1+R)\,L - L^* & \text{if } L^* \leq (1+R)\,L < (1+r_L)\,L \\ (1+R)\,L - L^* - M & \text{if } L^* < (1+r_L)\,L \leq (1+R)\,L \end{cases}. \tag{6.64}$$

The first case reflects the situation in which the renegotiated repayment is still exceeding the return generated by the company and thus the company would fail, losing its collateral. In the second case the company would not be able to repay the original loan, but is able to pay the renegotiated loan repayment. In the final case, the bank was able to repay the original loan, but decided to renegotiate their repayments despite this fact. In this case we assume that the company faces additional costs M, such as a loss of reputation and potentially higher loan rates in future borrowing. As the company in the previous case was actually not able to make the loan repayments, we assume that no additional costs are incurred as the renegotiation was initiated to avoid bankruptcy, while in the final case this was done to reduce the loan repayments.

A company will seek to renegotiate the loan repayment if it is more profitable to do so, $\hat{\Pi}_C \leq \Pi_C$. Comparing the expressions in equations (6.63) and (6.64), we immediately see that for $(1+R)\,L \geq (1+r_L)\,L$ this will always be fulfilled. If $(1+R)\,L < L^*$, the company fails in both cases and loses its collateral, making it whether to renegotiate the loan repayment or not. On the other hand, if $(1+R)\,L \geq L^*$ the company will benefit from the renegotiation as it can avoid bankruptcy, increasing its profits from $-C$ to $(1+R)\,L - L^* > 0$. The final case, where the company renegotiates the loan repayment even though the original loan can be repaid, requires $(1+R)\,L - L^* - M \geq (1+R)\,L - (1+r_L)\,L$ for such a renegotiation to be initiated by the bank. This condition can be solved for

$$L^* \leq \hat{L}^* = (1+r_L)\,L - M. \tag{6.65}$$

Thus, the renegotiated loan repayment must be sufficiently small such that it compensates the company for the additional costs M.

It i, however, not only that the company needs to initiate the renegotiations, the bank also has to agree to reducing the repayment of the loan to L^*. We will analyse the considerations of the bank next.

Banks accepting renegotiations We assume that the bank does not know the return generated by the investment of the company, it has therefore to rely on the distribution of these returns to evaluate their profits. If the return generated is below the renegotiated loan repayment, then the company will fail and bank obtains the collateral which it can sell at a fraction λ of its value. If the company does not fail, the bank will obtain the renegotiated loan repayments. The expected profits of the bank are thus given by

$$\begin{aligned} \hat{\Pi}_B &= \text{Prob}\left(\left(1+R\right)L < L^*\right)\lambda C \\ &\quad + \left(1 - \text{Prob}\left(\left(1+R\right)L < L^*\right)\right)L^* - (1+r_D)\,D \\ &= \frac{L^*}{\left(\overline{R} - \underline{R}\right)L}\left(\lambda C - L^*\right) + L^* - (1+r_D)\,D, \end{aligned} \tag{6.66}$$

where we note that the bank has financed the loan with deposits D on which interest r_D is payable. The second line is obtained by collecting terms and using that the returns being distributed uniformly on $\left[\underline{R}; \overline{R}\right]$.

The bank will select the renegotiated loan repayment that is optimal for their profits. Thus the first order condition becomes

$$\frac{\partial \hat{\Pi}_B}{\partial L^*} = \frac{\lambda C - L^*}{\left(\overline{R} - \underline{R}\right)L} - \frac{L^*}{\left(\overline{R} - \underline{R}\right)L} + 1 = 0, \tag{6.67}$$

such that the optimal loan repayment is given by

$$L^* = \frac{1}{2}\left(\lambda C + \left(\overline{R} - \underline{R}\right) L\right). \tag{6.68}$$

From this we obtain the profits of the bank to be

$$\hat{\Pi}_B = \frac{\left(\lambda C + \left(\overline{R} - \underline{R}\right) L\right)^2}{4\left(\overline{R} - \underline{R}\right) L}. \tag{6.69}$$

If banks were not to renegotiate the repayments of the loan, the bank's profits would be given in analogy to equation (6.66) as

$$\begin{aligned}\hat{\Pi}_B &= \text{Prob}\left(\left(1+R\right)L < \left(1+r_L\right)L\right)\lambda C \\ &\quad + \left(1 - \text{Prob}\left(\left(1+R\right)L < \left(1+r_L\right)L\right)\right)\left(1+r_L\right)L \\ &\quad - \left(1+r_D\right)D \\ &= \frac{\left(1+r_L\right)L}{\left(\overline{R} - \underline{R}\right)L}\left(\lambda C - \left(1+r_L\right)L\right) + \left(1+r_L\right)L \\ &\quad - \left(1+r_D\right)D.\end{aligned} \tag{6.70}$$

For bank willing to renegotiate, their profits from doing so has to increase, thus we require $\hat{\Pi}_B \geq \Pi_B$. Let us assume now that the parameters are such that this condition is fulfilled and hence banks renegotiate the repayments of loans by setting them optimally. The renegotiated loan repayment must be accepted by the company, hence inserting equation (6.68) into equation (6.65) will require that

$$\lambda C \leq 2\left(1+r_L\right)L - 2M - \left(\overline{R} - \underline{R}\right)L. \tag{6.71}$$

We thus see that loan renegotiations are successfully completed if the value of collateral to the bank is low, either because the collateral itself has a low value or the value to the bank itself is low. Higher uncertainty about the investment returns of the company, $\overline{R} - \underline{R}$, will make this condition more strict. In this case the uncertainty of the investment outcome makes the bank only offer a higher loan repayment as equation (6.68) shows; the reason is that the wider spread of outcomes is increasing the chances of the company being able to repay the loan in full. Not surprisingly, higher costs M from such a renegotiation makes this action less desirable and the bank will offer a higher repayment, while an initially high repayment, $\left(1+r_L\right)L$, especially a high loan rate r_L, makes renegotiation more attractive to companies.

If course, banks might offer conditions for the repayment of loans that, while not optimal to them, are nevertheless acceptable to the company by fulfilling the constraint in equation (6.65). Thus inserting the highest possible repayment fulfilling the company constraint in equation (6.65) with equality we get the profits of the banks with renegotiation to be higher than without renegotiation if

$$\lambda C \leq 2\left(1+r_L\right)L + M - \left(\overline{R} - \underline{R}\right)L, \tag{6.72}$$

which is less restrictive as the condition in equation (6.71) when using the repayment rate optimal for the bank. The basic relationship remain valid, however.

Summary We have thus seen that as long as the value of collateral to banks is not too high, banks are willing to renegotiate loan repayments and companies benefit from a reduction in their liabilities. This can be profitable to banks as such lower repayments avoid the failure of the company and them only obtaining the collateral rather than the now lower loan repayment. We would thus expect to see loan renegotiations in particular in cases where collateral is of low value to the bank and where there is a not too large uncertainty about the investment outcomes for the company. In practice the renegotiation may not directly include the repayment of the loan, but the loan rate which is an important element of the repayment, especially in long-term loans,

Reading Flynn, Ghent, & Tchistyi (2024)

6.6 Delegated monitoring of loans

Banks will regularly interact with companies that have borrowed from them; this might take the form of regular meetings, the submission of quarterly accounts, but also the monitoring of payments the company makes to ensure the loan provided is used as intended. As part of the monitoring process the bank may also provide advice to companies, particularly smaller companies owned by less experience managers, on concerns about the financial health of the company, which might help them to identify measures to take for improvements. Such monitoring by banks is intended to reduce the risk of the loan not being repaid, but will be costly to banks. Of course, banks do not themselves monitor companies, but their managers will do so and hence we will here look at the impact of cuch delegated monitoring.

Companies can obtain a loan L and make an investment that, if conducted diligently, yields a return of $(1+R)$ if successful , which is the case with probability π; if the investment is not successful it yield a return of zero. However, a company that does not conduct the investment diligently will only generates private benefits λL, but no funds will be available to repay the loan. We here assume that $\pi\left(1+R\right) > 1 > \lambda$ hence a diligently conducted investment would be socially desirable, while the not diligently conducted investment is not socially desirable. The private benefits λL to the company may include the use of funds to increase employee and management pay, improve the facilities for staff, or the gain of prestige for the manager through publicity from the investment, even if it was not successful.

Let us define $\xi = \pi\left(1+R\right) - \lambda$ as the quality of the company and we assume that $\lambda > \pi\left(\left(1+R\right) - \left(1+r_L\right)\right)$, where r_L denotes the loan rate, such that companies would not act diligently if given the opportunity as the private benefits they can obtain exceed the profits the investment could generate.

We consider a case where banks monitor companies directly, avoiding any conflict of interest between the bank and its employees, or the bank delegates this task to its employees. If a company is monitored by a single bank or a single employee, they can enforce diligent behaviour at the company with probability p_i. The company may also take up two loans from different banks, each to the amount of $\frac{1}{2}L$, and in this case may be monitored by both banks. In order to enforce diligent behaviour at the company, it is sufficient that one these banks is successfully monitoring them. As each bank fails to monitor the bank with probability $1-p_i$, both fail with probability $(1-p_i)\left(1-p_j\right)$ and hence companies are behaving diligently with probability $p = 1-(1-p_i)\left(1-p_j\right)$. The costs of this monitoring, when having given a loan of size L, are given by $C_i = \frac{1}{2}cp_i^2L$.

We will now consider the socially optimal monitoring for the case that the company borrows from a single bank as well as the case that the company borrows from two banks. We will compare this result with the level of monitoring banks monitor companies directly and then if monitoring is delegated to employees of the bank.

Socially optimal monitoring Let us initially consider a company who has obtained a loan L from a single bank. The loan generates a surplus if the investment is conducted diligently, thus if the bank successfully monitors the company, p_i, and the investment itself is successful, π; if the bank's monitoring is not successful, $1-p_i$, the company will generate their private benefits. The loan is fully financed by deposits at interest rate r_D and the bank incurs monitoring costs. Hence the welfare from providing this loan is given by

$$\Pi_W^1 = p_i\pi(1+R)L + (1-p_i)\lambda L - (1+r_D)L - \frac{1}{2}cp_i^2L. \tag{6.73}$$

The optimal monitoring level can be obtained from the first order condition maximizing the social welfare,

$$\frac{\partial \Pi_W^1}{\partial p_i} = (\pi(1+R) - \lambda - cp_i)L = 0, \tag{6.74}$$

which solves for or using the definition of ξ that $\xi - cp_i = 0$, hence

$$\hat{p}_i^* = \frac{\xi}{c}, \tag{6.75}$$

using that $\xi = \pi(1+R) - \lambda$.

If instead of using a single bank, the company seeks out two loans of $\frac{1}{2}L$ from two different banks. The company will be acting diligently unless both banks fail to monitor the company successfully, which was shown to be with probability $p = 1-(1-p_i)\left(1-p_j\right)$. Each bank monitors for their part of the loan, thus the social welfare becomes

$$\Pi_W^2 = p\pi(1+R)L + (1-p)\lambda L - (1+r_D)L - \frac{1}{2}cp_i^2\frac{L}{2} - \frac{1}{2}cp_j^2\frac{L}{2}. \tag{6.76}$$

The optimal monitoring level can be obtained from the first order condition maximizing the social welfare, which is

$$\frac{\partial \Pi_W^2}{\partial p_i} = \left(\left(1 - p_j\right) \left(\pi \left(1 + R\right) - \lambda\right) - \frac{1}{2} c p_i \right) L = 0, \tag{6.77}$$
$$\frac{\partial \Pi_W^2}{\partial p_j} = \left(\left(1 - p_i\right) \left(\pi \left(1 + R\right) - \lambda\right) - \frac{1}{2} c p_j \right) L = 0.$$

Considering only symmetric equilibria with $p_i = p_j$, this solves for

$$\hat{p}_i^{**} = \hat{p}_j^{**} = \frac{2\xi}{c + 2\xi}. \tag{6.78}$$

We see instantly that the monitoring effort of each bank individually is higher when a single bank is used, thus $\hat{p}_i^* > \hat{p}_i^{**}$ if $c < 2\xi$ and hence the costs of monitoring ar not too high. This socially optimal monitoring by banks can now be compared to the monitoring that is optimal for banks.

Direct monitoring If the loan is taken from a single bank, its profits are given by

$$\Pi_B = p_i \pi \left(1 + r_L\right) L - \left(1 + r_D\right) D - \frac{1}{2} c p_i^2 L, \tag{6.79}$$

where the investment is done diligently if the bank monitors successfully, p_i, and the investment itself is successful, π such that the loan can be repaid. The bank has to repay the deposits and faces their monitoring costs. The optimal monitoring is given by maximizing this expression and we obtain the first order condition as

$$\frac{\partial \Pi_B}{\partial p_i} = \left(\pi \left(1 + r_L\right) - c p_i\right) L = 0, \tag{6.80}$$

which solves for

$$p_i^* = \frac{\pi \left(1 + r_L\right)}{c}. \tag{6.81}$$

As we had assumed that $\lambda > \pi \left(\left(1 + R\right) - \left(1 + r_L\right)\right)$ and hence $\pi \left(1 + r_L\right) > \xi$ we see that compared to the social optimum, banks monitor more. The reason is that banks do not obtain the private benefits λL the company generates if its investments are not diligently conducted. It is therefore that the bank will monitor more to increase the chances of a diligent investment.

If banks obtain the same loan amount from two banks, then we have for$i$ that

$$\Pi_B^i = p \pi \left(1 + r_L\right) \frac{L}{2} - \left(1 + r_D\right) \frac{L}{2} - \frac{1}{2} c p_i^2 \frac{L}{2}, \tag{6.82}$$

where $p = 1 - \left(1 - p_i\right) \left(1 - p_j\right)$. Hence the optimal monitoring by banks directly is given from

$$\frac{\partial \Pi_B^i}{\partial p_i} = \left(\left(1 - p_j\right)\pi\left(1 + r_L\right) - \frac{1}{2}cp_i\right)L = 0, \tag{6.83}$$

hence

$$\begin{aligned} p_i &= 2\frac{\left(1 - p_j\right)\pi\left(1 + r_L\right)}{c}, \\ p_j &= 2\frac{\left(1 - p_i\right)\pi\left(1 + r_L\right)}{c}, \end{aligned} \tag{6.84}$$

where the second equation arises from the same derivation for bank j. Solving these two equations for the symmetric equilibrium we get

$$p_i^{**} = p_j^{**} = \frac{2\pi\left(1 + r_L\right)}{c + 2\pi\left(1 + r_L\right)}, \tag{6.85}$$

which is less monitoring than with a single bank as comparison with (6.81) easily shows, provided the costs of monitoring are sufficiently low such that $c < 2\pi\left(1 + r_L\right)$. For the same reasons outlined above, banks are monitoring more than is socially optimal.

We can now continue the analysis by considering the most realistic case where the bank does not directly monitor companies, but delegates this task to their employees.

Delegated monitoring from a single bank Let us now assume that the monitoring of companies is delegated to employees of the bank. As it is more beneficial for companies to not act diligently, they could offer bank employees a payment $\hat{\lambda}L$ if they do not enforce them acting diligently. In this case the profits to the company and managers, respectively, are given by

$$\begin{aligned} \Pi_C &= \lambda L - \hat{\lambda}L, \\ \Pi_M &= \hat{\lambda}L - \frac{1}{2}cp_i^2 L. \end{aligned} \tag{6.86}$$

If the loan is repaid, and thus the monitoring of the manager was successful, we assume the manager obtains a bonus from the bank which is calculated as a fraction α of the amount repaid. Hence if the manager forces the company to act diligently, we have the respective profits given by

$$\begin{aligned} \hat{\Pi}_C &= \pi\left(\left(1 + R\right) - \left(1 + r_L\right)\right)L, \\ \hat{\Pi}_M &= \alpha\pi\left(1 + r_L\right)L - \frac{1}{2}cp_i^2 L. \end{aligned} \tag{6.87}$$

If the aggregate surplus from the company and the manager in case of investments being conducted diligently is larger than that of a non-diligent conduct, the rewards of the manager can be set such that he will prefer diligence. Hence we need $\hat{\Pi}_C + \hat{\Pi}_M \geq \Pi_C + \Pi_M$, which when inserting from equations (6.86) and (6.87), becomes

$$\xi \geq \left(1 - \alpha\right)\pi\left(1 + r_L\right). \tag{6.88}$$

The payment to the manager if the company acts diligently will only be due if the monitoring is successful, hence his expected profits are given by

$$\Pi_M = p_i \alpha \pi (1 + r_L) L - \frac{1}{2} c p_i^2 L \geq 0, \tag{6.89}$$

which need to be positive for him to take up employment with the bank.

As the monitoring costs are covered by the managers but not incurred by the bank, managers receive a bonus from repaid loans to cover their monitoring costs, bank profits are given by

$$\Pi_B = (1 - \alpha) p_i \pi (1 + r_L) L - (1 + r_D) L, \tag{6.90}$$

noting that banks only retain a fraction $1 - \alpha$ of the repaid loan, the remainder paid to the manager as a bonus, provided the loan is repaid.

If we assume that $\alpha = 1$ and $p_i = 0$ we have $\Pi_B = -(1 + r_D) L < 0$ and hence banks would not provide loans, ruling out this possibility; for $\alpha = 0$, we see that managers would not take up employment as $\Pi_M < 0$.

Maximizing bank profits will require the optimal solution to comply with a number of constraints. Firstly we require that the probability of successful monitoring is less than 1, $p_i \leq 1$. In addition, equations (6.88) and (6.89) need to be met. With Lagrange multipliers η_i, we have our objective function given as

$$\mathcal{L} = \Pi_B + \eta_1 \Pi_M + \eta_2 (\xi - (1 - \alpha) \pi (1 + r_L)) + \eta_3 (1 - p_i). \tag{6.91}$$

This gives us the first order conditions for maximizing the bank profits by choosing the optimal monitoring and the optimal bonus payments to managers as

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial p_i} &= (1 - \alpha) \pi (1 + r_L) L + \eta_1 (\alpha \pi (1 + r_L) - c p_i) L - \eta_3 \\ &= 0, \\ \frac{\partial \mathcal{L}}{\partial \alpha} &= -p_i \pi (1 + r_L) L + \eta_1 p_i \pi (1 + r_l) L + \eta_2 \pi (1 + r_L) L \\ &= (p_i (\eta_1 - 1) - \eta_2) \pi (1 + r_L) L \\ &= 0. \end{aligned} \tag{6.92}$$

Let us now assume that $p_i = 1$ and hence the constraint is binding such that $\eta_3 > 0$. Should the second constraint not be binding such that $\eta_2 = 0$, the last line implies $\eta_1 = 1$ and $\Pi_M = 0$ as this constraint must be binding. In this case the first line in equation (6.92) becomes $(\pi (1 + r_L) - c) L = \eta_3$, when noting that $p_i = 1$. If monitoring costs are sufficiently high, $c > \pi (1 + r_L)$, then the left-hand side is negative, contradicting our assumption that $\eta_3 > 0$. We thus need $\eta_2 > 0$ and hence $\xi = (1 - \alpha) \pi (1 + r_L)$ as the constraint has to be binding. This is feasible only if banks are profitable, $\Pi_B \geq 0$. Inserting into equation (6.90) for $p_i = 1$ and using that $\xi = (1 - \alpha) \pi (1 + r_L)$, we get that for

$$\xi \geq \underline{\xi} = 1 + r_D \tag{6.93}$$

we have $p_i = 1$. The constraint that $\Pi_M \geq 0$, solves similarly for

$$\xi \leq \hat{\xi} = \pi (1 + r_L) - \frac{1}{2}c. \tag{6.94}$$

Hence if $\underline{\xi} \leq \xi \leq \hat{\xi}$ we have $p_i = 1$.

Consider now the case that $\eta_3 = 0$ and hence $p_i < 1$. We then need $\eta_1 > 0$ as for $\eta_1 = 0$ from the first line in equation (6.92) we obtain $\alpha = 1$, and hence $\Pi_B < 0$. Therefore $\Pi_M = 0$ is a binding constraint.

As a next step let us assume that $\eta_2 > 0$ hence $\xi = (1 - \alpha) \pi (1 + r_L)$, such that

$$\alpha = 1 - \frac{\xi}{\pi (1 + r_L)}. \tag{6.95}$$

Inserting this relationship into the constraint that $\Pi_M = 0$, we easily get

$$p_i = 2\frac{\pi (1 + r_L) - \xi}{c}. \tag{6.96}$$

Inserting equations (6.95) and (6.96) into the first line of the first order conditions in equation (6.92), we get that

$$\eta_1 = \frac{\xi}{\pi (1 + r_L) - \xi}. \tag{6.97}$$

Inserting this result in turn into the final line of equation (6.92), we obtain

$$\eta_2 = 2\frac{\pi (1 + r_L) - 2\xi}{c} \tag{6.98}$$

To meet our assumption we need that $\eta_2 > 0$, hence

$$\xi < \overline{\xi} = \frac{1}{2}\pi (1 + r_L). \tag{6.99}$$

If we instead assume that $\eta_2 = 0$, then with $\eta_3 = 0$ the last line of equation (6.92) becomes $p_i (\eta_1 - 1) = 0$, implying either $p_i = 0$, which means $\Pi_B < 0$ and is therefore not feasible, or $\eta_1 = 1$ and hence from the first line of equation (6.92) that

$$p_i = \frac{\pi (1 + r_L)}{c}. \tag{6.100}$$

As $\eta_1 > 0$, we have $\Pi_M = 0$ as a binding constraint. Inserting from equation (6.100) into equation (6.89) we easily solve for the bonus payment to be given by

$$\alpha = \frac{1}{2}. \tag{6.101}$$

Hence as $\eta_2 = 0$, we have

$$\xi \le \overline{\xi} = \frac{1}{2}\pi\left(1 + r_L\right). \tag{6.102}$$

We can now summarize the optimal monitoring by managers in the case that the company chooses a single bank given as

$$p_i^{***} = \begin{cases} 1 & \text{if} \quad \underline{\xi} \le \xi \le \hat{\xi} \\ 2\frac{\pi(1+r_L)-\xi}{c} & \text{if} \quad \hat{\xi} < \xi \le \overline{\xi} \\ \frac{\pi(1+r_L)}{c} & \text{if}\ \overline{\xi} < \xi \le \pi\left(1 + r_L\right) \end{cases}. \tag{6.103}$$

Figure 6.1 also visualises this result for the monitoring of the company if delegated to managers. We immediately see by comparing equations (6.81) and (6.103) that managers monitor more than banks would as we find that $p_i^{***} \ge p_i^*$. This result emerges as the manager participates not only in the case that companies are acting non-diligently through the payments they receive, but through the bonus paid by the bank they also benefit from the company acting diligently. The bonus payments are such that managers have incentives to ensure the loan is repaid and will thus enforce diligence more often. This is in turn the result of banks providing managers this incentive as they themselves only benefit from repaid loans. As the bank profits are reduced by the bonus paid, the banks want the diligence to be enforced even more rigorously than if they had been monitoring directly.

We also observe that the monitoring reduces the higher the quality ξ of the company is, while for direct monitoring by banks this did not affect the monitoring. With delegated monitoring, managers benefit from both, the repayment of loans in case companies invest diligently and the payments they receive from non-diligence; a higher quality company has a larger surplus to share between themselves and the manager, giving them a higher payment if companies are not diligent, reducing their incentives to monitor companies.

We can now proceed to consider the case that two banks provide the loan and monitor the company.

Delegated monitoring from two banks We commence by assuming that only one of the two managers detects non-diligence by the bank, then profits for non-diligence and diligence for this manager and the company are given by

$$\begin{aligned} \Pi_C &= \lambda L - \hat{\lambda}_i L, \\ \Pi_M^i &= \hat{\lambda}_i - \frac{1}{2}cp_i^2\frac{L}{2}, \\ \hat{\Pi}_C &= \pi\left(\left(1 + R\right) - \left(1 + r_L\right)\right)L, \\ \hat{\Pi}_M^i &= \alpha_i\pi\left(1 + r_L\right)\frac{L}{2} - \frac{1}{2}cp_i^2\frac{L}{2}. \end{aligned} \tag{6.104}$$

As before, if the joint surplus from enforcing diligence is higher than from not enforcing diligence, $\hat{\Pi}_C + \hat{\Pi}_M \ge \Pi_C + \Pi_M$, the manager will enforce diligence by the company. Thus we need

$$\xi \geq \left(1 - \frac{1}{2}\alpha_i\right)\pi\left(1 + r_L\right). \tag{6.105}$$

If on the other hand both managers detect non-diligence, we have the respective profits given by

$$\begin{aligned}
\Pi_C &= \lambda L - \hat{\lambda}_i \frac{L}{2} - \hat{\lambda}_j \frac{L}{2}, \qquad (6.106)\\
\Pi_M^i &= \hat{\lambda}_i \frac{L}{2} - \frac{1}{2}cp_i^2 \frac{L}{2},\\
\Pi_M^j &= \hat{\lambda}_j \frac{L}{2} - \frac{1}{2}cp_j^2 \frac{L}{2},\\
\hat{\Pi}_C &= \pi\left(\left(1 + R\right) - \left(1 + r_L\right)\right) L,\\
\hat{\Pi}_M^i &= \alpha_i \pi\left(1 + r_l\right) \frac{L}{2} - \frac{1}{2}cp_i^2 \frac{L}{2},\\
\hat{\Pi}_M^j &= \alpha_j \pi\left(1 + r_l\right) \frac{L}{2} - \frac{1}{2}cp_j^2 \frac{L}{2},
\end{aligned}$$

where we note that the company needs to payment $\hat{\lambda}_i$ to both managers as each could enforce diligence individually. Diligence will be enforced if the joint surplus of doing so is higher, $\hat{\Pi}_C + \hat{\Pi}_M^i + \hat{\Pi}_M^j \geq \Pi_C + \Pi_M^i + \Pi_M^j$, which requires

$$\xi \geq \left(1 - \frac{1}{2}\left(\alpha_i + \alpha_j\right)\right)\pi\left(1 + r_L\right). \tag{6.107}$$

As this condition is less restrictive than the restriction in equation (6.105) which only required one manager to detect the non-diligence of the company, we assume the more strict constraint in equation (6.105) to be applicable. The manager will only take up employment with the bank if he generates a surplus, thus

$$\Pi_M^i = \alpha_i p\pi\left(1 + r_L\right) \frac{L}{2} - \frac{1}{2}cp_i^2 \frac{L}{2} \geq 0, \tag{6.108}$$

where p denotes the detection of non-diligence by either manager as defined above. Bank profits are then given by

$$\Pi_B^i = \left(1 - \alpha_i\right) p\pi\left(1 + r_L\right) \frac{L}{2} - \left(1 + r_D\right) \frac{L}{2}. \tag{6.109}$$

The objective function for each bank then becomes

$$\begin{aligned}
\mathcal{L}_i = {} & \Pi_B^i + \eta_1^i \Pi_M^i \qquad (6.110)\\
& + \eta_2^i \left(\xi - \left(1 - \frac{1}{2}\left(1\alpha_i + \alpha_j\right)\right)\pi\left(1 + r_L\right)\right)\\
& + \eta_3^i \left(1 - p_i\right),
\end{aligned}$$

such that the first order conditions are

$$\frac{\partial \mathcal{L}_i}{\partial p_i} = (1-\alpha_i)\left(1-p_j\right)\pi(1+r_L)\frac{L}{2} \tag{6.111}$$
$$+\eta_1^i\left(\alpha\left(1-p_j\right)\pi(1+r_L)\frac{L}{2} - cp_i\frac{L}{2}\right) - \eta_3^i = 0,$$
$$\frac{\partial \mathcal{L}_i}{\partial \alpha_i} = -p\pi(1+r_L)\frac{L}{2} + \eta_1^i\pi(1+r_L)\frac{L}{2} + \frac{1}{2}\eta_2^i = 0.$$

As before, this optimization has to be conducted with all the constraints discussed, being fulfilled.

If we now assume $\eta_3^i > 0$ and hence $p_i = 1$, then the first line of equation (6.111) becomes

$$-\eta_1^i cp_i\frac{L}{2} = \eta_3^i. \tag{6.112}$$

This implies η_1^1 and η_3^i have different signs, which is impossible as Lagrange multiplier are non-negative. Hence we require that $\eta_3^i = 0$ and similarly $\eta_3^j = 0$, such that the monitoring is non-binding with $p_i < 1$ and $p_j < 1$.

If we now assume that $\eta_1^i = 0$, then the first line in equation (6.111) becomes $\left(1-p_j\right)(1-\alpha_i)\pi(1+r_L) = 0$ and as $p_j < 1$ due to $\eta_3^j = 0$, we need $\alpha_i = 1$, which implies $\Pi_B^i = -(1+r_D)\frac{L}{2} < 0$ which is not feasible and thus we need $\eta_1^i > 0$ to ensure banks participate. As this constraint is then binding we require $\Pi_M = 0$, giving us

$$\alpha_i\pi(1+r_L) = \frac{cp_i^2}{2p}. \tag{6.113}$$

Using equation (6.105), we obtain

$$2(\pi(1+r_L)-\xi) \leq \frac{cp_i^2}{2p}. \tag{6.114}$$

In the case this constraint is not binding, $\eta_2^i = 0$, no constraints are binding, and maximizing equation (6.110) becomes equivalent to the bank profits in equation (6.82), giving as the same result as when banks are monitoring directly, which was

$$p_i = p_j = \frac{2\pi(1+r_L)}{c+2\pi(1+r_L)}. \tag{6.115}$$

Inserting this result into equation (6.114), this requires

$$\xi \geq \overline{\xi} = \frac{3c+4\pi(1+r_L)}{4(c+\pi(1+r_L))}\pi(1+r_L). \tag{6.116}$$

If, in contrast, equation (6.114) is binding, then as this applies to both managers, we have

$$2(\pi(1+r_L)-\xi) = \frac{cp_i^2}{2p} = \frac{cp_j^2}{2p}$$

and hence $p_i = p_j$, such that

$$2\left(\pi\left(1+r_L\right)-\xi\right)=\frac{cp_i^2}{2\left(1-\left(1-p_i\right)^2\right)},$$

which solves for

$$p_i=p_j=8\frac{\pi\left(1+r_L\right)-\xi}{c+4\left(\pi\left(1+r_L\right)-\xi\right)}. \tag{6.117}$$

The second solution $p_i=p_j=0$ can be ruled out due to this resulting in $\Pi_B^i<0$.

As we need $p_i\leq 1$, we get from equation (6.117) that

$$\xi\geq\underline{\xi}=\pi\left(1+r_L\right)-\frac{c}{4}. \tag{6.118}$$

In summary, we have the optimal monitoring where companies choose to divide the loan between two banks, given as

$$p_i^{****}=p_j^{****}=\begin{cases}8\frac{\pi(1+r_L)-\xi}{c+4(\pi(1+r_L)-\xi)} & \text{if} \quad \underline{\xi}\leq\xi\leq\overline{\xi} \\ \frac{\pi(1+r_L)}{c+\pi(1+r_L)} & \text{if } \overline{\xi}<\xi\leq\pi\left(1+r_L\right)\end{cases}. \tag{6.119}$$

We have again visualised this result in figure 6.1. As before in the case of a single bank, we notice that with delegated monitoring the level of monitoring by each manager is either higher or equal to the level of monitoring that banks would seek to implement directly, and this was in turn higher than the socially optimal monitoring. An explanation for this result is identical to that given in the case of a single bank. Similarly, with two banks, a higher quality bank will also be monitored less due to larger surplus the company can share with the manager; as now the bank has to meet the requirements of two managers, of which only one needs to enforce diligence, the payments available in this case will be increasing more than with a single bank and hence the effect of company quality on the optimal level of monitoring will be stronger. Eventually the monitoring with two banks will be lower than with a single bank.

Borrower choice The profits for the company if choosing to borrow from one bank is given by

$$\begin{aligned}\Pi_C &= p_i\pi\left(\left(1+R\right)-\left(1+r_L\right)\right)L+\left(1-p_i\right)\lambda L \\ &= \left(\pi\left(1+R\right)-\xi\right)L-p_i\left(\pi\left(1+r_L\right)-\xi\right)L\end{aligned} \tag{6.120}$$

and for two banks it is

$$\begin{aligned}\hat{\Pi}_C &= p\pi\left(\left(1+R\right)-\left(1+r_L\right)\right)+\left(1-p\right)\lambda L \\ &= \left(\pi\left(1+R\right)-\xi\right)L-p\left(\pi\left(1+r_L\right)-\xi\right)L.\end{aligned} \tag{6.121}$$

The company prefers a single bank if its profits are higher, $\Pi_C\geq\hat{\Pi}_C$, or $p\geq p_i$. Inserting for p and p_i, this condition becomes

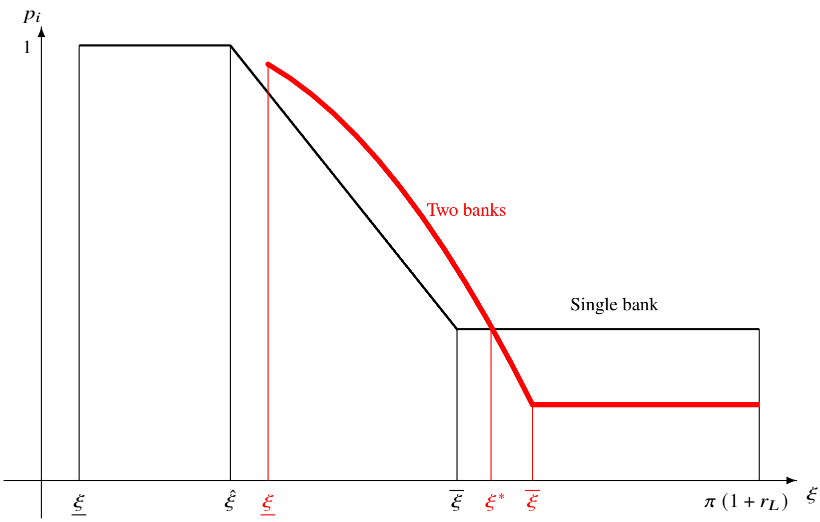


Fig. 6.1: Monitoring with a single and two banks

$$1 - \left(1 - 8\frac{\pi\,(1 + r_L) - \xi}{c + 4\,(\pi\,(1 + r_L) - \xi)}\right)^2 \geq \frac{\pi\,(1 + r_L)}{c} \qquad (6.122)$$

We see that this is only fulfilled for ξ being sufficiently small, but not too small, thus $\xi^{**}\xi < \xi^{*}$ as for ξ approaching $\pi\,(1 + r_L)$, the left hand side approaches zero, and the condition thus not being fulfilled; as ξ becomes small, the left-hand expression will increase again. However if we assume that $c > \pi\,(1 + r_L)$, then $\xi^{**} < \underline{\xi}$ and we neglect this possibility.

If the quality of companies is high, thus they exhibit a high value for $\xi = \pi\,(1 + R) - \lambda$, they prefer being monitored by two banks and thus they divide their loans up accordingly. Companies are monitored less, and hence more likely to obtain the higher benefits λL, as the high

Summary We have now established that when banks prefer to monitor more than is socially optimal and if this monitoring is delegated to managers, monitoring increases even more. It is therefore that companies are monitored overly strictly for their diligence in investing the proceeds of the loan. The higher the quality of the company is, as measured by the difference between the return in a diligent investment compared to the case where companies are not diligent, the less they are monitored. This effect is stronger if companies are monitored by multiple banks and if their quality is sufficiently high, being monitored by multiple banks is actually optimal for companies and they will divide their loans between banks to achieve this situation.

We have thus seen that companies that are deemed to generate higher surpluses are monitored less and that for very well-performing companies it will even be beneficial to obtain several smaller loans from different banks in order to reduce the level of monitoring and thus increasing their profits as they are more free to make their investments in their own interest rather than having their focus primarily on the ability to repay the loan, which banks are interested in.

Reading Dam & Chowdhury (2021)

Conclusions

We have seen that the commonly used loan contract in which the bank is repaid the initial funds provided by the bank plus interest if the outcome allows the company to do so, and repay as much as possible otherwise. This arrangement was shown to be optimal if the outcome of an investment could not be easily verified, while an equity-like loan contract would be optimal if any verification could be achieved at no cost. Such an arrangement allows to minimise auditing costs as the outcome of investments do not need to be verified if the company repays the full amount and this costly auditing is limited to situations where the loan is not repaid fully. The amount that is to be repaid, the initial loan plus interest payment, will have to be selected such that it covers and funding costs of banks and the auditing costs that may be incurred.

For decision-making, banks rely on their assessment of the risks companies pose in terms of their ability to repay the loan. In order to achieve this assessment, banks rely on information and more precise information allows them to price a loan more accurately, thus being able to extract more surplus from the company. But more precise information can also provide them with an informational advantage over their competitors, allowing them to quote loan rates that attract more loans to their bank, at the expense of their less well informed competitors. These benefits of information will induce banks to seek out more precise information.

Having previously established the optimal loan contract as being the traditional arrangement of repaying the initial amount lent plus interest, the time to maturity of such a loan might well exceed that of the investment the company seeks to fund with its proceeds. It turned out to be more cost-effective for low-risk borrowers to use a single long-term loan for a sequence of investments rather than a larger number of short-term loans matching the maturity of these investments. The origins of this result is that with long-term loans, the bank may obtain repayments on their loan even if some investments fail as we had assumed they can access past surplus, while with short-term loans as separate contracts this was not possible. For high-risk borrowers, banks will access such past surplus too frequently and long-term loans would be more expensive than short-term loans, despite having lower loan rates.

Companies can increase the amount of loans they can use to finance investments by arranging for a senior loan to be accompanied by a subordinated loan from a different bank. If the cost structure of banks are such that each of them has a competitive advantage in one area, funding costs and auditing costs, the company can seek to combine loans from both banks by exploiting these cost differences to take advantage of increased borrowing that increases their profits. When seeking loans from different banks, those companies with high returns may also reduce the level of monitoring banks impose on them.

Having shown that the commonly used loan contract is optimal and that it may be optimal for some companies to obtain long-term loans from multiple banks at different level of seniority, the following chapters will now explore how other aspects of the loan contract can be used to affect the willingness of banks to provide loans, but also how it may affect the behaviour of companies themselves and thereby the riskiness of the investment they finance with the loan they obtain.

Chapter 7

Strategic default

It is commonly assumed that borrowers are repaying their loans if they are able to do so. However, borrowers may well have an incentive to default on these obligations, even if they are able to meet them, which is referred to as a strategic default. The benefits of a strategic default are that the borrower can retain a larger proportion of the proceeds of their investment rather than repaying the loan, increasing their profits. Of course, banks would not agree to provide loans if the repayments are insufficient to generate them profits. Thus a mechanism needs to be provided that ensures strategic default does not happen or is at least sufficiently unlikely to ensure banks are willing to lend. One commonly used mechanism is that of auditing the borrower in case they default to identify whether they are unable or unwilling to repay the loan. We have seen in chapter 6.1.3 that such auditing leads to the standard debt contract, where repaying a fixed amount is optimal.

Banks could initiate audits of borrowers that do not repay their loans, but such audits are costly to banks and will bind resources that could instead be used to generate profits from additional lending. Hence the amount of resources a bank is willing to invest into this auditing process of failed loans will be limited. In chapter 7.1 we will see how the provision of limited auditing resources affects strategic defaults by borrowers.

Such audits are costly and it would be preferable if banks could provide incentives to avoid strategic defaults by companies even in the absence of audits. One way to incentivise borrowers to repay their loans is by excluding them from future loans and thus limit their future profitability. Chapter 7.2 looks at the conditions that need to be met such that borrowers do not default strategically. However, excluding borrowers from any future borrowing also reduces the profitability of the bank as they cannot generate profits from such future lending. Therefore chapter 7.3 will determine the optimal time period during which borrowers should be excluded from obtaining loans.

Strategic default does not commonly take the form of a company claiming that it has no resources available to repay a loan as their investments have been unsuccessful. Instead, companies may choose to use excessive dividend payments to reduce the capital a company has available, reward senior managers excessively to reduce the profitability of the company, or use transfer pricing between the company and its affiliated companies, often located abroad, to reduce the profits generated. Such practices is not only more difficult to detect during an audit, but it is also more difficult to prove that these measures have been implemented with the aim of avoiding the repayment of the loan. This would then be followed by legal problems to recover any funds transferred outside of the company.

7.1 Limited audit resources

It is common to assume that if a company fails to repay its loan, claiming this is due to their investments failing, it gets audited by the bank to assess the validity of their claim. Conducting such an audit is, however, costly to the bank and while it may recover these costs from companies if these are found to claim an inability to pay fraudulently, it first needs to obtain the resources to commence the audit process and cover their initial costs from these resources. The more audit resources are available, the more audits can be conducted. Putting aside such resources implies that banks have less funding available to provide loans, which will affect their profitability. Banks will therefore have to balance the number of audits, and hence the likelihood of detecting any fraudulent claims of not being able to repay loans against the loss of profits from reduced lending. This leads to a strategic interaction between companies deciding to default strategically and banks committing audit resources to detect such strategic defaults.

Company incentives Companies obtain a loan to make an investment that is successful and generates a value of V to the company with probability π. If the investment is not successful, it generates no value. A company which is not successful, cannot repay the loan and will therefore have to default on the loan. Hence we are only considering companies with a successful investment outcome as these are the companies that can default strategically. The investment of the company is financed by a loan L on which a loan rate of r_L is payable, hence the profits of a company not defaulting strategically are given by

$$\Pi_C = V - (1 + r_L) L. \tag{7.1}$$

If the investment is either not successful or the company defaults strategically, we assume that they are audited with probability p and that this audit will detect the strategic default with certainty. Each audit costs the bank C and this amount is charged to the company if a strategic default is detected, in addition to having to repay the loan in full. Hence the profits of a strategically defaulting company is given by

$$\hat{\Pi}_C = p\,(V - (1 + r_L) L - C) + (1 - p) V = V - p\,((1 + r_L) L + C)\,. \tag{7.2}$$

The first term denotes the case where the company is audited and the company repays the loan as well as compensates the bank for its auditing costs, while the second term denotes the case where the company is not audited and hence can keep the outcome of the investment without having to repay the loan.

The company will strategically default if it is more profitable to do so, $\hat{\Pi}_C \geq \Pi_C$, which easily solves for

$$p \leq p^* = \frac{(1 + r_L)\, L}{(1 + r_L)\, L + C}. \tag{7.3}$$

Thus, if the probability of being audited is sufficiently low, the company will default strategically. The bank can affect this probability of a company being audited and will determine audit resources optimally.

Bank incentives Banks are faced with N_D companies defaulting. These defaults can either be the result of unsuccessful investments that do not allow companies to repay their loans, or the result of strategic default. If a total of N loans are given, then a fraction $(1 - \pi)\, N$ would default due to unsuccessful investments. Of those companies whose investments are successful, a total of πN, we assume that a fraction κ defaults strategically. Hence the total number of defaulting companies is given by

$$N_D = (1 - \pi)\, N + \kappa\pi N = (1 - \pi\, (1 - \kappa))\, N. \tag{7.4}$$

Having set aside total audit resources of W and with each audit costing C, the bank can conduct $\frac{W}{C}$ audits. Having N_D defaults, a fraction

$$p = \frac{W}{CN_D} \tag{7.5}$$

of companies can be audited.

The bank now obtains their repayment $(1 + r_L)\, L$ of the loan from those with successful investment that are not strategically defaulting, $\pi\, (1 - \kappa)\, N$, and those strategically defaulting that are audited and hence identified as being able to repay the loan, $p\kappa\pi N$. Banks have to bear the audit costs for those companies that had unsuccessful investments and were being audited, $p\, (1 - \pi)\, N$. Hence with deposits D to finance the loans, on which interest r_D is paid, the bank profits are given by

$$\begin{aligned} \Pi_B &= (\pi\, (1 - \kappa)\, N + p\kappa\pi N)\, (1 + r_L)\, L - p\, (1 - \pi)\, NC \\ &\quad - (1 + r_D)\, D \\ &= \frac{\pi\, (\kappa L - (1 - \pi\, (1 - \kappa))\, (1 - \kappa)\, C)\, (1 + r_L) - (1 - \pi)\, C}{(1 - \pi\, (1 - \kappa))\, C} W \\ &\quad - ((1 + r_D) - \pi\, (1 - \kappa)\, (1 + r_L))\, D. \end{aligned} \tag{7.6}$$

The final expression we obtain by noting that the total amount of lending the amount of L to N companies, NL, will be constrained by the amount of deposits D the bank raises and the auditing resources W such that $NL = D - W$. Replacing the audit probability p with the expression in equation (7.5), the number of defaults N_D with the expression in equation (7.4) and the number of loans N with $N = \frac{D-W}{L}$, the final expression emerges.

Optimal auditing resources The optimal audit resources a bank sets aside are given by the bank maximizing its profits. The first order condition then becomes

$$\frac{\partial \Pi_B}{\partial W} = \frac{\pi \left(\kappa L - (1 - \pi (1 - \kappa)) (1 - \kappa) C\right) (1 + r_L) - (1 - \pi) C}{(1 - \pi (1 - \kappa)) C} \lesseqqgtr 0. \tag{7.7}$$

As we see, the optimal audit resources depend on the fraction of companies defaulting strategically, κ. We easily see that for small values of strategic default, the first order condition is negative, implying that the lowest possible audit resources are optimal, $W^* = 0$. In this case no audit can occur and we easily see from equation (7.5) that $p = 0$ and hence as equation (7.3) is fulfilled, all companies will strategically default and $\kappa = 1$. This would not be an equilibrium as we had assumed that strategic default κ was low, but the lack of audit resources committed in this case implies that strategic default is high. An equilibrium would require that for $W^* = 0$, we have $\kappa^* = 0$.

On the other hand, for high fractions of strategic default, the first order condition is positive, requiring banks to using the maximum amount of audit resources; this would be $W = CN_D = (1 - \pi (1 - \kappa)) CN$ as higher resources do not further increase the probability of companies being audited, which at that point reaches $p = 1$ as we see from equation (7.5). Inserting $N = \frac{D-W}{L}$, we obtain that the optimal audit resources are given by $W^{***} = \frac{(1-\pi(1-\kappa))\frac{C}{L}}{1+(1-\pi(1-\kappa))\frac{C}{L}}$. In this case $p = 1$ and hence equation (7.3) is never fulfilled, such that no company defaults strategically and $\kappa = 0$. Again, this would not be an equilibrium as we had assumed that strategic default κ was high, but the high audit resources committed in this case implies that strategic default is low. An equilibrium would require that for $W^{***} = \frac{(1-\pi(1-\kappa))\frac{C}{L}}{1+(1-\pi(1-\kappa))\frac{C}{L}}$, we have $\kappa^{***} = 1$.

In the case that $\frac{\partial \Pi_B}{\partial W} = 0$, which requires that

$$\kappa^{**} = -\frac{L + (1 - 2\pi) C}{2\pi C} + \sqrt{\left(\frac{L + (1 - 2\pi) C}{2\pi C}\right)^2 + \frac{1 + \pi (1 + r_L)}{\pi^2 (1 + r_L)} (1 - \pi)}, \tag{7.8}$$

banks are indifferent between all levels of audit resources. As only a fraction κ^{**} of companies will strategically default in this case, companies must be indifferent between strategically defaulting and not doing so. Hence, equation (7.3) must be fulfilled with equality. Inserting for p from equation (7.5) and using that $N = \frac{D-W}{L}$, we easily get that $W^{**} = \frac{p^*(1-\pi(1-\kappa))\frac{C}{L}}{1+p^*(1-\pi(1-\kappa))\frac{C}{L}}$.

As potential equilibria we therefore have

$$W = \begin{cases} W^* = 0 & \text{if } \kappa = \kappa^* = 0 \\ W^{**} & \text{if } \kappa = \kappa^{**} \\ W^{***} & \text{if } \kappa = \kappa^{***} = 1 \end{cases}. \tag{7.9}$$

Although banks make their decision on the amount of auditing resources at the time loans are given and companies decide to default only when the loan is due to be repaid, that is once the bank has already committed these auditing resources, we assume that companies do not know the amount of auditing resources a bank has committed. We can therefore treat the decision-making of companies and banks as simultaneous. This gives rise to a strategic interaction between banks and companies in the commitment of audit resources and strategic default, respectively, which we solve for its equilibrium.

Equilibrium The bank can decide on one of three levels of audit resources, W^*, W^{**}, or W^{***}, and the company can decide to not default strategically with certainty (κ^*), strategically default with probability κ^{**}, or to default strategically with certainty, κ^{***}. We can now distinguish nine possible combinations of these decisions and with the relevant parameters inserted, as well as taking into account whether companies default strategically, the profits that companies can achieve can be shown to fulfill the following inequalities:

$$\begin{aligned} \Pi_C\left(W^*,\kappa^*\right) < \Pi_C\left(W^*,\kappa^{**}\right) < \Pi_C\left(W^*,\kappa^{***}\right), \\ \Pi_C\left(W^{**},\kappa^*\right) < \Pi_C\left(W^{**},\kappa^{**}\right) < \Pi_C\left(W^{**},\kappa^{***}\right), \\ \Pi_C\left(W^{***},\kappa^*\right) > \Pi_C\left(W^{***},\kappa^{**}\right) > \Pi_C\left(W^{***},\kappa^{***}\right). \end{aligned} \tag{7.10}$$

We have added arguments W and κ to the company profits Π_C for notational clarity. Choosing κ^{**} is never a best response for the company and can for this reason be eliminated from further considerations as no company would make this choice. Similarly, having eliminated the possibility of companies choosing κ^{**}, for banks we obtain their profits to fulfill these inequalities:

$$\begin{aligned} \Pi_B\left(W^*,\kappa^*\right) > \Pi_B\left(W^{**},\kappa^*\right) > \Pi_B\left(W^{***},\kappa^*\right), \\ \Pi_B\left(W^*,\kappa^{***}\right) < \Pi_B\left(W^{**},\kappa^{***}\right) < \Pi_B\left(W^{***},\kappa^{***}\right). \end{aligned} \tag{7.11}$$

We have added arguments W and κ to the bank profits Π_B for notational clarity. The commitment of audit resources W^{**} is never a best response; it will therefore never be chosen by banks. This leaves us with banks choosing either audit resources $W^* = 0$ or $W^{***} = \frac{(1-\pi(1-\kappa))\frac{C}{L}}{1+(1-\pi(1-\kappa))\frac{C}{L}}$ as well as companies choosing to either never default strategically, $\kappa^* = 0$, or to always default strategically, $\kappa^{***} = 1$. In these cases the profits of the company and the bank, respectively, can be ordered as follows:

$$\begin{aligned} \Pi_C\left(0,1\right) > \Pi_C\left(0,0\right) = \Pi_C\left(W^{***},0\right) > \Pi_C\left(W^{***},1\right), \\ \Pi_B\left(0,0\right) > \Pi_B\left(W^{***},0\right) > \Pi_B\left(W^{***},1\right) > \Pi_B\left(0,1\right). \end{aligned} \tag{7.12}$$

No equilibrium in pure strategies exists in this strategic game between the bank and company, hence we have to determine a mixed strategy equilibrium. Defining the probability of the bank not committing audit resources as λ and the company to not strategically default as μ, we obtain that

$$\begin{aligned}
&\lambda \Pi_C(0,0) + (1-\lambda)\Pi_C(W^{***},0) \\
&\qquad = \lambda \Pi_C(0,1) + (1-\lambda)\Pi_C(W^{***},1), \\
&\mu \Pi_B(0,0) + (1-\mu)\Pi_B(0,1) \\
&\qquad = \mu \Pi_B(W^{***},0) + (1-\mu)\Pi_B(W^{***},1).
\end{aligned} \tag{7.13}$$

The right hand side shows the expected profits when choosing to not default strategically (for companies) and to not commit any audit resources (for banks), while the right-hand side shows the expected profits of choosing to default strategically (for companies) and commit audit resources W^{***} (for banks). The expected profits are calculated taking into account that the decision by the bank (company) is not known to the company (bank), but only the probability of the decision is known. In equilibrium the bank (company) would be indifferent about the decision the company (bank) makes.

Inserting for the expected profits, these equations easily solve for

$$\begin{aligned}
\lambda &= \frac{C}{(1+r_L)L + C}, \\
\mu &= \frac{\pi(1+r_L)L - (1-\pi)C}{\pi(1+r_L)(L+C) - (1-\pi)C}.
\end{aligned} \tag{7.14}$$

As $\mu < 1$, we see that strategic default occurs at a rate of $E[\kappa] = 1 - \mu = \frac{\pi(1+r_L)C}{\pi(1+r_L)(L+C)-(1-\pi)C}$. It is straightforward to see that higher auditing costs C increase strategic default as the higher auditing costs will reduce the number of audits a bank can conduct for given resources, increasing the likelihood of a strategic default remaining undetected. While the bank would increase the audit resources W^{***} and thus increase the likelihood λ of committing these resources, this effect only partially offsets the smaller number of audits it will be able to conduct.

If the success rates of investments are higher, the number of unsuccessful companies reduces, banks will more likely audit companies that default strategically, making it more likely that such defaults are detected. Hence strategic default is less attractive to companies as they are more likely to have to compensate the bank for their audit costs, even though banks reduce the auditing resources they commit.

Summary Strategic default can occur if the company anticipates that the auditing resources a bank will have available is not sufficient to conduct audits of all companies defaulting. Thus successful companies attempt to conceal their strategic default in the default of unsuccessful companies. If successful companies coordinate their strategic defaults, they can exhaust the audit resources of banks, lowering the probability of being detected. This coordination is limited due to the costs detection imposes on those companies that are strategically defaulting and are audited. While strategic default will be low for most realistic parameter settings, it will nevertheless be present due to imperfect auditing of companies defaulting.

Reading Krause (2022b)

7.2 The impact of future borrowing on strategic default

Companies continuously make investments, either updating existing projets or developing new investment opportunities. In many cases for each investment a new loan contract is signed and new loan conditions are agreed between the bank and the company. After the initial investment, the company needs to decide whether to repay the loan, assuming it is able to do so, or to default on its loan.

Let us assume that a company has the opportunity to pursue an investment requiring a loan of L at loan rate r_L and has access to identical investments for two time periods. In each time period the outcome will be either V_H with probability π_H or V_L with probability $\pi_L = 1 - \pi_H$. We assume $V_H > (1 + r_D) L > V_L$ such that the loan does not cover its funding costs r_D from deposits in full if the low outcome V_L is realised, but if the high outcome V_H is realised, the loan covers its costs. For convenience we define $\overline{V} = \pi_H V_H + \pi_L V_L > (1 + r_D) L$ as the expected value of the investment outcome; the investment is efficient in that the expected outcome exceeds the funding costs of the banks as represented by the deposit rate r_D. The loan rate r_L will have to cover at least the funding costs from deposits, r_D, for the bank to be profitable. Hence in the case of the low outcome V_L being realised, the loan cannot be repaid in full. The bank financing the loan only commits itself to financing it for the first time period and can decide whether to renew the loan for the second time period once it learns the outcome from the first time period.

As we assume that there is no possibility for the bank to verify the investment outcome the company declares, the company is free to declare the low outcome V_L and avoid repaying the loan in full, even though it realized the high outcome V_H. If the company declares the low outcome V_L and thus does not repay the loan in full, we assume that the bank will provide a loan to finance the investment in the second time period with probability p_L and if the company declares the high outcome V_H and repays the loan in full, the second investment will be financed by the bank through the provision of a loan with probability $p_H > p_L$.

The repayments in time period 1, depending on the outcome declared by the company, will be denoted R_H^1 and R_L^1, respectively, and for time period 2 R_H^2 and R_L^2, where we can easily show that $R_H^2 = R_L^2 = V_L$. This is because there is no incentive for the company to declare to have received the high outcome V_H and repay the loan in full. The company will declare the low outcome V_L as this reduces the repayment of the loan to $V_L < (1 + r_L) L$; the bank then obviously insists on the highest possible payment, V_L. The repayments in time period 1 also need to be affordable, thus we require $V_i \geq R_i^1$, implying that if the low outcome V_L is declared the repayment is $R_L^1 = V_L$ and for the high outcome it is $R_H^1 = (1 + r_L) L$.

Neglecting discounting, the expected profits of the bank when assuming that companies repay their loans if they are able to, are given by

$$\begin{aligned} \Pi_B &= \pi_H \left((1 + r_L) L + p_H \left(V_L - (1 + r_D) L \right) \right) \\ &\quad + \pi_L \left(V_L + p_L \left(V_L - (1 + r_D) L \right) \right) - (1 + r_D) L \\ &= \pi_H (1 + r_L) L + \pi_L V_L - (1 + r_D) L \\ &\quad + (\pi_H p_H + \pi_L p_L) \left(V_L - (1 + r_D) L \right). \end{aligned} \tag{7.15}$$

Initially the bank finances the loan L and fully with deposits that attract an interest rate of r_D. This initial loan is successfully generating the high outcome V_H with probability π_H, resulting in the loan being repaid in full. Based on this outcome a loan is provided in the second time period, again financed by deposits, on which the repayment to the bank will be V_L, either because the investment fails or because the company declares the low outcome V_L, even if the investment was successful. If the initial investment is not successful, the bank will only obtain the low outcome V_L as loan repayment and extend the loan with probability p_L. This loan then in return yields a repayment of the low outcome V_L for the same reasons as before.

The profits of a company declaring their investment outcomes truthfully, are obtained as

$$\begin{aligned} \Pi_C &= \pi_H \left(V_H - (1+r_L) L + p_H \left(\overline{V} - V_L\right)\right) \\ &\quad + \pi_L \left(V_L - V_L + p_L \left(\overline{V} - V_L\right)\right) \\ &= \pi_H (V_H - (1+r_L) L) + (\pi_H (p_H - p_L) + p_L) \left(\overline{V} - V_L\right), \end{aligned} \tag{7.16}$$

where we used that $\pi_L = 1 - \pi_H$ in the second equality. The investment is successful with probability π_H and generates the high outcome V_H from which the company repays the loan in full; it then obtains a loan for the second investment with probability p_H. The expected outcome of this investment is $\overline{V}$, from which it will repay the declared low outcome V_L. If the initial investment is not successful, it will obtain the low outcome V_L, which is paid to the bank, and then obtains a second loan as above.

A company not declaring the investment outcome truthfully in time period 1 will only repay V_L on their initial loan, but also obtain a second loan only with probability p_L. Hence the profits of these companies are given by

$$\begin{aligned} \hat{\Pi}_C &= \pi_H \left(V_H - V_L + p_L \left(\overline{V} - V_L\right)\right) \\ &\quad + \pi_L \left(V_L - V_L + p_L \left(\overline{V} - V_L\right)\right) \\ &= \pi_H (V_H - V_L) + p_L \left(\overline{V} - V_L\right), \end{aligned} \tag{7.17}$$

where we used that $\pi_L = 1 - \pi_H$ in the second equality.

In order to avoid a strategic default, i.e. the company declaring to have V_L when it has received V_H, we require that $\Pi_C \geq \hat{\Pi}_C$. Using equations (7.16) and (7.17), this becomes

$$(p_H - p_L) \left(\overline{V} - V_L\right) \geq (1 + r_L) L - V_L \tag{7.18}$$

As the bank will seek to extract the highest possible loan rate from the company, while avoiding the company to default strategically, and hence the right-hand side will become as high as possible, leading to this constraint becoming an equality. Inserting this equality into equation (7.15) after solving for $(1 + r_L) L$, we obtain the bank profits as

$$\Pi_B = V_L - (1 + r_D) L + \pi_H p_H \left(\overline{V} - (1 + r_D) L\right) \quad (7.19)$$
$$-p_L \left(\pi_L (1 + r_D) L + \pi_H \overline{V} - V_L\right).$$

The second expression is positive as we assumed that the average outcome, $\overline{V}$, will be sufficient to cover the funding costs of banks through deposits, $\overline{V} \geq (1 + r_D) L$. In this case it would be efficient for banks to provide loans to companies as the outcome these loans generate exceed the funding costs. With the second term positive, it would be optimal for the bank to choose the highest possible probability of providing a loan to the company if the first investment succeeds, thus $p_H = 1$. The final term will also be positive; to see this recall that $(1 + r_D) L \geq V_L$ and as we furthermore assumed that $\overline{V} \geq (1 + r_D) L$, we can, using that $\pi_L + \pi_H = 1$, rewrite the expression in brackets as $(1 + r_D) L - V_L + \pi_H \left(\overline{V} - (1 + r_D) L\right)$ and with the above mentioned assumptions see that the both differences are positive. With the final term being positive, the bank profits are maximised if the probability of banks providing loans to companies whose initial investment is not successful is as low as possible, thus we set $p_L = 0$.

Banks will therefore always provide loan to companies with previously successful investments, but will not provide loans if the investment has not been successful. This strategy of granting loans provides an incentive for companies to avoid strategic default. In order to ensure this avoidance of strategic default, banks cannot extract all surplus from companies if the investment is successful as equation (7.18) implies for $p_H = 1$ and $p_L = 0$ that $(1 + r_L) L = \overline{V}$ and hence companies repay $\overline{V}$ if the investment is successful, earning them profits of $V_H > \overline{V}$.

In case of the investment being successful, V_H, the bank will extend a loan, even though it is aware that it will make a loss in the second time period. The reason for this willingness to extend the loan is that it induces the company to not default in time period 1 if it realizes this high outcome. Hence a commitment to continue lending to a non-defaulting company avoids strategic default.

The size of the loan is restricted due to the requirement of bank profits to be positive. Inserting our results into equation (7.18), the condition that $\Pi_B \geq 0$ becomes

$$(1 + r_D) L \leq \frac{V_L + \pi_H \overline{V}}{1 + \pi_H} < \overline{V}, \quad (7.20)$$

where the final inequality arises from $V_L < \overline{V}$. This induces an inefficiency in that the deposit rate r_D is strictly less than the expected outcome of the investment, $\overline{V}$, even if the bank makes no profits and the relationship in equation (7.20) is fulfilled with equality. This arises from the need to retain some profits from lending in the first time period to compensate the bank for the losses emerging from the default of the companies in the second time period.

For companies these contractual arrangements are always profitable as inserting all results into equation (7.16) yields $\Pi_C = \pi_H (V_H - V_L) > 0$. We see that these profits are entirely based on the second time period, where the strategic default from a successful investment gives the company an outcome of V_L, from which it repays V_L, hence a profit of $V_H - V_L$, which is only realised if the investment is successful, which occurs with probability π_H. Profits in the first time period are extracted by the bank to cover their losses from the loan not being repaid fully in the second time period.

Companies are avoiding strategic default to secure a loan for follow-on investments in the second time period and thereby retaining the possibility of generating additional profits. Of course, if over time the economic conditions change during the first time period, the constraint in equation (7.18) might no longer be fulfilled. This could be the case if the investment outcomes are reduced, such as in recessions, increased competition for the company or additional regulatory burdens. In these cases, strategic default might be observed.

Reading Bolton & Scharfstein (1990)

7.3 Optimal exclusion length

After a company fails to repay its loan, it is common to assume that the company will be excluded from borrowing permanently. However, by excluding the company from obtaining loans, banks are reducing their own profits as they can no longer lend to this company. The exclusion of companies from future loans can be justified as a measure by banks to provide incentives to repay loans and not strategically default. However, the exclusion from loans does not only affect those companies that strategically defaulted, but also those whose investments genuinely failed. In many legislations, bankruptcy of companies or individuals imposes restrictions on borrowing for a certain period of time, amongst other constraints on their activities, but often borrowing can be resumed after the required time has elapsed.

Let us assume a company successfully completes an investment with probability π, giving a return on investment of $R \geq 0$. If the investment is not successful, which happens with probability $1 - \pi$, the company receives no revenue at all. Such an investment is available in each time period, *ad infinitum* and future revenue is discounted with a discount factor $\rho < 1$. The investment is fully financed by a bank loan of size L with interest r_L; in the case of an unsuccessful investment, this loan cannot be repaid. Furthermore, if the company does not repay the loan, it will be excluded from any further borrowing for $T \geq 0$ time periods.

As only companies whose investments are successful can repay the loan, it is only such companies that can consider to default strategically. Hence we only consider companies who in the current time period have completed investments successfully. A company with such a successful investment that repays its loan will make profits $\Pi_C = (1 + R) L - (1 + r_L) L$ in this time period and hence the value of the company will be given by

$$V_C = \Pi_C + \rho\pi V_C + (1 - \pi) \rho^T V_C, \tag{7.21}$$

where the first term represents the profits of the successful company in the current time period. The future discounted profits are given in the following two terms;

the second term covers the case where the subsequent investment is also successful and therefore the company will continue to receive loans and be able to make these investments, generating value V_C in the future, while the third term denotes the case of an unsuccessful investment that yields no revenue in the next period, and is then followed by exclusion for T time periods, after which the investment may resume and the company generates value V_C again.

Using the definition of Π_C, we can solve equation (7.21) such that we obtain

$$V_C = \frac{R - r_L}{1 - \pi\rho - (1 - \pi)\,\rho^T} L. \tag{7.22}$$

The bank cannot verify the cause of a loan not being repaid, a successful company could claim to have been unsuccessful to avoid repaying the loan, a strategic default. This would result in profits of $(1 + R)\,L$ in the current time period as the loan is not repaid, and the resumption of borrowing after the exclusion period in T time period, valued at $\rho^T V_C$. We here assume that the company resumes repaying their loans in the future and plans to strategically default only in the current time period. The value of the company if defaulting strategically is thus given by

$$\hat{V}_C = (1 + R)\,L + \rho^T V_C. \tag{7.23}$$

To avoid such strategic default, we require that the value of a company repaying the loan exceeds that of a company defaulting strategically, $V_C \geq \hat{V}_C$, which gives us the condition that

$$1 + r_L \leq 1 + r_L^* = \rho\pi \frac{1 - \rho^{T-1}}{1 - \rho^T} (1 + R)\,. \tag{7.24}$$

Banks fully finance their loans L by deposits on which an interest rate of r_D is payable, giving them expected profits of $\Pi_B = \pi\,(1 + R_L)\,L - (1 + r_D)\,L$ in each time period, taking into account that loans are only repaid by companies with successful investments. Analogously to equation (7.22) the bank value is then given by

$$V_B = \Pi_B + \rho\pi V_B + (1 - \pi)\,\rho^T V_B, \tag{7.25}$$

which using the definition of Π_B solves for

$$V_B = \frac{\pi\,(1 + r_L) - (1 + r_D)}{1 - \pi\rho - (1 - \pi)\,\rho^T} L. \tag{7.26}$$

Let us now assume that competitive forces between banks are such that banks generate no economic profits, hence $V_B = 0$. This then implies that the loan rate is given by

$$1 + r_L^{**} = \frac{1 + r_D}{\pi}. \tag{7.27}$$

As the bank value increases in the loan rate r_L, banks would choose the highest possible loan rate that avoids strategic default. This loan rate is given in equation

(7.24) and inserting this equation and equalling it with the competitive loan rate from equation (7.27), we easily obtain the optimal exclusion period as

$$T^* = \frac{\ln \frac{\pi^2\rho(1+R)-(1+r_D)}{\pi^2(1+R)-(1+r_D)}}{\ln \rho}. \tag{7.28}$$

Hence it is optimal for the exclusion period to be limited in time as $T^* < +\infty$ as long as $\pi^2\rho\,(1+R) > (1+r_D)$, which we assume to be fulfilled. This allows a trade-off between avoiding strategic default and the bank generating profits from the company through the provision of future loans.

We easily obtain that

$$\begin{aligned} \frac{\partial T^*}{\partial \pi} &< 0, \\ \frac{\partial T^*}{\partial\,(1+R)} &< 0. \end{aligned} \tag{7.29}$$

We thus find that more risky companies are excluded from the loan market for longer, while more profitable companies face shorter exclusions. The reason for these findings is that more risky companies generate less profits for the bank due to the more frequent unsuccessful investments, thus excluding such companies for longer does not affect the bank as negatively as companies with higher success rates. Companies with more profitable investments are defaulting less likely strategically as they would lose these high profits during the exclusion period, allowing banks to reduce this time period without inducing strategic default.

Banks optimally exclude companies defaulting only for a limited period of time. While exclusion from lending ensures companies have less incentives to default strategically, it also reduces the potential future profits of banks from lending to such companies. Banks can balance the generation of future profits with the incentives to avoid strategic default by limiting the length companies are excluded from borrowing. If this exclusion period is sufficiently long, it deters strategic defaults by companies, while allowing the bank to earn future profits from continuing to lend to these companies.

Reading Krause (2022a)

Conclusions

Even if their investment is successful, companies have incentives to default on their loan if the benefits of doing so outweighs the cost. The benefits are usually immediate in the form of the loan repayments not being required. The costs of such strategic default can be the exclusion from future loans and thus foregone profit opportunities if investment cannot be made. As banks are unaware whether a default is the result of a genuine inability to repay the loan due to failing investments or strategic default, banks lose substantial future profits by excluding all companies that fail to repay their loans. Maintaining the benefits of reducing future profits to companies defaulting on their loan, banks would optimally only exclude them for a specific period of time, thereby imposing some losses on the company, but retaining their ability to retain future profits from lending to this company.

Strategic default can best be mitigated through auditing of defaulting companies, which should detect whether a company is defaulting strategically or unable to repay its loan. However, auditing is costly to banks, and thus with limited resources available to banks, not all defaults can be audited. This gives companies an incentive to exploit the limited resources banks are willing to commit to auditing. Banks will balance the costs of committing such resources against an the frequency of strategic defaults and the associated losses.

Hence, while auditing will be able to reduce the instances of strategic default, it cannot eliminate them completely. In addition, audits might not be able to detect all strategic defaults as companies have many ways to reduce their ability to repay loans without reducing the wealth to its owners. Using exclusion from borrowing for a specific time period imposes costs on companies to default strategically and might be the most effective way of addressing this possibility. However, changing economic conditions might make strategic default more attractive to companies, especially if an effective auditing system has not been established.

Chapter 8

Credit rationing

Companies often apply for loans that are larger than what banks are willing to provide them with, even when taking into account the loan rate they are willing to pay. This gives rise to a situation in which companies obtain a loan that is smaller than what they applied for and even when offering a higher loan rate, banks do not increase the size of the loan offer. We therefore face a situation where the demand for loan exceeds the supply of loans. Following conventional economic theory, in equilibrium the demand and supply for loans should be balanced and the price, the loan rate, be used as a toll to achieve such a balance. If, however, excess demand for loans cannot be eliminated through increasing the loan rate, this excess demand can be interpreted as an equilibrium. We refer to such an equilibrium as credit rationing.

Credit rationing occurs if banks are not meeting the demand of companies for loans, even if they are offered higher loan rates by companies. It thus has to be the case that the profits of banks are higher with a smaller loan at a lower loan rate. Such a situation can arise if banks are less likely to be repaid the larger loan, reducing their profits even if the loan rate would be higher. As chapter 8.1 will show, this can be the result of companies defaulting more often due to a higher leverage of the company when obtaining a larger loan. Alternatively we will see in chapter 8.2 that companies increasing the risk of their investments if they are granted larger loans, will also affect banks negatively and might induce them to limit the size of the loan in order that companies are making low-risk investments. Finally, banks my limit the size of loans as to prevent strategic default by companies as chapter 8.3 shows. Finally, chapter 8.4 investigates the effect competition has on credit rationing.

8.1 The consequences of uncertain outcomes

Companies can fund their investments using their own funds, equity, or a bank loan. If we assume that there are no constraints on the availability of equity, companies will choose the optimal combination of these two funding sources. Of course, when deciding on the size of the loan they seek, companies will take into account the loan rate they are offered. Banks, providing such loans, will consider the ability of the company to repay their loan. With the outcomes of investments uncertain, banks cannot be sure to be repaid their loan and will take into account the possibility of default when offering loans. This default will not only be taken into account when setting the loan rate, but also when deciding the size of the loan. A larger loan implies a higher repayment is required to the bank, which required the company to obtain a higher return on its investment to avoid default. Banks will seek to balance these possible defaults in their loan offers with the profits they obtain in cases where the loan is repaid.

We assume that companies make investments I, financed though a combination of bank loans $L \geq I$ and equity E, such that $I = L + E$. The expected investment yields a return of R if it is successful and no return otherwise, where success is achieved with probability π. This probability as well as the return in the case of a successful investment are not known in advance to either the bank or the company; however, it is known that the expected outcome, $\pi (1 + R) I$ has a distribution function $F(\cdot)$. Companies will obtain the outcome only once they have repaid their bank loan, including interest r_L and hence their profits are given by

$$\begin{aligned} \Pi_C &= \int_{(1+r_L)L}^{+\infty} \pi (1 + R) L dF (\pi (1 + R) L) - E \qquad (8.1) \\ &= L + \int_{(1+r_L)L}^{+\infty} \pi (1 + R) L dF (\pi (1 + R) L) - I. \end{aligned}$$

For a given loan rate, the optimal amount of bank loans will be given by maximizing their profits and solving the first order condition $\frac{\partial \Pi_C}{\partial L} = 0$, we obtain that $(1 + r_L)^2 Lf (\pi (1 + r_L) L) = 1$, where $f(\cdot)$ denotes the density function. We clearly see that the loan demand is decreasing in the loan rate.

Companies will only demand loans if it is profitable to do so, thus $\Pi_C \geq 0$. It is obvious that the bank profits are decreasing in the loan rate as the lower boundary of the integration in equation (8.1) is increasing. Hence let us define $\overline{r}_L$ as the loan rate at which the company breaks even, $\Pi_C = 0$. We then have

$$\begin{aligned} \frac{\partial \Pi_C}{\partial L} &= 1 - (1 + \overline{r}_L)^2 Lf (\pi (1 + \overline{r}_L) L), \qquad (8.2) \\ \frac{\partial \Pi_C}{\partial (1 + \overline{r}_L)} &= -(1 + \overline{r}_L) L^2 f (\pi (1 + \overline{r} + L) L). \end{aligned}$$

Using the implicit function theorem, we easily get that

$$\frac{\partial (1 + \overline{r}_L)}{\partial L} = -\frac{\frac{\partial \Pi_C}{\partial L}}{\frac{\partial \Pi_C}{\partial (1+\overline{r}_L)}} = \frac{1 - (1 + \overline{r}_L)^2 Lf (\pi (1 + \overline{r}_L) L)}{(1 + \overline{r}_L) L^2 f (\pi (1 + \overline{r} + L) L)}. \qquad (8.3)$$

The bank will obtain the outcome of the investment if the company cannot repay its loan in full and if the outcome is sufficiently high, will repaid the loan, where we know that the highest possible loan rate is given by $\overline{r}_L$ for companies to demand loans. If we assume that loans are financed fully by deposits with a deposit rate r_D, the bank profits are given by

$$\begin{aligned}
\Pi_B = & \int_0^{(1+r_L)L} \pi(1+R)\, LdF(\pi(1+R)L) \\
& + \int_{(1+r_L)L}^{(1+\overline{r})L} (1+r_L)\, LdF(\pi(1+R)L) - (1+r_D)L \\
= & \int_0^{(1+r_L)L} \pi(1+R)\, LdF(\pi(1+R)L) \\
& + (F((1+\overline{r}_L)L) - F((1+r_L)L))(1+r_L)L - (1+r_D)L.
\end{aligned} \tag{8.4}$$

Using the Leibniz integral rule, we easily obtain that

$$\begin{aligned}
\frac{\partial \Pi_B}{\partial(1+r_L)} &= (F((1+\overline{r}_L)L) - F((1+r_L)L))L > 0, \\
\frac{\partial \Pi_B}{\partial L} &= (F((1+\overline{r}_L)L) - F((1+r_L)L))(1+r_L) \\
& \quad + \frac{1+r_L}{1+\overline{r}_L} - (1+r_D).
\end{aligned} \tag{8.5}$$

The first term is positive as $\overline{r}_L \geq r_L$ and hence the term in bracket must be positive. The second term will be negative for some $L \geq \hat{L}$. This is because if the amount lend is very small, then $(1+\overline{r}_L)L \approx (1+r_L)L$ and hence $F((1+\overline{r}_L)L) \approx F((1+r_L)L) \approx 0$, while the second term will be less than 1 due to $r_L \leq \overline{r}_L$ and hence the second and final term will be jointly negative. Similarly, for very large bank loans, we have $F((1+\overline{r}_L)L) \approx F((1+r_L)L) \approx 1$, and the first time vanishes again, making the expression negative for $L > \hat{\hat{L}}$. For intermediate sizes of bank loans, this expression might well be positive as long as $F((1+\overline{r}_L)L)$ is sufficiently larger $F((1+r_L)L)$. Hence the expression is positive if $\hat{L} < L \leq \hat{\hat{L}}$.

Assuming that banks are competing such that $\Pi_B = 0$, we can use the implicit function theorem to get

$$\frac{\partial(1+r_L)}{\partial L} = -\frac{\frac{\partial \Pi_B}{\partial L}}{\frac{\partial \Pi_B}{\partial(1+r_L)}}, \tag{8.6}$$

which is positive for $L \leq \hat{L}$ and $L > \hat{\hat{L}}$ and negative for $\hat{L} < L \leq \hat{\hat{L}}$. Figure 8.2 shows this relationship between the loan rate and the amount of loans offered. We clearly see that the loan rate is not monotonically increasing in the amount of loans offered, but downward slowing for an intermediate range of loan rates. This is the case because as loan rates are increased, the amount the company needs to repay will also increase; such an increased repayment be possible for some outcomes and banks reduce the size of the bank loan to avoid the company defaulting.

Banks will maximize their profits by choosing the optimal loan repayment, $(1+r_L)L$. The first order condition $\frac{\partial \Pi_B}{\partial(1+r_L)L} = 0$ solves for

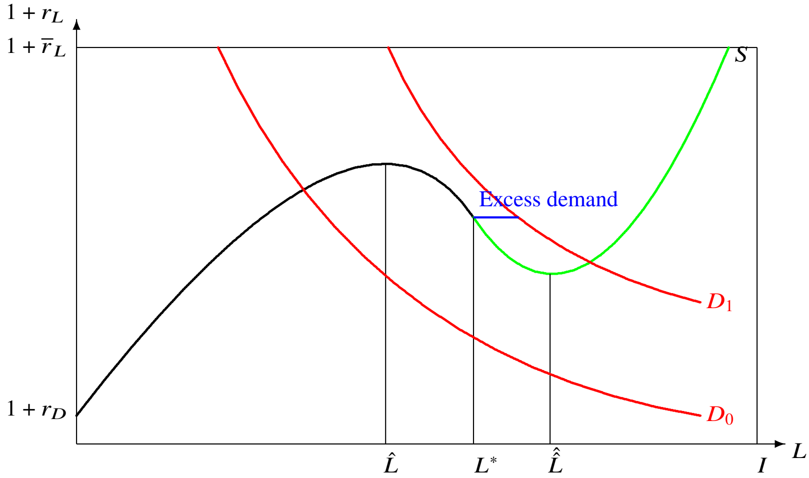


Fig. 8.1: Credit rationing due to uncertain outcomes

$$1 + r_D = \left(F\left(\left(1 + \overline{r}_L\right) L\right) - F\left(\left(1 + r_L\right) L\right)\right)\left(1 + r_L\right), \tag{8.7}$$

where we used that $\frac{\partial L}{\partial(1+r_L)L} = \frac{1}{\frac{\partial(1+r_L)L}{\partial L}} = \frac{1}{1+r_L}$. Inserting this optimal solution into equation (8.5), we easily obtain that $\frac{\partial \Pi_B}{\partial L} = \frac{1+r_L}{1+\overline{r}_L} > 0$ and hence the optimal amount lend, L^*, will be such that $\hat{L} < L^* \leq \hat{\hat{L}}$. Providing larger loans would reduce profits to the bank and they would therefore not be doing so, thus there will be no supply of loans beyond L^*. This has direct implications for the equilibrium loan amount.

If the loan demand is low, indicated by D_0 in figure 8.2, then an equilibrium can easily be found where demand equals supply. However, if the demand increases to D_1, we see that demand and supply only meet at a point which would require a loan size exceeding the optimal loan size for the bank, L^*, which they therefore would not offer; this area of the loan supply is indicated in green. Banks would only offer a loan of size L^*. However, at this point, the demand for loans exceeds that of the supply of loans, causing loans to be rationed.

In times of low demand, an equilibrium can be reached in which demand for loans and their supply are matched, even though the bank supplies less than their optimal amount of loans. They would not be able to provide their optimal size of loans, L^*, as this would necessitate a loan rate that would not be profitable. The supply curve S in figure 8.2 represents the line in which bank profits are equal and any point below this line would cause the bank to make losses. With the demand at L^* requiring a lower loan rate, the bank would make a loss. Thus the equilibrium would be at the point demand and supply equal. If the demand is high, D_1, demand and supply are equal only for a loan size $L > L^*$, but as the bank would not offer loans above L^*, this cannot be an equilibrium. Banks will offer their optimal loan size L^* and competition between banks ensures that the loan rate associated with this loan offer is not raised, but this results in an excess demand for loans as companies would prefer to obtain larger loans at that loan rate. The competition between banks prevents them from raising the loan rate to a level where the demand for loans by companies would be L^*. The result is an equilibrium with credit rationing; companies are allocated a lower loan than they demand even though they would be willing to pay a higher loan rate.

If banks are less competitive, the supply curve would shift upwards as banks will be able to make some profits, this might alleviate credit rationing, although if the demand would increase further, credit rationing would emerge again.

We thus see that in times of high demand for loans, credit rationing may occur and companies cannot secure the amount of loans they seek at the loan rate they are quoted by banks. Such credit rationing emerges from the uncertainty of the investments companies conduct and hence the uncertainty about the repayment of the loan to banks. Providing companies with larger loans increases the amount that needs to be repaid, making a default more likely as the company needs to obtain a larger return on their investments than with a lower loan rate. In order to reduce defaults, banks may lower the loan rate and thereby lower the amount the company needs to repay, balancing these two aspects to maintain their profitability.

Readings Stiglitz & Weiss (1981), Arnold & Riley (2009)

8.2 Credit rationing caused by moral hazard

Companies can often choose between investments with different risk profiles. This indices a moral hazard in that banks would prefer companies to choose investments of lower risk, as this increases the chances of the loan bing repaid, while for companies it might be more profitable to choose a more risky investment. Banks can use their loan conditions to provide incentives for companies to make the low-risk investments they prefer. When taking the incentives of companies into account, banks might find themselves in a situation where they cannot make more profits from changing the loan conditions without companies changing to more risky investments. This might lead to a mismatch between the loan conditions offered by banks and the loan conditions companies would be willing to accept.

Let us assume that companies have the choice between two investments, one yields a return of R_H if the project is successful, which happens with probability π_H, while the other investment yields $R_L > R_H$ if successful, which occurs with probability $\pi_L < \pi_H$. In both cases an unsuccessful investment yields no return. However, if the investment is successful, the return on the high-risk investment is higher. Banks are aware that companies have these two investment opportunities, but are not able to influence the decision of the company directly.

Companies have limited liability and we assume that the investment is fully financed through a bank loan L on which a loan rate of r_L is payable. With companies only able to repay the loan of their investment is successful, their profits are given by

$$\Pi_C^i = \pi_i \left(\left(1 + R_i\right) L - \left(1 + r_L\right) L \right). \tag{8.8}$$

They will seek the low-risk investment H over the high-risk investment L if it is more profitable to do so, hence we require that $\Pi_C^H \geq \Pi_C^L$. This easily solves for

$$1 + r_L \leq 1 + \hat{r}_L = \frac{\pi_H \left(1 + R_H\right) - \pi_L \left(1 + R_L\right)}{\pi_H - \pi_L}. \tag{8.9}$$

As long as the loan rate is not too high, companies will prefer to choose the low-risk investment.

Banks are repaid the loan if the investment is successful and they themselves have to repay deposits on which interest r_D is payable. With loans fully financed by deposits, we obtain the bank profits as

$$\Pi_B^i = \pi_i \left(1 + r_L\right) L - \left(1 + r_D\right) L, \tag{8.10}$$

depending on the choice if investments made by the company. The bank knows that for $r_L \leq \hat{r}_L$ the company chooses the low risk investment and for higher loan rates the high-risk investment. Thus their profits are given by

$$\Pi_B = \begin{cases} \Pi_B^L \text{ if } r_L \leq \hat{r}_L \\ \Pi_B^L \text{ if } r_L > \hat{r}_L \end{cases}. \tag{8.11}$$

The lower right panel in figure 8.2 illustrates this profit function of the bank. We see that at $r_L = \hat{r}_L$ the profits shift downwards as the company switches from the low-risk investment to the high-risk investment and the probability of the loan being repaid reduces. In the area colored green, the bank will make lower profits from charging a higher loan rate due to the switch of investments by the company. Thus, banks would not choose a loan rate in this area. Only once the bank raises the loan rate above r_L^* will their profits from granting loans to companies making high-risk investments be higher than when offering a lower loan rate of $\hat{r}$ and ensuring the company makes the low-risk investment. Of course, loans can only be given if the company makes profits, which from equation (8.8) implies that $r_L \leq R_i$. We assume that $R_H > \hat{r}_L$, which leaves us with the constraint that $r_L \leq R_L$.

The lower left panel shows how the bank's profits evolve with the loan size L if the loan they provide is granted to a company choosing the low-risk investment, H, and the high-risk investment, L, respectively. Using this information, we can now determine the supply curve for the loans of banks as indicated in the upper left panel of figure 8.2. We see that the supply of loans is increasing in the loan rate, however, not all loan rates are feasible. We note that loan rates indicated by the green line correspond to those loan rates where banks obtain a lower profit than when charging $\hat{r}_L$ and hence these loan rates are not offered by banks.

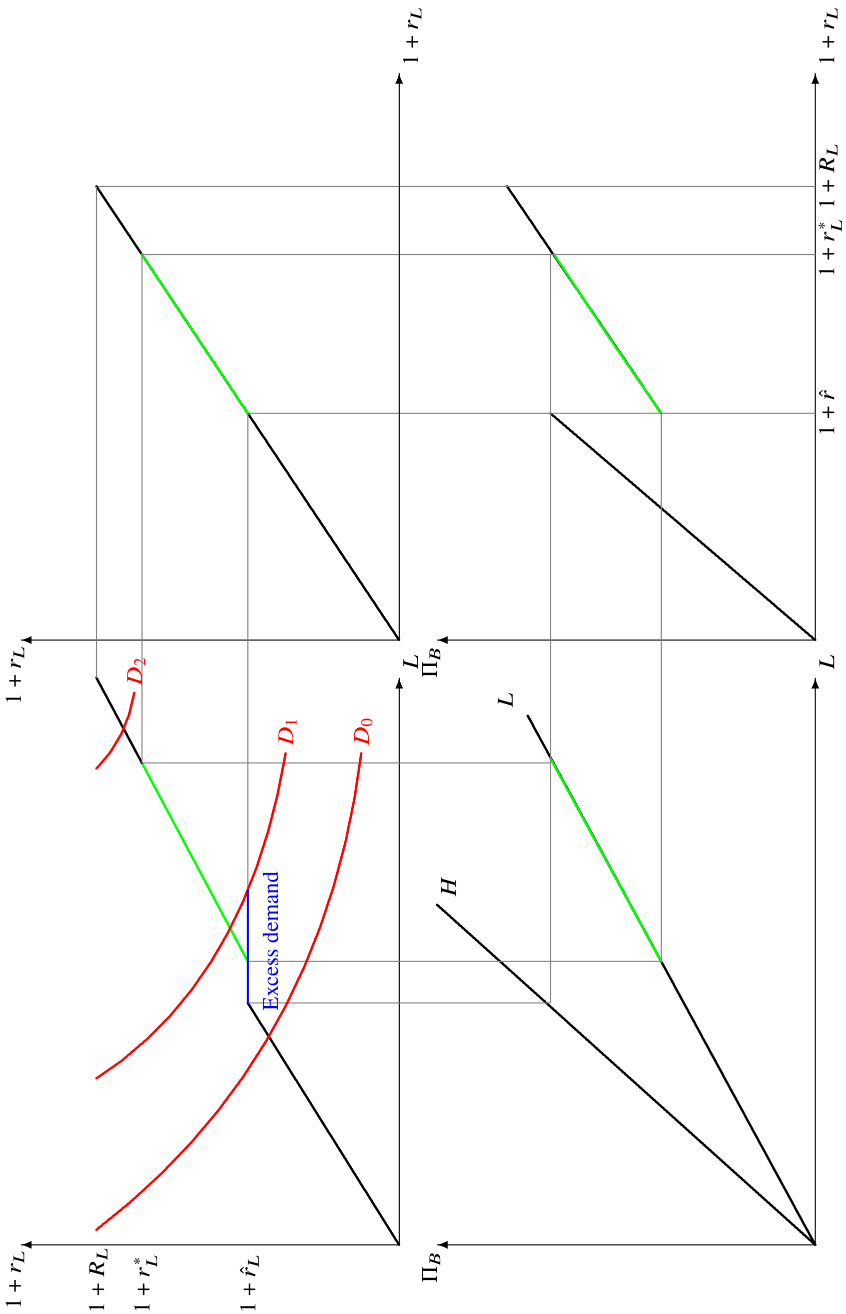


Fig. 8.2: Credit rationing in the presence of moral hazard

The demand curve of an individual company is given by $\frac{\partial \Pi_C^i}{\partial L} = \pi_i \left((1 + R_i) - (1 + r_L)\right) =$ 0 and is as such flat at R_i for any loan size L. However, if we assume that companies overall have access to different investments with different returns, then we easily can establish that the demand curve by companies will have a negative slope as with higher loan rates, more and more investments become unprofitable. The read lines in figure 8.2 indicate such demand for loans by companies. We see that if the demand is low, such as in D_0, we obtain an equilibrium where demand and supply meet. However, if demand is higher at D_1, demand and supply cannot be matched. Demand and supply would be matched at a loan rate where the bank would make lower profits than when charging the lower loan rate $\hat{r}_L$, indicated by the green line. However, if charging this loan rate, the demand for loans exceeds that what banks are willing to supply. We thus observe credit rationing. Only if the demand increases further to D_2 will an equilibrium emerge again at which supply and demand are matched.

We thus see that the moral hazard induced by the company switching from low-risk investments to high-risk investments can cause credit rationing. Banks are not increasing loan rates such that demand and supply are matched as this would induce companies to change their investment into the more risky one, reducing bank profits due to the increased risks. It is therefore that banks maintain the loan rate at the highest level at which the company would choose the low-risk investment, even though the demand by companies is such that they could charge a higher loan rate; this behaviour induces credit rationing.

Reading Bester & Hellwig (1987)

8.3 Credit rationing reducing strategic default

Companies seeking loans have a strong incentive to not repay them, and hence increase their profits, assuming that banks cannot easily verify the true outcomes of any investments they have conducted. The larger the loan the larger the benefits to companies to strategically default on their loans. While it might be optimal for companies to seek large loans in order to obtain the highest possible profits from their investment opportunities, banks might want to restrict the size of their loans in order to ensure companies do not default strategically.

Let us assume that a company obtains a loan L at loan rate r_L and using these funds makes an investments that generates income V with probability π and with probability $1 - \pi$ no income is generated. The size of the outcome in the case the investment is successful, V, will depend on the size of the loan such that $\frac{\partial V}{\partial L} > 0$ and $\frac{\partial^2 V}{\partial L^2} < 0$. Hence the outcome is increasing in the loan size, but this increase is diminishing, for example due to exhausting their investment opportunities. Thus the company profits are given by

$$\Pi_C = \pi \left(V - (1 + r_L)L\right). \tag{8.12}$$

Companies maintain an identical investment opportunity in every time period and if they discount future profits using a discount factor ρ, their total profits are given by

$$\hat{\Pi}_C = \sum_{t=0}^{+\infty} \rho^t \Pi_C = \frac{1}{1-\rho}\Pi_C = \frac{\pi\left(V - (1+r_L)L\right)}{1-\rho}, \tag{8.13}$$

with the last equality arising by inserting from equation (8.12). Companies are maximizing their profits by demanding a loan of the optimal size, thus requiring $\frac{\partial \hat{\Pi}_C}{\partial L} = 0$, which easily solves for

$$\frac{\partial V}{\partial L} = 1 + r_L. \tag{8.14}$$

The optimal loan would be such that the marginal benefits of the loan, $\frac{\partial V}{\partial L}$, equal its marginal costs of repaying the loan, $1 + r_L$.

Banks cannot observe the outcome of the investment, V, and hence the company could declare that its investment was not successful and hence avoid repaying the loan, saving $(1 + r_L)\, L$. This is commonly referred to a strategic default. However, if defaulting on its loan, the bank would not provide them with any future loans, hence they would lose all future profits $\sum_{t=1}^{+\infty} \rho^t \Pi_C = \frac{\rho}{1-\rho}\Pi_C$. Companies would not default if their future profits exceed the instant saving of the loan repayment. We this require $(1 + r_L)\, L \leq \frac{\rho}{1-\rho}\Pi_C$, which solves for

$$(1 + r_L)L \leq \frac{\rho\pi}{1-\rho\,(1-\pi)} V. \tag{8.15}$$

Banks would only provide a loan if they know that their loan is repaid as long as the company is able to do so, thus it wants to ensure that the condition in equation (8.15) is fulfilled. With banks maximizing their profits, they would charge the highest possible loan rate such that equation (8.15) is fulfilled with equality. If the condition in equation (8.15) is not fulfilled, then companies anticipating that banks would not end if strategic default is going to occur, will maximize their profits subject to the constraint in equation (8.15) to ensure banks are providing loans. Thus with a Lagrangian multiplier ξ, we get as our objective function for the company

$$\mathcal{L} = \frac{1}{1-\rho}\Pi_C + \xi\left((1 + r_L)\, L\, (1 - \rho(1-\pi)) - \rho\pi V\right). \tag{8.16}$$

The first order condition for a maximum becomes

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial L} = &\frac{1}{1-\rho}\pi\left(\frac{\partial V}{\partial L} - (1+r_L)\right) \\ &+\xi\left((1 + r_L)\,(1 - \rho(1-\pi)) - \rho\pi\frac{\partial V}{\partial L}\right) \\ = &\ 0, \end{aligned} \tag{8.17}$$

in addition to equation (8.15) being met with equality. We can solve the first order condition (8.17), using equation (8.15) for ξ and obtain

$$\xi = -\frac{1}{\rho(1-\pi)} \frac{\frac{\partial V}{\partial L} - \frac{\rho\pi}{1-\rho(1-\pi)}}{\frac{V}{L} - \frac{\partial V}{\partial L}}, \tag{8.18}$$

which we insert back into equation (8.17) to obtain

$$V = \frac{1-\rho(1-\pi)}{\rho\pi}(1+r_L). \tag{8.19}$$

Using this solution for the successful outcome, we easily get

$$\frac{\partial V}{\partial L} = \frac{1-\rho(1-\pi)}{\rho\pi}(1+r_L) > 1+r_L. \tag{8.20}$$

Therefore the marginal benefits of the investment, $\frac{\partial V}{\partial L}$, will be higher than in the unconstrained optimum as given by equation (8.14). Given the reducing marginal benefits, $\frac{\partial^2 V}{\partial L^2} < 0$, this implies a smaller loan will be optimal.

Hence if the solution $\frac{\partial V}{\partial L} = 1 + r_L$ from equation (8.14) violates the constraint in equation (8.15), then banks will only be willing to provide this smaller loan, while companies would a higher loan, resulting in credit rationing as the demand by companies, exceeds the loan provided by banks. Increasing the loan rate will not alleviate this imbalance in the demand and supply of loans as the inequality in equation $\frac{\partial V}{\partial L} = 1 + r_L$ remain. Thus, an increase in the loan rate would not align the demand and supply as the marginal benefits always have to be higher in the constrained case. In addition, a higher interest rate would make the constraint more binding as we easily see from equation (8.15).

Hence banks will ration credit if for the optimal loan amount leading to $\frac{\partial V}{\partial L} = 1+r_L$ the condition in equation (8.15) is violated. This condition is not fulfilled if the outcome, V, is low or the probability of success, π, is low, implying that loans are rationed for low quality and high risk investments. To see this asserting, let us rewrite the constraint from equation (8.15) as

$$\begin{aligned} \Psi &= (1+r_L)L\,(1-\rho(1-\pi)) - \rho\pi V \\ &= \frac{\partial V}{\partial L}\,(1-\rho(1-\pi)) - \rho\pi V \leq 0, \end{aligned} \tag{8.21}$$

where we used the result on the optimal loan amount for companies from equation (8.14) in the final equality. We then have with $V > (1+r_L)L$ to ensure investments are profitable in case of their success that

$$\frac{\partial \Psi}{\partial \pi} = -\rho\left(V - (1 + r_L)L\right) \leq 0 \tag{8.22}$$

$$\frac{\partial \Psi}{\partial V} = \frac{\partial \Psi}{\partial L}\frac{\partial L}{\partial V} = \frac{\frac{\partial^2 V}{\partial L^2} L\left(1 - \rho(1-\pi)\right) - \rho\pi\frac{\partial V}{\partial L}}{\frac{\partial V}{\partial L}},$$

$$= \frac{\frac{\partial^2 V}{\partial L^2}}{\frac{\partial V}{\partial L}} L\left(1 - \rho(1-\pi)\right) - \rho\pi \leq 0.$$

Reducing the successful outcome V or the probability of success π will increase this term and may therefore more easily lead to a breach of condition (8.15), making the imposition of credit rationing by banks necessary.

We thus see that credit rationing occurs if companies would strategically default if they obtain their optimal loan size; banks reduce the loan size such that defaulting becomes unattractive. Therefore, credit rationing can be used as a tool to avoid strategic default by companies. Such potential strategic default, and hence credit rationing, becomes more prevalent if investments are yield a lower outcome to companies, for example in less profitable companies, or companies take substantial risks. We can therefore expect companies to experience credit rationing in times of recessions or in industries that take substantial risks.

Reading Allen (1983)

8.4 The effect of competition on credit rationing

Banks provide loans to companies and their investments might succeed or fail, imposing risks on the ability of companies to repay their loans. With a high degree of competition between banks, their profits will be low and this might make them more cautious about providing loans to companies and they might offer only a smaller loan than what companies would like to obtain. Through smaller loans, and thus a larger contribution of equity by companies that can absorb at least some losses from failed investments and therefore increase the repayment of loans to banks, they are able to protect their profits, but this may well result in credit rationing. In less competitive markets, such concerns by banks might be less pronounced due to the higher profits banks make due to higher loan rates, which should reduce the prospect of credit rationing to occur.

A company finances its investment I through loans L and equity E such that $I = L + E$, where the amount of equity is exogenously given, allowing the company to increase its investment through loans. It faces one of two possible outcomes of their investments, with probability π the investment is a success and the company achieves a return R_H and with probability $1-\pi$ the investment fails by giving a return of $R_L < R_H$. Let us now assume that with high returns R_H being realised, the loan can always be repaid in full, while for the low return R_L this cannot be guaranteed. We thus assume that $(1 + R_H)\,I > (1 + r_L)\,L$, while $(1 + R_L)\,I \gtreqless (1 + r_L)\,L$, such that for high loan amounts the loan cannot be repaid in full, where r_L denotes the loan rate.

The expected profits of the company are then easily given by

$$\Pi_C = \pi\left(\left(1+R_H\right)I-\left(1+r_L\right)L\right) + \left(1-\pi\right)\max\left\{\left(1+R_L\right)I-\left(1+r_L\right)L,0\right\}-E. \tag{8.23}$$

Using these profits, we can now derive the optimal loan demand by companies.

Optimal loan demand Using equation (8.23) and noting that $I = L+E$, we obtain the isoprofit curve for companies as $\frac{\partial \Pi_C}{\partial L}dL + \frac{\partial \Pi_C}{\partial r_L}dr_L = 0$, from which we then get the slope of the isoprofit curve as

$$\frac{dr_L}{dL} = \begin{cases} \frac{\pi(R_H-R_L)+(R_L-r_L)}{L} & \text{if } L \leq \frac{1+R_L}{1+r_L}I \\ \frac{R_H-r_L}{L} & \text{if } L > \frac{1+R_L}{1+r_L}I \end{cases}. \tag{8.24}$$

We assume that companies conduct their investments using loans only if their expected return $\pi R_H + (1-\pi)R_L$ exceeds the funding costs of r_L. This assumption ensures that the slope of this isoprofit curve is positive and we also observe that at $L = \frac{1+R_L}{1+r_L}I$ the slope of the isoprofit curves increases. It is at this point that the loan becomes risky to the bank as the company will not be able to repay its loan fully if the low return R_L is realised.

Banks finance the loan they provide fully through deposits on which they pay interest r_D. If return the company realises the high return R_H, the loan will be repaid with certainty and if the low return R_L is realised, the loan is only repaid in full if the bank has sufficient assets and otherwise obtains these assets. We thus have bank profits given by

$$\Pi_B = \pi\left(1+r_L\right)L + \left(1-\pi\right)\min\left\{\left(1+R_L\right)I,\left(1+r_L\right)L\right\} - \left(1+r_D\right)L. \tag{8.25}$$

Noting that $I = L+E$, we obtain the isoprofit curve for banks as $\frac{\partial \Pi_B}{\partial L}dL+\frac{\partial \Pi_B}{\partial r_L}dr_L = 0$, from which we then get the slope of the isoprofit curve as

$$\frac{dr_L}{dL} = \begin{cases} -\frac{r_L-r_D}{L} & \text{if } L \leq \frac{1+R_L}{1+r_L}I \\ -\frac{\pi(1+r_L)+(1-\pi)(1+R_L)-(1+r_D)}{\pi L} & \text{if } L > \frac{1+R_L}{1+r_L}I \end{cases}. \tag{8.26}$$

An equilibrium would emerge if the two slopes of the isoprofit curves were identical. In order to evaluate this equilibrium, we distinguish the cases of competitive banks and a monopolistic bank.

Competitive banks If banks are competitive, they will make no profits, thus we require $\Pi_B = 0$. From equation (8.25) this allows us to solve for the loan rate banks will apply, which becomes

$$1 + r_L = \begin{cases} 1 + r_D & \text{if } L \leq \frac{1+R_L}{1+r_L} I \\ \frac{(1+r_D)-(1-\pi)(1+R_L)\frac{L+E}{L}}{\pi} & \text{if } L > \frac{1+R_L}{1+r_L} I \end{cases}. \tag{8.27}$$

Using this result in the slope of the isoprofit curve of banks from equation (8.26), we easily see that for $L \leq \frac{1+R_L}{1+r_L} I$ we obtain $\frac{dr_L}{dL} = 0$. Setting this equal to the isoprofit curve of the company from equation (8.24) we require that $\pi (R_H - R_L) + (R_L - r_D) = 0$, having inserted that $r_L = r_D$. Such a parameter constellation is unlikely to be fulfilled and this would generally not be an equilibrium. Thus banks would lend a larger amount of $L > \frac{1+R_L}{1+r_L} I$.

In this case, we would obtain from setting the slopes of the indifference curves by banks and companies equal that $\pi (1 + r_L) + (1 - \pi) (1 + R_L) - (1 + r_D) = \pi (r_L - R_H)$ as we easily see from equations (8.24) and (8.26). Requiring that $\Pi_B = 0$ for competitive banks, we obtain after inserting $I = L + E$ that $\pi (1 + r_L) L + (1 - \pi) (1 + R_L) (L + E) - (1 + r_D) L = 0$. Using the left-hand side of the previous equation, this can be rewritten as $\pi (r_L - R_H) + (1 - \pi) (1 + r_L) E = 0$, implying that $\pi (R_H - r_L) = (1 - \pi) (1 + r_L) E > 0$. The first derivative of the company profits is given as $\frac{\partial \Pi_C}{\partial L} = \pi (R_H - r_L)$. We thus see that in equilibrium we would have $\frac{\partial \Pi_C}{\partial L} > 0$, implying that companies would prefer a larger loan than they obtain in equilibrium. We can interpret this excess demand by companies in equilibrium as credit rationing.

Monopolistic banks If banks are not competitive but are monopolistic, they would extract all surplus from companies such that $\Pi_C = 0$. From this we obtain that in the case of $L > \frac{1+R_L}{1+r_L} I$ it is $\pi (R_H - r_L) L = -(\pi (1 + R_H) - 1) E < 0$ after inserting for $I = L + E$. Given that $\frac{\partial \Pi_C}{\partial L} = \pi (R_H - r_L)$, we immediately see that $\frac{\partial \Pi_C}{\partial L} < 0$ and hence the equilibrium amount of loans exceeds the optimal loans that companies seek, thus credit rationing cannot occur.

In the case that $L \leq \frac{1+R_L}{1+r_L} I$ we have from $\Pi_C = 0$ that $\pi (R_H - r_L) = (1 - \pi) (r_L - R_L) - (\pi R_H + (1 - \pi) R_L) E$. As $L \leq \frac{1+R_L}{1+r_L} I$, we have $L \leq \frac{1+R_L}{r_L - R_L} E$ and hence the right-hand side will be less than $-((1 - \pi) (R_H - R_L) - \pi R_L) E < 0$, giving us again that $\pi (R_H - r_L) < 0$ and credit rationing cannot emerge in this case. Hence with monopolistic banks, credit rationing cannot occur.

Summary We have seen that in case of perfect competition between banks credit rationing can occur, while for monopolistic banks no such credit rationing can be observed. It is thus that competition makes the occurrence of credit rationing more likely as competition between banks reduces their profits and makes them vulnerable to losses from companies taking higher risk and subsequently not being able to repay their loan fully. Banks will subsequently limit their risk exposure by reducing the size of the loan they give, thus reducing the leverage companies can obtain and making their default less likely. If banking markets are less competitive and banks make higher profits, the risks to banks are mitigated through these higher profits banks can obtain from lending. This might lead to the observation that in competitive banking markets companies might struggle more to obtain a loan whose size meets their demand, while in less competitive markets their demands might be more easily met; however, they will pay higher loan rates in such less competitive markets.

Reading Meza & Webb (1987)

Conclusions

Credit rationing occurs if banks are not willing to provide a larger loan, even though the companies are willing to pay a higher loan rate than the loan rate offered with the smaller loan. The reason that banks are not willing to offer larger loans is that when doing so their profits are reducing. This reduction in profits is the result of companies being more likely to default when obtaining a larger loan, either through the higher repayment that is required, exacerbated by the higher loan rate, or through choosing more risky investments.

By reducing the loan amount, banks can provide incentives for companies to choose less risky investments as they repayments they have to make are reduced due to the lower loan amount as well as lower interest payments. This reduction in loan payments will allow banks to retain a larger fraction of their investment returns and induces them to pursue less risky projects. Similarly will the reduced loan payment make a default by the company less likely and hence the bank will a full repayment of their loan more often. Banks will not benefit by providing companies with larger loans at higher loan rates if these are less likely to be repaid than a smaller loan at lower loan rates.

Arising from the moral hazard of companies choosing more risky investments as well as the uncertainty surrounding the ability of the company to repay its loans, banks will not always meet the loan demands of companies. While they are willing to provide loans, the size of the loan might be smaller and offering to pay a higher loan rate will not induce banks to increase its loan size. The company will feel rationed in the loan amount they can obtain. Similarly, a larger loan can also provide incentives to strategically default, given the large benefits arising from such a decision. Banks will then provide only a smaller loan to reduce the incentives to companies for such strategic defaults.

Competition between banks makes credit rationing more likely to occur. The lower profits banks make in competitive markets, will make it difficult for banks to compensate for the risks banks face when making larger loans. As increasing the loan rate increases the risk of default as the amount that is to be repaid to the bank increases, such an increase will not necessarily increase the profits of banks, it may actually reduce them as incentives for companies become such that more risky investments are pursued or strategic default becomes more attractive. Hence banks will reduce their risk exposure by reducing the size of the loan, which has the additional benefit of affecting the incentives of companies positively.

We thus see that on many occasions banks will not provide companies with loans of the size they find optimal. Taking into account the likelihood of a loan being repaid, in addition to the amount that is being repaid, can lead bank to the an assessment that only a smaller loan should be provided. Using smaller loan allows banks to affect the incentives of companies to pursue more risky investments or default strategically and is therefore used as an incentive device by banks to reduce the risks they are exposed to through the company's decisions.

Chapter 9

Collateral

A collateral is an asset that the bank can obtain if a company cannot repay its loan. Until such a default occurs, the asset remains the property of the company providing this collateral, and if no default occurs it is never transferred to the bank. Collateral can take many forms, best known is the use of real estate for mortgages, a name for loans that use real estate as collateral. Many other assets can be used as collateral, such as retaining an interest in a car if this is financed by a loan, any machinery a company might hold, securities held in a portfolio, or account balances at the same or other banks. Other assets might include future payments the company receives (receivables) from already agreed or yet to agree sales. A common feature of these forms of collateral is that in most cases these assets are owned by the company and in principle the bank would have access to these assets if the company defaults and it is liquidated. However, the value of such assets to banks is small as firstly the realisation of their value in the liquidation process takes considerable time. Secondly, the value of these assets need to be shared with any other creditors, making it often difficult to obtain a substantial payout, especially after the costs of the liquidation have been taken into account. Having a collateral has the effect, if done following due legal process, that the assets earmarked as collateral are taken out of the liquidation procedure and given to the bank directly. This process is not only faster than following the normal liquidation procedure, the bank can also be assured that they receive the full value of the assets that have been pledges as collateral as they do not have to be shared with other creditors.

While the benefits to banks of having such collateral is obvious, the costs to companies of providing is less obvious. As defaulting companies will in principle be liquidated, it should make no difference whether assets are liquidated in the normal process or transferred to the bank at the earliest point in time; in either case the assets are lost to the company. Indeed models using collateral assume that companies defaulting face additional costs through losing their collateral. Firstly these costs may actually arise if the collateral was owned by another legal entity within the company, for example a subsidiary or parent company. Such assets might not or only with difficulty be seized by any liquidators. Therefore their seizure would impose an actual loss to the company. But even if liquidators had access to the assets, there might be an additional loss to the company. In many cases companies are not fully liquidated, but an arrangement is made between the liquidators (or persons with comparable roles) and any creditors to ensure the company can survive. Often a write-down of loans is agreed, a conversion of loans into equity, or a postponement of loan repayments until the company is profitable again after a restructuring. Having been pledged a collateral, the bank does not have to participate in this process and could insist on the collateral being handed over to them, unless they want to provide support in allowing the company to continue operating. If they insist on obtaining their collateral, this might affect negatively the company's chances of survival and would thus be a cost to the company.

There are other forms of collateral that would impose actual losses on the company or associated entities. Most common among such form of collateral is the guarantee, often given as a personal guarantee by the owner of the company, backed up by his private wealth, that would otherwise not be part of the liquidation process. This guarantee might also be in the form of a mortgage on a private property of the owner, or any other person agreeing to such an arrangement. Guarantees might also be given by other companies, either legally independent of the company seeking the loan but controlled by the same owner, or it might be a parent company guaranteeing the loan of a subsidiary. In all cases a default of the company would impose actual losses on those providing the guarantee. If we assume that these costs are internalised and thus taken into account in the decision-making of the company, guarantees can be treated as collateral.

This chapter will investigate the implications the use of collateral has on the decision-making of companies and banks alike. Chapter 9.1 will discuss the benefits arising from the use of collateral in terms of the cost of loans to companies, before in chapter 9.2 wider implications on the incentives of companies are considered, in terms of the ability of collateral to reduce adverse selection between companies and banks as well as a reduction in the moral hazard of companies when providing effort to reduce the risks of their investments. Once collateral has been provided, banks do not only retain them as an insurance in the case a loan is not repaid, but can use them for their own benefit as we will see in chapter 9.3.

Not strictly a collateral, but another form to companies do not adversely affect their ability to repay the loan once it has been granted, is a debt covenant. Covenant impose constraints on the behaviour of companies with the aim to ensure that the risks to the bank are not increasing during the life time of a loan. In chapter 9.4 we discuss the implications of such debt covenants and chapter 9.5 assess the desire by banks for loan guarantees.

9.1 The benefits of collateral

The widespread use of collateral suggests that there are inherent benefits to its use. Chapter 9.1.1 will show that the use of collateral will reduce the loan rate banks charge companies, but that companies are neither better or worse off when agreeing to provide such collateral, although there might be secondary benefits or costs.

Similarly banks are not in itself better or worse off when obtain a collateral, but may have indirect benefits allowing them to expand lending. However, the use of collateral can be beneficial to the company providing the collateral if the bank and the company disagree on the risks of the investment that is financed. As chapter 9.1.2 shows, collateral can be used to transfer risk from banks, who perceive them to be high, to companies, who perceive them to be low.

9.1.1 Risk reduction through collateral

If companies provide banks with a collateral, the bank can use this collateral to reduce their losses in the case that the company is not able to repay the loan. This will reduce the risks banks are taking when providing loans, which should be reflected in the loan rate they are charging. Assume companies are making an investment that returns V if it is successful, which happens with probability π; in this case the company can repay the loan L, including interest r_L. If the investment is not successful, it yields no return and the company is unable to repay the loan, but it will loose its collateral C. This gives the company a profit of

$$\Pi_C = \pi\left(V - (1 + r_L)\,L\right) - (1 - \pi)\,C. \tag{9.1}$$

The bank will obtain the loan repayment if the investment is successful and if it is not successful it will obtain the collateral. Having funded their loan fully through deposits on which interest r_D is payable, the bank profits are given by

$$\Pi_B = \pi\,(1 + r_L)\,L + (1 - \pi)\,C - (1 + r_D)\,L. \tag{9.2}$$

In a competitive markets banks make no profits, $\Pi_B = 0$ and the loan rate will be given by

$$1 + r_L = \frac{1 + r_D}{\pi} - \frac{1 - \pi}{\pi}\frac{C}{L}. \tag{9.3}$$

We see that the more collateral is required, the lower the loan rate will be; this reflects the lower risk the bank is exposed to. The company, on the other hand, faces a higher risk as they might lose their collateral if the investment fails. This risk, however, is compensated fully by the lower loan rate the bank charges. Inserting equation (9.3) into equation (9.1), the company profits are

$$\Pi_C = \pi V - (1 + r_D)\,L, \tag{9.4}$$

which does not depend on amount of collateral the company had to provide. Thus the company should be indifferent whether it provides the bank with a collateral or not.

Providing collateral has the advantage that the interest to be paid on the loan is reduced, preserving the cash position of companies, and - as long as the investment is not failing - increasing the profits they can show to their investors. This is particularly attractive if the collateral is an asset that either cannot be used otherwise productively or can still generate the same return, even if pledged as a collateral. On the other hand, if the investment fails, the company will lose the collateral; at a time of failing investments facing the loss of potentially important assets, might be more detrimental to the company than paying higher interest on their loan.

For banks the main benefit is that the potential losses they face are significantly reduced by being provided with a collateral. This may lead to lower capital costs due to taking lower risks, but also a lower capital requirement on this loan, allowing the bank to increase lending. There is not an immediate impact on the profitability of banks; while the interest they earn is reduced, the potential losses are reduced and hence any loan write-offs will also be smaller. Hence the main attraction of collateral to banks is the reduced risk.

Reading Jappelli, Pagano, & Bianco (2005)

9.1.2 Collateral overcoming different risk assessments

In many cases the assessment of the prospect of investments differ between the company and the bank. It is common that the company assesses the risks associated with an investment as smaller than their bank. The reason for this difference in the risk assessment might be found in the lack of credible information the bank has on an investment, but it might also reflect an over-optimistic assessment of the company. Whatever the origins of this discrepancy between the assessment companies and their bank, it will have implications for the loan rate the company is charged by its bank, making the loan more expensive than the company would expect given its own analysis.

Let us assume that companies assess the likelihood that the investment is successful and yields an outcome of V, as being π_C, while banks assign this likelihood a value of $\pi_B < \pi_C$. Assume now that bank were to share the assessment of the company, hence its profits would be

$$\Pi_B = \pi_C (1 + r_L) L - (1 + r_D) L, \tag{9.5}$$

where the bank provides a loan of size L at loan rate r_L, fully financed by deposits on which they have to pay a deposit rate of r_D. If banks are competitive, we would have $\Pi_B = 0$ and hence

$$1 + r_L = \frac{1 + r_D}{\pi_C}. \tag{9.6}$$

This is the loan rate a company would expect to receive. Its profits are then given by

$$\Pi_C = \pi_C (V - (1 + r_L) L) = \pi_C V - (1 + r_D) L. \tag{9.7}$$

The company obtains the investment outcome V and repays the loan, if the investment is successful; the second equality arises from inserting for $1 + r_L$ from equation (9.6).

However, the bank disagrees with the risk assessment of the company and would actually charge a loan rate of $1 + r_L = \frac{1+r_D}{\pi_B}$, as can easily be verified, which is higher due to our assumption of $\pi_B < \pi_C$, and hence the company would make less profits. Suppose now that the bank offers the company a loan contract with collateral requirements C. In this case the bank would charge a loan rate $\hat{r}_L$ and obtain the collateral if the company fails to repay its loan. Thus the bank profits are given by

$$\hat{\Pi}_B = \pi_B (1 + \hat{r}_L) L + (1 - \pi_B) C - (1 + r_D) L. \tag{9.8}$$

With banks being competitive, the requirement that $\hat{\Pi}_B = 0$ gives us a loan rate of

$$1 + \hat{r}_L = \frac{1 + r_D}{\pi_B} - \frac{1 - \pi_B}{\pi_B} \frac{C}{L}. \tag{9.9}$$

The company profits are now reduced by the loss of the collateral if the loan is not repaid, hence

$$\begin{aligned} \hat{\Pi}_C &= \pi_C (V - (1 + \hat{r}_L) L) - (1 - \pi_C) C \\ &= \pi_C V - \frac{\pi_C}{\pi_B} (1 + r_D) L + \frac{\pi_C - \pi_B}{\pi_B} C, \end{aligned} \tag{9.10}$$

where the second equality arises from inserting equation (9.6). For a company to be as well off if the bank disagrees with the company on the risks of their investment compared as to when they would agreed on the company assessment, we would want to set the collateral such that $\hat{\Pi}_C = \Pi_C$. This solves for

$$C = (1 + r_D) D. \tag{9.11}$$

Hence, if the company provides collateral to the extent that the bank can repay its depositors, the loan rate is reduced sufficiently to increase the profits of the company to the level it would be if the bank shared their risk assessment, despite the possible loss of the collateral.

We can use collateral to overcome the losses a company might face from higher loan rates if the bank do not agree on their assessment that the risks associated with an investment is low, thus the likelihood of success being high. By using collateral, the risk to the bank, perceived by them to be high, reduces, allowing it to reduce the loan rate to its funding costs, r_D. In turn, the collateral exposes the company to the risks of their investment, which they perceive to be low. Hence the high risk from the bank's perspective has been exchanged for a low risk from the company's perspective.

Reading Chan & Kanatas (1985)

Résumé

As long as the company and its bank agree on the risks associated with the investment that is financed, there is no direct economic benefit or cost to the use of collateral. The expected profits, taking into account that banks will obtain the collateral from the company if the company cannot repay its loan. The reduced loan rate that a loan using collateral demands will be offset exactly by the possible transfer of the collateral from the company to the bank. There may well be indirect benefits arising from the use of collateral; these may include lower capital requirements for banks due to lower risks and lower interest payments during the life time of the loan for the company. Direct benefits will only be observed if the company and its bank disagree on the risks the company takes. In this case the collateral allows banks to reduce the loan rate and the transfer of the risk to the company, who perceives this risk as being lower, allows the company to make the same profits as if the bank would agree on its risk assessment.

9.2 Collateral as an incentive device

Banks are often in a position where they are less well informed about the risks of investments than the companies they are lending to. This naturally arises from the familiarity of companies with their business and the difficulty of banks in assessing the information they have been able to obtain, consequently they are often not able to distinguish the risks companies face. This asymmetric information can lead to adverse selection where low-risk companies are priced out of the loan market and the bank is faced only with high-risk companies seeking loans. Offering loan contracts that include the possibility of providing collateral, banks can be able to distinguish between companies facing different levels of risk as we will see in chapter 9.2.1. Adverse selection is not the only problem banks face when providing loans to companies. It might not be in the best interest of companies to exert a high level of costly effort to reduce the risk of their investment. Such moral hazard can lead to suboptimal allocation of resources and we will see in chapter 9.2.2 how collateral can be used to align the interest of companies the social optimum.

9.2.1 Identifying company types through collateral

Banks cannot always distinguish clearly the likelihood a company is repaying the loan, while the company itself might have better knowledge about their own ability. A bank setting loan rates that account for the average repayment rate of companies would face adverse selection in that such loan rates are only attractive to companies with low abilities to repay, while companies with high abilities to repay will not seek a loan. Banks will therefore grant loans only to companies with low abilities to repay loans, facing a loss of doing so due to the low repayments they will receive. By offering a loan contract that requires collateral, the bank will be able to distinguish between companies of different abilities to repay their loans.

Let us assume that companies succeed with their investment with probability of π_i giving a return of R, and otherwise they fail; if they have provided collateral, they will lose it to the bank providing the loan. If the investment is fully financed by a loan L, on which the bank charges interest r_L^i, and the company provides collateral C_i, we get the company profits as

$$\Pi_C^i = \pi_i \left((1+R)\,L - \left(1+r_L^i\right)L\right) - (1-\pi_i)\,C_i. \tag{9.12}$$

In order to assess the trade-off between the loan rate, r_L^i, and the amount of collateral provided, C_i, we assume that we hold the company profits constant and taking the total differential gives us $d\Pi_C^i = -\pi_i L dr_L^i - (1-\pi_i)\, dC_i = 0$ and hence

$$\frac{dr_L^i}{dC_i} = -\frac{1-\pi_i}{\pi_i L}. \tag{9.13}$$

We thus see that companies providing a higher collateral would need a lower loan rate to retain the same profits.

Let us assume that the bank finances its loans fully through deposits on which they ay interest r_D. Any collateral they are provided with they obtain if the company fails and cannot repay its loan. However, banks can only sell the collateral at a loss as they are obtaining an asset which will be sold into a market they are not familiar with; we thus assume that banks only obtain a fraction $\lambda \leq 1$ of the value of the collateral. The bank profits are now given as

$$\Pi_B^i = \pi_i \left(1 + r_L^i\right) L + (1-\pi_i)\, \lambda C_i - (1 + r_D)\, L. \tag{9.14}$$

We propose that banks are competitive such that $\Pi_B^i = 0$ and totally differentiating their profits yields $d\Pi_B^i = \pi_i L dr_L^i + \lambda\,(1-\pi_i)\, dC_i = 0$, from which we obtain that

$$\frac{dr_L^i}{dC_i} = -\lambda \frac{1-\pi_i}{\pi_i L} \tag{9.15}$$

If the bank is provided with a larger collateral, it will charge a lower loan rate to maintain its competitive profits of zero. The relationship between the loan rate and the collateral is less strong for banks compared to companies due to the factor λ, which accounts for the losses the bank would make when selling the collateral.

For simplicity let us assume that there only two types of companies, one type makes low-risk investments, which has a probability of success π_H and the other type of companies makes high-risk investments which succeed with probability $\pi_L < \pi_H$. We immediately see from equations (9.13) and (9.15) that the relationship between the loan rate and collateral is stronger, thus has a lower value, for the high-risk company having a probability of success pi_L.

We illustrate in figure 9.1 the iso-profit curves of a bank lending to the high-risk companies, depicted in black, and the low-risk companies, shown in green. These isoprofit curves assume that banks know the type of company they are providing a loan to; even though they are unaware of this property, we will see from the argument that follows, that they can make a correct inference about the companies. The area blow the isoprofit curve of banks, thus charging a lower loan rate or requiring lower collateral, will induce losses to the bank, while the area above the isoprofit curve generates profits to the bank. Note that if the value of the collateral to the bank, λC_i exceeds the amount to be repaid, $\left(1 + r_L^i\right) L$, the bank always obtains full repayment, regardless of the success of the investment of the company. Providing a larger collateral would be not beneficial and with banks supposed to be competitive making no profits, the loan rate will reflect the costs of funding by the banks, its deposit rate. No loan will be offered below this interest rate.

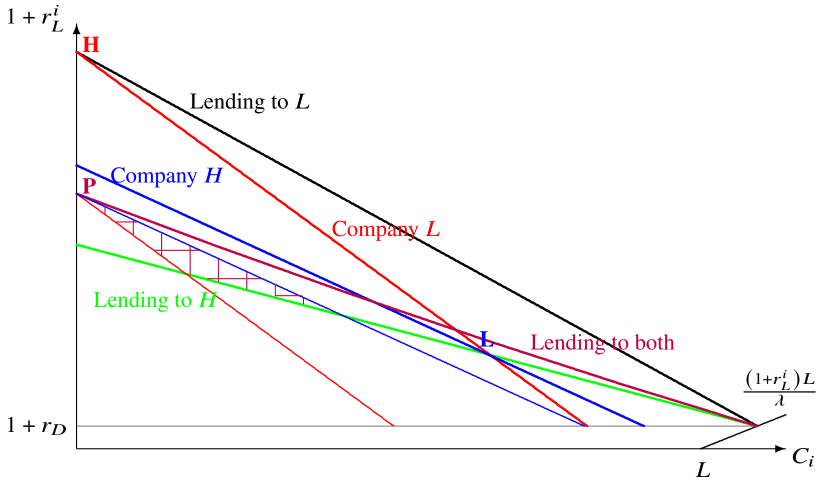


Fig. 9.1: Separating equilibrium with collateral

The isoprofits curves of companies making low-risk investments, allowing loans to be repaid with probability π_H, are indicated in red and those of companies making high-risk investments, corresponding to a success rate of π_L, in blue. As indicated above, the slopes of these isoprofit curves are steeper than those of the banks. Companies prefer lower loan rates and providing less collateral, thus profits are increasing the lower or more left the isoprofit curve is located.

With perfect competition between banks requiring any loan conditions to be located on the isoprofit curve of banks, we see that for high-risk companies, the best solution, providing the highest profits to companies, is to not seek collateral and charge a high loan rate. This solution is indicated as H in figure 9.1. Similarly, banks lending to low-risk companies would break even and companies enjoyed the highest possible profits, if they also did not require collateral and charged a loan rate indicated by the point where the green line crosses the vertical axis. However, banks cannot distinguish between companies of different types, hence the loan rate and amount of collateral required can be accepted by either the high-risk or the low-risk company. Clearly, if lending at this loan rate to a high-risk company, the bank would make a loss, thus such a solution is not feasible, given that it does not make any profits when lending to low-risk companies. The bank seeking to lend only to low-risk companies would have to offer loan conditions that are worse than H for high risk companies, but better than H for low-risk companies. At point L the isoprofit curve of high-risk companies crosses the isoprofit curve of banks lending to low-risk companies. If a bank would charge the loan rate and require collateral for this point, or a point just marginally to the right, this loan contract would not be selected by high-risk companies, who prefer H, but clearly it is preferred by low-risk companies as it provides them with a higher profits.

It is therefore that banks may offer two loan contracts, H and L. The high-risk company will seek loan contract H where it pays a high loan rate, but does not provide collateral, while the low-risk company will select loan contract L, enjoying a lower loan rate but having to provide collateral. High-risk companies will prefer to not provide collateral as the higher risks they are exposed to increases the probability of them losing their collateral; this makes the use of collateral unattractive and companies instead prefer to pay a higher loan rate. We have thus achieve a separation of companies and from the choice of loan contract, banks know the type of company they are lending to. This is commonly referred to as a separating equilibrium.

We can now consider a bank which only offers a single loan contract to both types of companies. If the bank knows that there is a fraction p of low-risk companies and a fraction $1 - p$ of high-risk companies in the market, the expected success rate is given by $\pi = p\pi_H + (1 - p)\,\pi_L$ and hence bank profits are

$$\Pi_B^P = \pi\,(1 + r_L)\,L + (1 - \pi)\,\lambda C - (1 + r_D)\,L, \tag{9.16}$$

giving us isoprofit curves with slope

$$\frac{dr_L}{dC} = -\lambda\frac{1 - \pi}{\pi L}. \tag{9.17}$$

As $\pi_H \geq \pi \geq \pi_L$, this slope will be between the slope of the isoprofit curves of bank lending to low-risk and high-risk companies, respectively. The resulting optimal loan contract is indicated in figure 9.1 by P, and the isoprofit curves of the high-risk and low-risk companies for this loan contract are indicated as well. Such a loan contract is often called a pooling equilibrium. We see that this pooling equilibrium is preferred by both high-risk and low-risk companies as their respective isoprofit curves are below those of the separating equilibrium. Hence offering this loan contract is feasible.

It could now be that a bank seeks to deviate from providing a single loan contract to both types of companies. By offering a loan contract in the hatched area, the bank would only attract low-risk companies. For them a loan contract in this area represents an increase in profits as the loan contract is below their iso-profit curve, while for high-risk companies this loan contract is above their current isoprofit curve, making it less attractive. As the loan contract is also above the isoprofit curve of a bank lending to low-risk companies only, the bank would make a profit and offering such a loan contract is viable. This would, however, leave the bank offering a single loan contract only with only high-risk borrowers, and would thus induce a loss to them.

A bank seeking to offer such a loan contract to low-risk companies, would do so by choosing a loan contract at point P, or marginally to the lower right of this point. The loan rate at this point is chosen such that the bank offering loan contracts to both types of companies breaks even, $\Pi_B^P = 0$ and knowing that no collateral is required, $C = 0$, we get from equation (9.16) that $1 + r_L = \frac{1+r_D}{\pi}$. The bank offering the alternative loan contract to low-risk companies would then make profits of

$$\Pi_B^* = \pi_H (1 + r_L) L - (1 + r_D) L = (1 - p) \frac{\pi_H - \pi_L}{\pi} (1 + r_D) L. \tag{9.18}$$

The company offering the loan contract to both types of companies will only be providing loans to high-risk companies as any low-risk companies will seek a loan from the other bank. Its profits are therefore given by

$$\Pi_B^{**} = \pi_L (1 + r_L) L - (1 + r_D) L = -p \frac{\pi_H - \pi_L}{\pi} (1 + r_D) L. \tag{9.19}$$

If both banks offer the pooling equilibrium or the separating equilibrium, they will be making zero profits due to our assumption of perfect competition between banks.

		Bank 1	
		pooling	separating
Bank 2	pooling	0, 0	Π_B^*, Π_B^{**}
	separating	Π_B^{**}, Π_B^*	0, 0

Fig. 9.2: Strategic choice of loan contracts

Banks now need to decide whether to offer a single loan contract to both types of companies, the pooling equilibrium, or offer different types of loan contracts to companies. We can interpret this as a strategic interaction between two banks, whose resulting profits are depicted in figure 9.2, where pooling indicated the offer of loan contract P and separating the offer of loan contracts H and L. It is easy to confirm that the only equilibrium is that both banks choose the separating equilibrium. Hence, while a pooling contract may be desirable for companies, it is not an equilibrium and we will observe a separating equilibrium.

We have thus established that by offering two types of loan contracts, one with a high loan rate without collateral requirements and another contract with a lower loan rate with collateral requirements, banks can distinguish between companies taking different levels of risk. Hence, collateral can be used extract information from companies and reduce adverse selection between banks and companies. It is low-risk companies that are willing to provide collateral, while high-risk companies prefer to pay higher loan rates instead of providing collateral. Of course, such a separating equilibrium can only emerge if low-risk companies are able to provide the collateral required, and thus collateral might now always be able to distinguish between companies of different risks. Furthermore, while we here assumed that companies know their own risks, companies might assess their own risks wrongly and hence the choice of collateral would only reflect the beliefs of the company rather than the actual risks it faces.

Reading Bester (1985)

9.2.2 Collateral and moral hazard

The success of investments companies make, will not only depend on the ability of the company, but also the effort they exert. However, exerting effort will impose costs on the company and while it might be desirable that such effort is exerted, the private incentives of companies might be such that the costs they have to bear make this not the best available option. This leads to moral hazard in that the exertion of effort is socially desirable, but not profitable to the company having to bear its costs.

Let us assume that the quality of the company can be either good, indexed by G or bad, B, and this quality is known to the company itself as well as the bank providing the loan for their investments. The probability of the investment being successful and generating value V will be higher for companies with a high ability. In addition, companies can exert effort to increase the probability of success, this effort incurs costs E to the company. We assume that the increase in the success rate is more pronounced for the bad company than the good company. Hence we find that

$$\hat{\pi}_B - \pi_B \geq \hat{\pi}_G - \pi_G, \tag{9.20}$$

where π_B (π_G) denotes the probability of success of the bad (good) company if effort is exerted, while $\hat{\pi}_B$ ($\hat{\pi}_G$) denotes the probability of success of the bad (good) company if no effort is exerted.

Banks provide a loan of size L, which fully finances the investment of the company and finances this loan fully through deposits on which interest r_D is payable.

Social optimum In the social optimum the total welfare is composed of the investment outcome, provided the investment is successful, less the costs of funding the loan. In the cases the good company is conducting the investment, the social welfare with and without the exertion of effort is given by

$$\begin{aligned} \Pi_W^G &= \pi_G V - (1 + r_D) L, \\ \hat{\Pi}_W^G &= \hat{\pi}_G V - (1 + r_D) L - E. \end{aligned} \tag{9.21}$$

It is optimal for the good company to not exert effort if $\Pi_W^G \geq \hat{\Pi}_W^G$, which easily solves for

$$\hat{\pi}_G - \pi_G \leq \frac{E}{V} \tag{9.22}$$

Similarly, for bad companies we have the welfare given by

$$\begin{aligned} \Pi_W^B &= \pi_B V - (1 + r_D) L, \\ \hat{\Pi}_W^B &= \hat{\pi}_B V - (1 + r_D) L - E \end{aligned} \tag{9.23}$$

and the bad company would exert effort if $\Pi_W^B \leq \hat{\Pi}_W^B$, from which we obtain

$$\hat{\pi}_B - \pi_B \geq \frac{E}{V}. \tag{9.24}$$

Combining equations (9.22) and (9.24) into $\hat{\pi}_B - \pi_B \geq \frac{E}{V} \geq \hat{\pi}_G - \pi_G$, we see that such an allocation of effort is consistent with our assumption in equation (9.20). Let us therefore now assume that effort costs are such that this condition is fulfilled. We then have the social optimum as bad companies exerting effort and good companies exerting no effort. While this result represents the social optimum, its implementation will depend on the incentives of each company.

No collateral The good company will obtain the investment outcome V and repays its loan, including interest r_L with probability $\hat{\pi}_G$ if it exerts effort and with probability π_G if it does not exert effort. Its profits are thus given by

$$\begin{aligned} \hat{\Pi}_C^G &= \hat{\pi}_G \left(V - (1 + r_L) L\right) - E, \\ \Pi_C^G &= \pi_G \left(V - (1 + r_L) L\right). \end{aligned} \tag{9.25}$$

If $\hat{\Pi}_C^G \leq \Pi_C^G$, the company will choose to not exert effort. This requires

$$\hat{\pi}_G - \pi_G \leq \frac{E}{V - (1 + r_L) L}. \tag{9.26}$$

Compared to the condition in the social optimum in equation (9.22), the increase in the success rates from exerting effort may be substantially lower. This implies that the exertion of effort by good companies may occur even though it is not socially optimal, they exert too much effort.

The bad company will obtain the investment outcome V and repays its loan, including interest, with probability $\hat{\pi}_B$ if it exerts effort and with probability π_B if it does not exert effort. Its profits are thus given by

$$\begin{aligned} \hat{\Pi}_C^B &= \hat{\pi}_B \left(V - (1 + r_L) L\right) - E, \\ \Pi_C^B &= \pi_B \left(V - (1 + r_L) L\right). \end{aligned} \tag{9.27}$$

If $\hat{\Pi}_C^B \geq \Pi_C^B$, the company will choose to exert effort. This requires

$$\hat{\pi}_B - \pi_B \geq \frac{E}{V - (1 + r_L) L}. \tag{9.28}$$

Compared to the condition in the social optimum in equation (9.24), the increase in the success rates from exerting effort must be substantially higher. This implies that the exertion of effort by bad companies is not always guaranteed where it would be socially optimal, they exert too little effort.

Hence we find that good companies exert too much effort, while bad companies exert too little effort, compared to the social optimum.

Using collateral Banks might demand a collateral when providing loans, which is lost to the company if the investment is not successful and the loan cannot be repaid. Of course, the bank might charge a different loan rate $\hat{r}_L$ compared to a loan without such collateral. In general the loan rate will be lower if collateral is provided.

As before, the good company will obtain the investment outcome V and repays its loan, including interest, with probability $\hat{\pi}_G$ if it exerts effort and with probability π_G if it does not exert effort. Given the use of collateral C, the company will lose this collateral if the investment is not successful. Its profits are thus given by

$$\begin{aligned} \hat{\Pi}_C^G &= \hat{\pi}_G \left(V - (1 + \hat{r}_L) L\right) - (1 - \hat{\pi}_G) C - E, \\ \Pi_C^G &= \pi_G \left(V - (1 + \hat{r}_L) L\right) - (1 - \pi_G) C. \end{aligned} \tag{9.29}$$

If $\hat{\Pi}_C^G \leq \Pi_C^G$, the company will choose to not exert effort. This requires

$$\hat{\pi}_G - \pi_G \leq \frac{E}{V - (1 + \hat{r}_L) L + C}. \tag{9.30}$$

Compared to the condition in the social optimum in equation (9.22), this constraint is identical if $C = (1 + \hat{r}_L) L$. Hence, if the loan is fully collateralized, the bad company exerts effort consistent with the social optimum.

Similarly, the bad company will obtain the investment outcome V and repays its loan, including interest, with probability $\hat{\pi}_B$ if the company exerts effort and with probability π_B if it does not exert effort. Given the use of collateral C, it will lose this collateral if the investment is not successful. Its profits are thus given by

$$\begin{aligned} \hat{\Pi}_C^B &= \hat{\pi}_B \left(V - (1 + \hat{r}_L) L\right) - (1 - \hat{\pi}_B) C - E, \\ \Pi_C^B &= \pi_B \left(V - (1 + \hat{r}_L) L\right) - (1 - \pi_B) C. \end{aligned} \tag{9.31}$$

If $\hat{\Pi}_C^B \geq \Pi_C^B$, the company will choose to exert effort. This requires

$$\hat{\pi}_B - \pi_B \geq \frac{E}{V - (1 + \hat{r}_L) L + C}. \tag{9.32}$$

Compared to the condition in the social optimum in equation (9.24), this constraint is identical if $C = (1 + \hat{r}_L) L$. Hence, if the loan is fully collateralized, the bad company exerts effort consistent with the social optimum.

Thus by fully collateralising the loan, the social optimum in the exertion of effort can be implemented. The reason the company will exert effort in a socially optimal way is that due to the limited liability of the company, the it would ignore the losses imposed on banks from it defaulting on its loan if no collateral is provided. As the company is loosing the collateral if it defaults, this loss is internalised and the social optimum obtained.

Summary Collateral can be used to overcome the moral hazard of companies not exerting sufficient or too much effort. Through the loss of the collateral when defaulting on their loans, companies internalise the costs of their default and as such their behaviour will align with that of the social optimum. This social optimum is only achieved if the loan is fully collateralised, a partial collateralisation of the loan will result in a closer, but still imperfect alignment of the incentives on exerting efforts with the social optimum.

While banks in general are not concerned about achieving the social optimum, but rather seek to reduce the risks arising from providing loans, they would find it particularly useful to require high-risk companies to provide collateral in order to provide incentives to exert more effort and reduce the risks of the loan to the bank. On the other hand, the incentives to reduce efforts by low-risk companies towards the social optimum, would not necessarily in the interest of banks and hence they might not ask for collateral from such companies.

Reading Boot, Thakor, & Udell (1991)

Résumé

Collateral can be useful in allowing banks to distinguish between companies having different levels of risk, enabling banks to tailor the loan conditions to their specific risk profile, and it can provide incentives for companies to exert optimal levels of effort to reduce the risks of their investments. Collateral has this effect as a high risk will increase the likelihood of losing the collateral, given higher probability of not being able to repay the loan. This loss can be reduced if the risk is reduced, decreasing the moral hazard problem, but also inducing companies that cannot reduce their risks to not offer collateral at all, allowing banks to distinguish between companies of high and low risk, lessening the problem of adverse selection.

9.3 Rehypothecation

Companies pledge collateral to their bank and the bank will take control of this collateral if the loan cannot be repaid; this is done to reduce the losses to the bank. If the bank itself would require a loan, for example to provide more loans to other companies, they could be required to provide collateral themselves in order to obtain this loan. If we assume that the bank has no collateral itself, it could use the collateral that is provided to the bank by the company. It is thus that the bank will use the collateral they have received and pledge the same collateral to a lender of theirs. Such a process is referred to as rehypothecation. Of course, the company providing the collateral originally has to agree to this arrangement.

We will evaluate how the ability of banks to rehypothecate the collateral affects the company providing the collateral and whether they would agree to such an arrangement, as well as whether rehypothecation is desirable for the bank. To fully assess the impact the use of collateral has on companies, we initially assess a situation in which no collateral is offered, then introduce the use of collateral, before extending the framework to include rehypothecation.

Borrowing without rehypothecation A company has an investment available that would need to be fully financed by a bank loan L and which yields a return of R, if successful; if the company is not successful it receives no payment. The investment has a probability of success π_L if the company does not exert any additional effort and a probability of success of $\pi_H > \pi_L$ if the company exerts additional effort, which costs them E. These costs might comprise the building of expertise, managerial capacities, additional staffing, or longer and more intense working hours.

With an interest rate r_L on the loan, the expected profits of the company are given by

$$\begin{aligned} \Pi_C^H &= \pi_H \left((1+R) L - (1+r_L) L \right) - E = \pi_H (R - r_L) L - E, \\ \Pi_C^L &= \pi_L \left((1+R) L - (1+r_L) L \right) = \pi_L (R - r_L) L, \end{aligned} \tag{9.33}$$

for exerting effort and not exerting effort, respectively. The company will choose to exert effort if this is more profitable, thus $\Pi_C^H \geq \Pi_C^L$, which solves for

$$L \geq L^* = \frac{E}{(\pi_H - \pi_L)(R - r_L)}. \tag{9.34}$$

Hence, as long as the loan is large enough to spread the costs of effort sufficiently, companies will exert effort. Equivalently, we could state that as long as the effort costs are not too high, effort will be exerted.

Banks finance their loan entirely with deposits on which interest r_D is payable and we assume that banks want to induce companies to exert efforts as lending to companies not exerting this effort is not profitable, even in the presence of collateral, which we introduce below. Thus, in light of the constraint in equation (9.34), banks would provide only loans to companies seeking a sufficiently large loan of at least L^* to ensure the incentives to the company induce it to exert effort.

If, however, companies provide collateral C, their respective profits are reduced by $(1 - \pi_i)C$ as they would lose this collateral in case their investment fails. Hence their profits when exerting effort and not exerting effort, respectively, are given by

$$\begin{aligned} \hat{\Pi}_C^H &= \pi_H \left((1+R) L - (1+r_L) L \right) - (1 - \pi_H) C - E \\ &= \pi_H (R - r_L) L - (1 - \pi_H) C - E, \\ \hat{\Pi}_C^L &= \pi_L \left((1+R) L - (1+r_L) L \right) - (1 - \pi_L) C \\ &= \pi_L (R - r_L) L - (1 - \pi_L) C. \end{aligned} \tag{9.35}$$

The company will choose to exert effort if it is more profitable to do so, thus $\hat{\Pi}_C^H \geq \hat{\Pi}_C^L$. This gives us

$$L \geq L^{**} = \frac{E - (\pi_H - \pi_L) C}{(\pi_H - \pi_L)(R - r_L)}. \tag{9.36}$$

In comparison with the constraint in the absence of collateral from equation (9.34), it is obvious that this requirement on the minimum loan size is less stringent than without the provision of collateral, making the provision of loans possible for smaller loans than without the provision of collateral. The loss of the collateral in the case that the investment is not successful, provides stronger incentives for the company to exert effort and reduce this possibility. This in turn allows companies to obtain smaller loans compared to the case where no collateral was required. Similarly to before, we can interpret this result as companies that face higher effort costs are able to obtain a loan of the same size compared to a situation in which no collateral is provided.

It is trivial to show that for identical loan rates, banks would always prefer the company to provide collateral; retaining the collateral if the investment of the company fails, increases the profits of the bank without reducing its revenue from the loan repayment. Hence banks will always ask for collateral. We will now compares these results with that where banks can rehypothecate the collateral they have obtained from the company.

Allowing rehypothecation Let us now assume that the collateral pledged by the company can be used by the bank to gain access to a loan, similarly as the original company had to provide collateral to the bank in order to obtain its loan. The collateral would ensure the lender obtains sufficient repayments from the bank, either by them repaying their loan or forfeiting the collateral the provided.

The probability of the success of the bank's investment is denoted $\hat{\pi}$, its return if successful $\hat{R}$, and the loan obtained has size $\hat{L}$ at an interest rate $\hat{r}_L$. The bank taking this additional loan would make profits

$$\begin{aligned}\hat{\Pi}_B = {} & \hat{\pi}\left(\left(1 + \hat{R}\right)\hat{L} - (1 + \hat{r}_L)\hat{L} + \pi_H (1 + r_L) L + (1 - \pi_H) C\right) \\ & - (1 + r_D) L.\end{aligned} \tag{9.37}$$

The bank is only able to repay their loan if their investment is successful, $\hat{\pi}$, in which case they obtain the return $\hat{R}$ and repay their loan. In addition, they will retain the loan from the company, if repaid and if not repaid obtain the collateral the company provided. If the investment of the bank is not successful, it cannot repay its loan and as it forfeits the collateral it has pledged, which originally belonged to the company, it cannot return the collateral, causing the company to not repay its loan. This last assumption implies that if the bank does not return the collateral to the bank, the company is under no obligation to repay the loan. As the company will have agreed for the bank to use its collateral, such an arrangement would be enforceable.

If not rehypothecating the collateral, the bank will make profits of

$$\Pi_B = \pi_H (1 + r_L) L + (1 - \pi_H) C - (1 + r_D) L. \tag{9.38}$$

The bank obtain the loan repayment if the company investment is successful and if it is not successful retains the collateral before paying its depositors that financed the loan. Rehypothecation would be preferable to the bank if it generates higher profits. We thus require that $\hat{\Pi}_B \geq \Pi_B$, from which we obtain that

$$\hat{\pi} \geq \hat{\pi}^* = \frac{\pi_H (1 + r_L) L + (1 - \pi_H) C}{\left(\hat{R} - \hat{r}_L\right) \hat{L} + \pi_H (1 + r_L) L + (1 - \pi_H) C}. \tag{9.39}$$

Thus the investment of the bank must have a sufficiently high success rate to merit rehypothecation. We easily see that for high values of the company success rate, $\pi_H \approx 1$, this requirement allows for bank investments that are riskier than the loan they provided as $\hat{\pi}^* < \pi_H$, while for more risky loans with a low value of π_H, we have $\hat{\pi}^* > \pi_H$ and the bank investment has to be less risky than the loan it provides. A realistic scenario is that the bank borrows at its deposit rate such that $\hat{r}_L = r_D$ and the return on investment is the loan rate, $\hat{R} = r_L$. In this case, the bank would need to find less risky loans to grant if the original loan is high-risk, while higher risks can be taken if the original loan was low-risk. Assuming that banks have access to investments (loans) for which the condition in equation (9.39) is fulfilled, they would like to engage in the rehypothecation of collateral companies have provided them with.

The company will now have to repay the loan only if the investment is successful and bank is able to return the collateral and will in turn lose collateral unless both the company itself and the bank are able to repay their respective loans. The company profits for exerting and not exerting efforts, respectively, are therefore given by

$$\begin{aligned} \hat{\hat{\Pi}}_C^H &= \pi_H \left((1 + R) L - \hat{\pi} (1 + r_L) L\right) - (1 - \pi_H \hat{\pi}) C - E, \\ \hat{\hat{\Pi}}_C^L &= \pi_L \left((1 + R) L - \hat{\pi} (1 + r_L) L\right) - (1 - \pi_L \hat{\pi}) C. \end{aligned} \tag{9.40}$$

Again, the effort is exerted if it is profitable for the company to do so. Requiring that $\hat{\hat{\Pi}}_C^H \geq \hat{\hat{\Pi}}_C^L$ solves for

$$L \geq L^{***} = \frac{E - \hat{\pi} (\pi_H - \pi_L) C}{(\pi_H - \pi_L) \left((1 + R) - \hat{\pi} (1 + r_L)\right)}. \tag{9.41}$$

A smaller loan is viable with rehypothecation if this constraint is less binding than the constraint without rehypthecation, $L^{***} \leq L^{**}$. Using equations (9.36) and (9.41), this solves for

$$E \geq E^* = (\pi_H - \pi_L) (1 + r_L) C. \tag{9.42}$$

If the effort costs are sufficiently high, or the collateral requirements sufficiently low, rehypothecation allows for smaller loans to be provided. Thus rehypothecation benefits those companies that seek small loans and have relatively high effort costs. Similarly, equation (9.41) can be interpreted that for a given loan size, loans can be provided to company with sufficiently low effort costs, but these effort costs must not be too small in light of equation (9.42).

In addition, we can easily show that $L^{***} \leq L^{*}$ as a comparison of equations (9.34) and (9.41) shows. Hence smaller loans are always available with rehypothecation compared to a situation in which no collateral is used. Equivalently, loans to companies with higher effort costs can be supported with rehypothecation.

The benefits to companies being able to secure smaller loans in the presence of rehypothecation arise from the additional incentive to exert effort. The likelihood of losing the collateral is increased as the company must succeed with its investment as well as the bank. While the company is compensated for that possibility by not repaying the loan if the bank's investment fails and does not return the collateral, it provides stronger incentives to reduce its own probability of the investment failing. The marginal effect this has is reduced by the factor $\hat{\pi}$ in equation (9.40), and hence more efforts are optimally to be exerted.

We can easily show that companies prefer rehypothecation as $\hat{\hat{\Pi}}_C^H \geq \hat{\Pi}_C^H$ for $(1 + r_L) L \geq C$. Hence, as long as the loan is not over-collaterised the profits of the company when allowing rehypothecation will be higher than when rehypothecation is not allowed and companies will agree to such arrangements. This arises from the fact that not having to repay the loan if the bank cannot return the collateral, which is larger than the collateral, increases the profits of the company. It is thus beneficial to the bank, if it finds a suitable investment fulfilling constrains (9.39), as well as the company.

Summary Banks may reuse the collateral they have been provided with by companies as collateral in their own borrowing. This rehypothecation allows banks to generate additional profits and companies may benefit from having easier access to loans, as well as making higher profits. The presence of moral hazard in that companies need to be incentivised to exert efforts in reducing the risk to their investments, requires that the effort costs have to be spread across a sufficiently large loan. Rehypothecation allows this loan to be smaller, or equivalently the effort costs to be higher, enabling a wider range of companies to obtain a loan.

Companies that would otherwise have no access to loan due to either their high effort costs or small loan size providing no incentives to exert effort reducing the loan risk to the bank, will readily not only agree to provide a collateral, but also agree to the bank using their collateral in rehypothecation. It can increase bank and company profits alike. In deriving these results, we have not relied on the fact that commonly collateralised loans are requiring a lower loan rate, making them more attractive to companies, but less attractive to banks. However, the presence of the collateral would compensate for these differences as chapter 9.1.1 as shown. Rehypothecation, however, would increase the value of the collateral to the bank, making it more valuable to banks, and inducing them to offer even better conditions to companies for providing collateral.

While rehypothecation may be beneficial to companies and banks alike, there are clear limits to its feasibility in practice. Unless the collateral consists of well known assets, such as securities, it is difficult to evaluate the value of the collateral for a bank; the lender to the bank will be further removed from the company owning the collateral, making it even more difficult for them to evaluate its value. While there is nothing to stop this lender to hand on the collateral to another lender to obtain a loan themselves and thereby create a collateral chain, the difficulty in evaluating the collateral becomes ever more pronounced. It is therefore most likely that we find common securities or real estate used in rehypothecation due the ease of assessing their value.

Reading Park & Kahn (2019)

9.4 Debt covenants

It is not unusual for banks and companies to agree specific conditions the company must adhere to in order to secure a loan. Such conditions might compel the borrower to refrain from certain activities, such as the selling of specific assets or expanding into new business areas. Alternatively, these conditions require the company to conduct specific activities, such as maintaining a minimum amount of liquid assets. maintain their main accounts with the lending bank, or to limit the risks of their investments. Such conditions are referred to as debt covenants. If debt covenants are broken, the bank usually has the right to require the instant repayment of the loan. The aim of debt covenants is to reduce the risks banks face and ensure the likelihood of the bank loan being repaid is increased.

In contrast to traditional collateral, with debt covenants there are no additional losses to the company if they do not repay their loan. We can nevertheless interpret debt covenants as a form of collateral as it provides additional safeguards for the bank against losses and the restrictions imposed on the decisions of the company are costly in that it will limit the profits they can generate.

Let us now assume that a company seeks a loan of size L, paying interest r_L, to make an investment. The company has available a risky investment in which it will invest L_R that will yield a return of R with probability π and will yield no return otherwise; we assume that the expected return of this investment covers the loan costs, such that $\pi(1+R) \geq 1 + r_L$ and hence providing the loan for the risky investment is efficient. The other investment is safe in that its investment L_S will always return L_S; as this investment yields no profits, it is not efficient for the bank to provide a loan for this safe investment as long as $r_L > 0$. Of course we require that the loan is fully split between these two investments such that $L = L_R + L_S$.

The company has limited liability and will only be able to repay its loan if the realised value of their investments are sufficient. If the risky investment is not yielding a return, it will be impossible for the bank to repay the loan fully, leading to zero profits to the company; this is because the safe investment does not increase in value. Hence the profits are given by

$$\Pi_C = \pi\left(\max\left\{(1+R)\,L_R + L_S; (1+r_L)\,L\right\} - (1+r_L)\,L\right). \tag{9.43}$$

The first term specifies that the company will be able to retain their assets, consisting of the successful risky investment, $(1+R) L_R$, and the safe investment, L_S, as long as it exceeds their obligation for the repayment of the loan, $(1+r_L) L$. If the assets are not sufficient to cover the loan repayment, even if the risky investment is successful, the bank will seize all assets and the company will not make any profits.

We can now distinguish two cases, firstly if $(1+R) L_R + L_S < (1+r_L) L$, which when using that $L = L_R + L_S$ becomes $L_R < \frac{r_L}{R} L$, we find that $\Pi_C = 0$. The second case of $(1+R) L_R + L_S \geq (1+r_L) L$, or $L_R \geq \frac{r_L}{R} L$, yields $\Pi_C = \pi (R L_R - r_L L)$, again using that $L = L_R + L_S$. The profits of the company are increasing in the risky investment L_R and it is optimal for companies to invest fully into this investment such that $L_R = L$. As for $L_R < \frac{r_L}{R} L < L$, we have $\Pi_C = 0$, choosing $L_R = L$, and hence $L_S = 0$, is the optimal choice of companies.

Banks will either receive the agreed loan repayment or seize the available assets of the company if these are not sufficient to repay the loan. Any repayment of the loan involving the risky asset can only be successful if this investment is successful and the bank obtains either $(1+R) L_R + L_S$ if the value of these assets are not sufficient to repay the loan, or they receive the full loan repayment $(1+r_L) L$, whichever is the smaller. If the risky investment is not successful, the bank can only seize the safe investment L_S as the risky investment has no value. Financing the loan fully by deposits on which interest r_D is payable, the bank profits are thus given by

$$\begin{aligned} \Pi_B = {} & \pi \min \{(1+R) L_R + L_S; (1+r_L) L\} \\ & + (1-\pi) L_B - (1+r_D) L. \end{aligned} \tag{9.44}$$

If we again distinguish two cases, the first being $(1+R) L_R + L_S < (1+r_L) L$, or $L_R < \frac{r_L}{R} L$, we easily get the bank profits as $\Pi_B = (\pi (1+R) - 1) L_R - r_D L$ after inserting from $L = L_R + L_S$. As by assumption $\pi (1+R) \geq 1 + r_L \geq 1$, these profits are increasing in L_R and hence the bank would like the company to maximize the risky investment; given the constraint for this case, this gives us $L_R = \frac{r_L}{R} L$.

The second case requires $(1+R) L_R + L_S \geq (1+r_L) L$, or $L_R \geq \frac{r_L}{R} L$, and hence $\Pi_B = (\pi R - r_D) L - (1-\pi) L_R$, which is decreasing in the risky investment L_R. Consequently, the bank would want the company to invest as little as possible into the risky investment; given the constraint in this case, this will be $L_R = \frac{r_L}{R} L$. Hence in both cases, the bank wants the company to make a risky investment of $L_R = \frac{r_L}{R} L$ and therefore make a safe investment of $L_S = \frac{R - r_L}{R} L$.

We can now interpret the bank's requirement for a safe investment of $L_S = \frac{R - r_L}{R} L$ as a debt covenant in which the company is prevented from using the risky investment to maximize their profits, which would have implied $L_S = 0$. The bank here insists on such a safe investment to protect partially the repayment of the loan and as companies would make a profit of $\Pi_C = 0$, they thus accept this debt covenant. If the bank's market position does not allow it to impose, through a debt covenant, its profit maximizing choice, it would be able to insist on a smaller, but nevertheless positive amount of safe investment. The bank would not insist on the company to make only the safe investment as the low return of such investments would not allow the bank to earn interest on the loan, hence they have to allow some degree of risk-taking by companies.

Therefore, banks are able to reduce the risks of loans they are providing by imposing debt covenants on companies. Requiring a certain amount of low-risk investments to safeguard the repayment of the loan, while at the same time allowing some more risky investments to generate returns that can then be used to pay interest on the loan, allows the bank to balance the risks they are exposed to and the returns that are needed to be profitable. The benefit of using debt covenants to reduce the risks for banks, is that they can be agreed with companies even if these companies do not have access to collateral and companies do not face additional costs from failing to repay their loans. On th other hand, they might limit the scope of investments the company is able to conduct, affecting their profitability. Debt covenants have the further benefit of having the potential to reduce moral hazard in making investment decisions by limiting the amount of risky investments the company can make.

Reading Berlin & Mester (1992)

9.5 Loan guarantees

It is common for banks to seek guarantees on loans they have provided to their customers; it is even often only because they can obtain such guarantees that the loan is provided. Guarantees are granted by governments or government-backed organisations which usually have the aim of promoting investments into certain regions of a country, into specific industries, or the promotion the export of goods. Facilitating the provision of loans is seen as encouraging investments by companies. If the loan the bank provides to a company is not repaid, the bank will be compensated for their losses by the guarantor; this guarantee can often only be obtained against the payment of an insurance fee that is paid by the bank. We will here investigate the incentives of banks to seek such guarantees.

We consider a bank that provides a loan L at a loan rate r_L to a company. This company use the loan proceeds to finance an investment which will succeed with probability π_i; the probability of success is either high, pi_H, or low, $\pi_L < \pi_H$. The bank does not ex-ante know the probability of success of the bank, it only knows that a fraction p of companies have a high success rate, π_H, while a fraction $1 - p$ of companies has a low success rate, π_L. They can obtain information on the type of companies they are lending to by conducting a screening of the companies, which imposes costs C on them. While these screening costs are known to the bank itself, they are not known to an outside observer, where it is only known that the screening costs follow a distribution $G(\cdot)$.

In addition, after having provided the loan, the bank faces a liquidity shock with probability γ, which will force them to sell their loans to generate additional cash reserves. The value of loans that are guaranteed are easily determined as $P_G = (1 + r_L) L$, given that any shortfall from the company not repaying the loan will be covered by the guarantee. The value of the loans that are not guaranteed, P_N, will depend on the composition of loans the bank will sell.

We will in a first step determine the condition under which the bank screens their customers and thus becomes informed about their type. If banks face a liquidity shock, which occurs with probability γ and is unobservable to the public, they sell all loans and obtain the market price, P_N; if the bank faces no shock, $1-\gamma$, they retain the fraction p of the loans they have screened and where they found them to have a high probability of being repaid, the value being $\pi_H(1+r_L)L$, and sell the fraction $1-p$ of loans with a low success rate at the market price P_N. Banks will also sell all loans that have not been screened. We assume here that the market cannot distinguish between the different types of loans. In addition the bank will have to face the screening costs and repay the deposits D used to finance the loans, including interest r_D. We thus have the profits of the bank given by

$$\Pi_B = \gamma P_N + (1-\gamma)(p\pi_H(1+r_L)L + (1-p)P_N) - C - (1+r_D)D. \quad (9.45)$$

If the bank does not screen the companies they are lending to, they sell all their loans regardless whether they face a liquidity shock or not, hence we obtain the bank profits as

$$\hat{\Pi} = P_N - (1+r_D)D. \quad (9.46)$$

We will argue below that this strategy of selling all unscreened loans and selling only screened loans with low probabilities of being repaid if no liquidity shock occurs, is optimal.

The bank will screen the companies they are lending to if this is more profitable than not screening companies, $\Pi_B \geq \hat{\Pi}_B$, which requires

$$C \leq C^* = p(1-\gamma)(\pi_H(1+r_L)L - P_N). \quad (9.47)$$

Hence, if the screening costs are not too high, banks will conduct screening of companies and thus learn the probability of default. As the screening costs are not known to an outside observer, including the purchasers of loans, they do not know whether they obtain a screened or an unscreened loan; they can only assign a probability of $G(C^*)$ that the loan has been screened.

We can now continue to determine the market value of the loan. If we assume that a fraction θ of loans are guaranteed and thus traded separately, we can determine that the amount of loans sold consists of a fraction p of loans with a low default rate, if they are screened, $G(C^*)$, and a liquidity shock occurs, γ. We will see below that none of these loans will be guaranteed. With the above said, all non-screened loans will be sold, $1-G(C^*)$ and all screened loans $G(C^*)$ that have been found with a high default rate, $1-p$. It is only those loans that are not guaranteed, $1-\theta$, that are considered in this market. We thus in total have sales of loans to the amount of $\gamma pG(C^*) + (1-pG(C^*))(1-\theta)$, with the last term has been obtained by adding $1-G(C^*)$ and $(1-p)G(C^*)$.

We now can determine the amount of revenue a purchaser would obtain from purchasing such a loan. If the bank faces a liquidity shock and sells the screened loan with a low default rate, $\gamma p G(C^*)$, the purchaser obtains the revenue of this loan, $\pi_H(1+r_L)L$. If the loan has not been screened, $1-G(C^*)$, the expected repayment of the loan is $\pi(1+r_L)L$, where $\pi = p\pi_H + (1-p)\pi_L$ denotes the average repayment rate. If the loan has been screened and found to have a low repayment probability, $(1-p)G(C^*)$, the purchaser will obtain $\pi_L(1+r_L)L$. Combining all terms we obtain the revenue to the purchaser to be $\gamma p G(C^*)\pi_H(1+L)L + ((1-G(C^*))\pi(1+r_L)L + (1-p)G(C^*)\pi_L(1+r_L)L)(1-\theta)$.

The market price of the loan is then the amount received by the purchasers, divided by the total amount of loans sold, thus

$$P_N = \frac{\left\{\begin{array}{c}\gamma p G(C^*)\pi_H(1+L)L + ((1-G(C^*))\pi(1+r_L)L \\ +(1-p)G(C^*)\pi_L(1+r_L)L)(1-\theta)\end{array}\right\}}{\gamma p G(C^*) + (1-pG(C^*))(1-\theta)}. \tag{9.48}$$

We can now easily see that $P_N \leq \pi_H(1+r_L)L$ and as the price obtained is less than the value of loans with a high probability of repayment, a bank would not sell these loans unless it has to do so if experiencing a liquidity shock; the bank would only have knowledge of the probability of repayment if it has screened the company. Similarly we have $P_N \geq \pi_L(1+r_L)L$ and banks would sell loans they have assessed as having a low probability of repayment. For unscreened companies where the bank does not know the probability of the loan being repaid, we have $P_N \leq \pi(1+R_L)L$ and hence all unscreened loans will be sold. This aligns with the assumptions we made on the sale of loans as we determined the threshold for screening, C^* in equations (9.45) and (9.46).

We need to ensure that the market prices of guaranteed and not guaranteed loans are consistent with each other for the bank to indifferent between selling either type of loan. The only difference for the bank when selling these loans is that they had to pay a fee F to obtain the guarantee. Thus the net revenue for a loan with a guarantee is $P_G - F$, where we argued above that $P_G = (1+r_L)L$, and the revenue when selling a loan that is not guaranteed, the bank obtains P_N. Using that $P_N = P_G - F$ and inserting from equation (9.48), we get the optimal fraction of loans the bank obtains a guarantee for as

$$\theta^* = 1 - \gamma p G(C^*)\frac{F-(1-\pi_H)(1+r_L)L}{\left\{\begin{array}{c}(1-pG(C^*))((1+R_L)L-F) \\ -(\pi - p\pi_H G(C^*))(1+r_L)L\end{array}\right\}}. \tag{9.49}$$

As this condition has to be met in order for the market prices of guaranteed and not guaranteed loans to be consistent with each other, we can only admit solutions that meet the obvious requirement that the fraction of loans for which banks seek a guaranteed is in the interval $[0;1]$. With our result in equation (9.49), this condition becomes

$$
\begin{aligned}
\underline{F} &= (1-\pi_H)(1+r_L)L \\
&\leq F \\
&\leq \frac{(1-\pi) - pG(C^*)(1-\gamma)(1-\pi_H)}{1-pG(C^*)(1-\gamma)}(1+r_L)L \\
&= \overline{F}.
\end{aligned} \tag{9.50}
$$

We thus see that for guarantees that charge a fee below $\underline{F}$, which represents the losses in the best-case scenario of a low-risk loan, banks would insure all loans. This is because the fee paid will be less than the losses from loans not being repaid and thus banks would make a profits from seeking guarantees. On the other hand, the fee must not be too large as otherwise no bank would seek a guarantee. As banks would insure preferably those loans they have identified as high-risk, thus having a low repayment rate, the benefits they obtain from the guarantee are $(1-\pi_L)(1+r_L)L$, which we can show to be larger than $\overline{F}$. It is thus that the guarantees to banks will not be cost-covering to ensure an orderly market for loan exist that allows banks to increase their cash reserves through the sale of loans if they face a liquidity shortage.

The guarantees provided increase the market price of the non-guaranteed loans as they reduce the adverse selection between banks and the purchasers of the loans. What induces this adverse selection is that banks sell high-risk loans and non-screened loans only if they do not face a liquidity shock, but will not sell low-risk loans; these will be sold only in the case of a liquidity shock. As the purchasers of the loans do not know whether a liquidity shock has occurred, the market price for loans if a liquidity shock occurs is low to take into account this adverse selection; the guarantee reduces the number of high-risk loans in the market as they are insured and removed from this part of the market. As a consequence, the price obtained by banks in the market increases, benefitting them. As the fee charged will have to be below the benefits banks obtain from the guarantee, they are seeking such guarantees.

Reading Ahnert & Kuncl (2024)

Conclusions

The most obvious benefit of employing collateral is that the risks banks face is reduced and companies should benefit from lower loan rates. If a company does not repay its loans, the bank will seize the asset and thereby ensure the (partial) repayment of the loan; this reduces the banks' losses if the investments of companies are not successful. These reduced losses should be reflected in a lower loan rate, which will benefit the company. On the other hand the company will lose the collateral if they are unable to repay their loan, imposing losses onto the company in addition to the losses arising from unsuccessful investments.

The impact of collateral has goes beyond this reduction in risks and loan rates. Firstly does the provision of collateral by companies allow to overcome differences in opinions between companies and their banks on the prospects of the financed investment. By providing a collateral, the risks for the bank recede sufficiently to reduce the loan rate substantially, while the increased risk of losing the collateral increases less due the company's perceived lower risk. This benefits the banks and the company alike. The requirement to provide the bank with collateral also provides incentives to the company to exert high level of effort ensuring the investment succeeds and the collateral is not lost. Thus collateral does not only reduce the risk to banks, but also affects the behaviour of the company itself. While collateral may affect the risk-taking behaviour of companies, banks may use debt covenants to limit the risk-taking of companies. By restricting the type of investment a company can make, the bank can increase the company's ability to repay its loan. Using such a debt covenant does not rely on incentives, but instead imposes a direct constraint on the behaviour of the company. It might be particularly attractive to banks where the company they are lending to has no collateral or the collateral they could provide is of limited value to the bank, for example because it is difficult to sell.

While collateral, and debt covenants, are able to affect the risk-taking behaviour of companies, banks are often struggling to identify the risks of companies properly. This might not only lead to a difference in opinion, but a situation in which banks are not able to distinguish companies taking on different risk levels. The use of collateral can allow a distinction between companies of different risk levels as high-risk firms are preferring to not offer collateral, given the high risk of losing the collateral due to their l=high likelihood of failing to repay the loan, while those companies taking lower risks, will provide collateral. This allows banks to distinguish companies of different risks by observing their willing ness to provide collateral in exchange for a lower loan rate.

Bank having been provided with collateral, may use this collateral to secure loan they themselves obtain, a process called rehypothecation. While collateral might be lost if the bank cannot repay its loan, the company originating the collateral might benefit from such an arrangement as the loss of the collateral would absolve it from repaying the original loan, increasing its profits. In the same way, the company also has more incentives to exert effort, as long as the costs of doing so are not too high, to reduce the risk of the company itself not being able to repay the loan and thus lose the collateral. The risk of losing the collateral now has two sources, the failure of the company to repay its loan and the failure of the bank to repay their own loan. This reduces the marginal impact of the company's effort on the likelihood of losing the collateral and companies will compensate for this by increasing their effort levels. Provided the costs of such effort is not too high, this will result in increased efforts and the rehypothecation of collateral will reduce the risks companies take.

We have seen that collateral can have more widespread effect than merely reducing the risks to banks. While this effect is clearly present, collateral also affects the moral hazard in companies' investment decision. Taking into account the addition costs from losing the collateral if not repaying the loan, the company will take additional measures to reduce this risk. Collateral thus affects the risk-levels taken by companies. In addition, the willing ness to provide collateral can also provide information to banks on the riskiness of a company and thereby reduce adverse selection between the company and its bank, helping the loan market to function properly.

Chapter 10

Credit reference agencies

Credit reference agencies, also called credit bureaus in the Unites States, collects financial information of individuals and companies. This information is provided by banks or other companies that provide consumer finance, and typically encompasses information on the existence of current accounts, loans and similar credit arrangements, such as arranged overdrafts, leases, or mobile phone contracts, but also loans applied for and not taken up or refused. They are also provided with repayment habits of the borrower, such as missed or late repayments or exceeding any overdraft arrangements. This information is then provided to other banks and finance companies to allowed these a better assessment of the creditworthiness of their borrowers. Especially for private individuals, credit reference agencies often combine this type of information with other personal data, for example the occupation, salary, location, and age, to determine a credit score, which aims at providing an assessment of the risks this borrower might pose to a bank. However, frequently banks will complement this assessment by the credit reference agency with their own credit risk assessment rather than relying solely on the assessment of the credit reference agency.

In this chapter we will assess the willingness of banks to share information with credit reference agencies and thereby indirectly with their competitors. Any information banks have on their own customers will provide them with an advantage over competitors without this information. If, based on the information of its bank, a company is of lower risk than other banks would assess the company at, the bank has the advantage that it could provide the company with a loan offer, that competitors could not match, while still making profits. On the other hand, if the company is assessed to be of higher risk than a competitor would assess the company as, the bank would lose this company as competitors could provide them with better loan conditions. However, assuming that the assessment of the company is correct, their competitors would make a loss from companies switching to them, causing an adverse selection problem between banks.

We will evaluate why banks are sharing information with their competitors and reduce the competitive advantage they have from access to information about their companies. In particular, we will explore in chapter 10.1 how adverse selection between banks due to the different levels of information they have about a company affects which companies would prefer banks to disclose information about them and which companies would prefer that such information is not disclosed. Not being able to offer loans that accurately reflect the risks a company is taking, may provide incentives to companies to increase such risks. Information disclosure can be used to reduce such moral hazard, as we will see in chapter 10.2, as it allows to take into account the risks companies take and a higher loan rate to account for these risks might well incentivize companies to not take on higher risks. Information disclosure does not only affect the profits of companies as it reduces adverse selection and moral hazard, but the informational advantage a bank can gain from having more information on a company will affect the competition between banks. Therefore, chapter 10.3 will explore the impact information disclosure has in this respect.

10.1 Preferences for information disclosure

By a bank providing information to credit reference agencies, other banks can make better inferences about the risks this company faces. Such information disclosure can only occur if companies agree, usually as part of the terms and conditions of entering any contract with the bank. In order for banks to obtain such an agreement, it must be beneficial for companies for other banks to hold this information, while at the same time be at least not detrimental to the bank itself to provide this information to the credit reference agency.

Let us assume that there are two types of companies in the market. A fraction ν of the companies will use the loan L to make an investment that generates a return of R with some probability π, it is thus capable of generating successful investments. The remaining fraction of $1 - \nu$ companies cannot make investments that allow the company to repay its loan, they are thus not able to generate successful investments. Due to non-pecuniary benefits, companies that cannot generate any successful investments are nevertheless demanding loans; however, when assessing the incentives of companies we will only explore those of companies that are able to generate successful investments. Each company can make identical investments for two subsequent time periods and a failure to repay their loan in time period 1 after the investment has not been successful does not affect their ability to obtain another loan for their investment in time period 2. While companies know their type, banks only learn the type of company after they have lent to the company in time period 1, thus a bank who has not lent to the company in time period 1 has no information about the type of the company unless the initial bank decides to disclose any information through credit reference agencies.

After time period 1, companies can switch their loan to another bank, but we assume that this involves costs of S. Such costs may arise from the prolonged assessment of their credit worthiness by the new bank, the set-up of new accounts, or the work involved in providing the new bank with all relevant information. In time period 1, companies do not know these costs and only learn them in time period 2 as they make the decision whether to change their bank or not. However, we assume that these costs are distributed uniformally with a minimum of zero and a maximum cost of $\overline{S}$, hence $S \in \left[0; \overline{S}\right]$. The distribution function is therefore given by

$$F\left(S^*\right) = \text{Prob}\left(S \leq S^*\right) = \frac{1}{\overline{S}} \int_0^{S^*} dS = \frac{S^*}{\overline{S}}. \tag{10.1}$$

Of course, if a bank knows that the company will not repay their loan as their type is such that the investment will never succeed, they will not lend to them; consequently all companies that cannot generate a successful investment will switch banks to secure a loan from another bank. As banks does not know the type of company when lending commences in time period 1, banks cannot discriminate between companies of different types until they have learned this type prior to any lending in time period 2, in time period 2 banks can discriminate their loan rates between those companies that have switched to them and those that have not switched and hence whose type they know, where, as noted above, companies not able to generate successful investments will switch banks..

Let us first consider the case where banks do not disclose any information about the company they are lending to before then considering the disclosure of information to credit reference agencies.

No information disclosure Analysing the lending decision in time period 2 first, we know that companies will generate a successful investment with probability π, which then allows them to repay their loan. If they stay with their existing bank, they will be charged a loan rate r_L^2 and if they change to another bank, they will be charged a loan rate of $\hat{r}_L^2$, in addition to facing switching costs S. The profits of the company in the second time period for staying with their existing bank and switching to another bank, respectively, are thus given by

$$\begin{aligned} \Pi_C^2 &= \pi\left((1+R)L - \left(1+r_L^2\right)L\right), \qquad (10.2) \\ \hat{\Pi}_C^2 &= \pi\left((1+R)L - \left(1+\hat{r}_L^2\right)L\right) - S. \end{aligned}$$

If $\hat{\Pi}_C^2 \geq \Pi_C^2$, the company is better off switching to another bank as its profits will be higher. We can rewrite this condition as

$$S \leq S^* = \pi\left(\left(1+r_L^2\right) - \left(1+\hat{r}_L^2\right)\right)L. \tag{10.3}$$

If banks do not know the switching costs of companies, but are only aware of their distribution, the bank can infer from the distribution of switching costs in equation (10.1) that the probability of a company switching banks is given as $F\left(S^*\right) = \frac{S^*}{\overline{S}}$. Similarly with the company not knowing their switching costs in time period 1, they will assign the same probability that they themselves will switch banks in time period 2.

The initial bank will in time period 2 only lend to the fraction ν of companies it has been identified as being able to generate successful investments. Hence they will lend again to these companies, provided they do not switch. With a fraction of $1 - F(S^*)$ remaining with their initial bank, the profits the bank will make from their existing companies is given by

$$\begin{aligned} \Pi_B^{2,A} &= \nu\left(\pi\left(1+r_L^2\right)L - (1+r_D)L\right)(1-F(S^*)) \qquad (10.4) \\ &= \frac{\nu}{\overline{S}}\left(\pi\left(1+r_L^2\right) - (1+r_D)\right)L \\ &\quad\times\left(\overline{S} - \pi\left(\left(1+r_L^2\right) - \left(1+\hat{r}_L^2\right)\right)L\right), \end{aligned}$$

where r_D denotes the interest on deposits that finance the loan and we have used the expression for S^* from equation (10.3), together with the probability distribution in equation (10.1). Maximising these profits over the optimal loan rate to charge their existing companies, we get the first order condition that

$$\begin{aligned} \frac{\partial \Pi_B^{2,A}}{\partial\left(1+r_L^2\right)} &= \frac{\nu\pi L}{\overline{S}}\left(\overline{S} - \pi\left(\left(1+r_L^2\right) - \left(1+\hat{r}_L^2\right)\right)L\right. \\ &\qquad \left. - \left(\pi\left(1+r_L^2\right) - (1+r_D)\right)L\right) \\ &= = 0, \end{aligned}$$

which easily solves for the loan rate to become

$$1 + r_L^2 = \frac{\pi\left(1+\hat{r}_L^2\right) + (1+r_D) - \frac{\overline{S}}{L}}{2\pi}. \qquad (10.5)$$

In addition to their existing companies, the bank will also attract companies switching from other banks, but it will not know its type. Therefore, it will make a loss from all those who are unable to repay their loan, a fraction of $1 - \nu$, as all of them will switch after being denied loans by their initial bank. On the other hand, only a fraction $F(S^*)$ of companies able to generate successful investments are switching banks, where S^* is again defined in equation (10.3). Hence we have

$$\begin{aligned} \Pi_B^{2,B} &= \nu\left(\pi\left(1+\hat{r}_L^2\right)L - (1+r_D)L\right)F(S^*) - (1-\nu)(1+r_D)L \qquad (10.6) \\ &= \frac{\nu\pi}{\overline{S}}\left(\pi\left(1+\hat{r}_L^2\right) - (1+r_D)\right)L \\ &\quad\times\left(\left(1+r_L^2\right)L - \left(1+\hat{r}_L^2\right)L\right) - (1-\nu)(1+r_D)L. \end{aligned}$$

Maximizing profits the bank can make from those companies that switch to them, gives rise to the first order condition

$$\frac{\partial \Pi_B^{2,B}}{\partial\left(1+\hat{r}_L^2\right)} = \frac{\nu\pi}{\overline{S}}\left(\pi\left(1+r_L^2\right)-\pi\left(1+\hat{R}_L^2\right)\right.$$
$$\left.-\pi\left(1+\hat{r}_L^2\right)+\left(1+r_D\right)\right)L^2=0,$$

and hence

$$1+\hat{r}_L^2 = \frac{\pi\left(1+r_L^2\right)+\left(1+r_D\right)}{2\pi}. \tag{10.7}$$

Combining equations (10.5) and (10.7) we get the equilibrium loan rates as

$$\begin{aligned} 1+r_L^2 &= \frac{1+r_D}{\pi}+\frac{2}{3}\frac{\overline{S}}{\pi L}, \\ 1+\hat{r}_L^2 &= \frac{1+r_D}{\pi}+\frac{1}{3}\frac{\overline{S}}{\pi L}. \end{aligned} \tag{10.8}$$

We easily see that $1+r_L^2 > 1+\hat{r}_L^2$ and the initial bank charges a higher interest rate as it exploits its market power arising from the switching costs S. As we can easily derive when inserting equations (10.8) into equation (10.3), we have $S^* = \frac{1}{3}\overline{S}$ and $\frac{1}{3}$ of companies will switch banks.

The total profits of banks are from those companies that stay with them as well as those that switch to them, hence the total period 2 profits of banks are given by

$$\Pi_B^2 = \Pi_B^{2,A}+\Pi_B^{2,B} = \frac{5}{9}\nu\overline{S}-(1-\nu)\left(1+r_D\right)L \tag{10.9}$$

as we insert the solutions for the loan rate from equations (10.8) into equations (10.4) and (10.6).

In time period 1, the bank does not know the type a company is, hence it can only make profits if it provides a loan to a company that is able to generate successful investments and the investment is actually successful. Thus

$$\Pi_B^1 = \nu\pi\left(1+r_L^1\right)L-\left(1+r_D\right)L. \tag{10.10}$$

If we assume that banks are competitive, they will compete for customers in period 1 such that $\Pi_B = \Pi_B^1+\Pi_B^2 = 0$, hence

$$\Pi_B = \nu\pi\left(1+r_L^1\right)L-\left(1+r_D\right)L+\frac{5}{9}\nu\overline{S}-(1-\nu)\left(1+r_D\right)L=0, \tag{10.11}$$

which gives rise to a loan rate in time period 1 of

$$1+r_L^1 = \frac{2-\nu}{\nu\pi}\left(1+r_D\right)-\frac{5}{9}\frac{\overline{S}}{\pi L}. \tag{10.12}$$

The profits to companies that do not switch are consisting of the profits in time period 1 and time period 2, giving us

$$\begin{aligned}\Pi_C &= \pi\left((1+R)L - \left(1+r_L^1\right)L\right) \\ &\quad +\pi\left((1+R)L - \left(1+r_L^2\right)L\right) \\ &= 2\pi(1+R)L - 2\frac{1-\nu}{\nu}(1+r_D)L - \frac{1}{9}\overline{S},\end{aligned} \tag{10.13}$$

inserting for the loan rates from equations (10.8) and (10.12). Similarly, for those companies that do switch banks, their profits are given by

$$\begin{aligned}\hat{\Pi}_C &= \pi\left((1+R)L - \left(1+r_L^1\right)L\right) \\ &\quad +\pi\left((1+R)L - \left(1+\hat{r}_L^2\right)L\right) - S \\ &= 2\pi(1+R)L - 2\frac{1-\nu}{\nu}(1+r_D)L + \frac{2}{9}\overline{S} - S.\end{aligned} \tag{10.14}$$

We only consider the profits those companies that are able to generate successful investments as the other type of companies will be indifferent to any loan conditions, given they will never be able to repay the loan.

Companies do not know their switching costs in time period 1, hence can only infer the likelihood of switching banks, given by $F(S^*)$, such their expected profits are given by

$$\begin{aligned}\overline{\Pi}_C &= \Pi_C\left(1 - F(S^*)\right) + \hat{\Pi}_C F(S^*) \\ &= 2\pi(1+R)L - 2\frac{1-\nu}{\nu}(1+r_D)L - \frac{1}{18}\overline{S},\end{aligned} \tag{10.15}$$

noting that the last term in equation (10.14) arises from $\frac{1}{\overline{S}}\int_0^{S^*} SdS = \frac{1}{18}\overline{S}$, as the switching costs are only incurred if $S \leq S^* = \frac{1}{3}$.

for companies to demand loans in time period 1, we would require that it is profitable to do so, thus we require that $\overline{\Pi} \geq 0$. Hence for loan demand to exist we find that the fraction of companies that are able to generate successful investments has to exceed at least

$$\nu \geq \frac{36(1+r_D)}{36(1+r_D) + 36\pi(1+R) - \frac{\overline{S}}{L}}. \tag{10.16}$$

As for a viable solution we obviously need $\pi(1+R) - (1+r_D) \geq 0$ such that investments earn at least their costs of funding, we see that the requirements with a small switching costs $\frac{\overline{S}}{L}$ and not too high return on investment R are close to at least $\frac{1}{2}$ of companies being able to generate successful investments.

Using this result as a benchmark, we can now consider the case where the bank uses credit reference agencies to disclose information about the company. We will consider cases where a defaulting company is assessed as being not creditworthy first, before than looking at the case of creditworthy companies.

Information disclosure if companies are not creditworthy Banks may disclose whether a company has repaid their loan or not. If the bank reports that the company has repaid its loan, it is obvious that it is a company that is able to generate successful investments and hence the other banks can infer for the second time period that the probability of the company being able to generate successful investments is $\hat{\nu}_N = 1$. Companies not repaying their loan cannot be readily assigned a type as this might be due to companies not being able to generate successful investments or they are able to generate such investments, but have not been successful this time. Bayesian learning allows banks to update their beliefs about the likelihood of the company being able to generate successful investments. Acknowledging that the prior belief of such banks on the likelihood of companies being able to generate successful investments is ν and the probability of a success being π, we obtain the new belief as

$$\hat{\nu}_D = \frac{\nu (1 - \pi)}{\nu (1 - \pi) + (1 - \nu)}. \tag{10.17}$$

The numerator represents the likelihood that a company is able to generate successful investments, ν, but defaults, $1 - \pi$, and the denominator the likelihood of observing a default, which consists of the company being able to generate successful investments but failing in the first time period, in addition to the company not being able to generate successful investments at all.

The initial bank will know the type of company and lend to them if they are able to generate successful investments, provided they do not switch. For those companies not defaulting, the bank will face competition from other banks, while for those defaulting we here assume that $\hat{\nu}_D \pi (1 + R) L - (1 + r_D) L < 0$ and other banks would not provide a loan as on average the company will not be able to repay the loan and they would make a loss. Such companies are regarded as not creditworthy. Hence, due to a lack of competition, the initial bank can charge the maximum interest rate $1 + r_L^2 = 1 + R$ to these companies. Thus we have the profits of the initial bank in time period 2 given as

$$\begin{aligned} \Pi_B^{2,A} &= \nu\pi \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) (1 - F(S^*)) \\ &\quad + \nu (1 - \pi) (\pi (1 + R) L - (1 + r_D) L) \\ &= \frac{\nu\pi}{\overline{S}} \left(\pi \left(1 + r_L^2\right) - (1 + r_D)\right) L \\ &\quad \times \left(\overline{S} - \pi \left(\left(1 + r_L^2\right) - \left(1 + \hat{r}_L^2\right)\right) L\right) \\ &\quad + \nu (1 - \pi) (\pi (1 + R) L - (1 + r_D) L), \end{aligned} \tag{10.18}$$

where S^* is defined as in equation (10.3); the provision of information does not alter the incentives to switch banks. The first term denotes those companies that have been successful in the first time period, of which a fraction $F(S^*)$ do switch, and the second term encompasses those companies that have not repaid their loans in the first time period and who therefore cannot switch as they are regarded as not creditworthy by other banks. The initial bank, however, knows their type and therefore assess them as creditworthy.

Maximizing these profits for the loan rate in time period 2 yields the first order condition

$$\begin{aligned}\frac{\partial \Pi_B^{2,A}}{\partial\left(1+r_L^2\right)} = \frac{\nu\pi^2 L^2}{\overline{S}} & \left(\overline{S} - \pi\left(1+r_L^2\right)L + \pi\left(1+\hat{r}_L^2\right)L\right. \\ & \left. - \left(\pi\left(1+r_L^2\right) - (1+r_D)\right)\right) = 0,\end{aligned}$$

which easily solves for

$$1 + r_L^2 = \frac{\pi\left(1+\hat{r}_L^2\right) + (1+r_D) - \frac{\overline{S}}{L}}{2\pi}. \tag{10.19}$$

By assumption, it is not profitable for the other bank to lend to those companies that have defaulted, hence none of these companies are switching away from the initial bank. This gives us bank profits for the other banks that rely on those companies having succeeded in the first time period only and switching banks, such that the bank profits are given by

$$\begin{aligned}\Pi_B^{2,B} &= \nu\pi\left(\pi\left(1+\hat{r}_L^2\right)L - (1+r_D)L\right)F\left(S^*\right) \qquad (10.20)\\ &= \frac{\nu\pi^2}{\overline{S}}\left(\pi\left(1+\hat{r}_L^2\right)L - (1+r_D)L\right)\left(\left(1+r_L^2\right) - \left(1+\hat{r}_L^2\right)\right)L.\end{aligned}$$

Maximizing these profits over the loan rate charged to switching companies gives us the first order condition as

$$\frac{\partial \Pi_B^{2,B}}{\partial\left(1+\hat{r}_L^2\right)} = \left(\pi\left(\left(1+r_L^2\right) - \left(1+\hat{r}_L^2\right)\right) - \left(\pi\left(1+\hat{r}_L^2\right) - (1+r_D)\right)\right)L^2 = 0,$$

which gives us the loan rate as

$$1 + \hat{r}_L^2 = \frac{\pi\left(1+r_L^2\right) + (1+r_D)}{2\pi} \tag{10.21}$$

From equations (10.19) and (10.21) we get the same loan rates as in the absence of information sharing. Thus as in equation (10.8), we have the loan rates given by

$$\begin{aligned}1 + r_L^2 &= \frac{1+r_D}{\pi} + \frac{2}{3}\frac{\overline{S}}{\pi} \qquad (10.22)\\ 1 + \hat{r}_L^2 &= \frac{1+r_D}{\pi} + \frac{1}{3}\frac{\overline{S}}{\pi}.\end{aligned}$$

Using these loan rates, we get the profits of banks in time period 2 from both existing and switching companies, given by

$$\Pi_B^2 = \Pi_B^{2,A} + \Pi_B^{2,B} = \frac{5}{9}\nu\pi\overline{S} + \nu\left(1-\pi\right)\left(\pi\left(1+R\right) - \left(1+r_D\right)\right)L. \tag{10.23}$$

Competition between banks will again lead to competitive loan rates in time period 1 such that $\Pi_B = 0$ with Π_B^1, as given in equation (10.10) for the profits of the first time period, because the profits are unaffected by the disclosure of information in the future. This requirement solves for

$$1 + r_L^1 = \frac{1+\nu\left(1-\pi\right)}{\pi\nu}\left(1+r_D\right) - \frac{5}{9}\overline{s} - \left(1-\pi\right)\left(1+R\right). \tag{10.24}$$

A company being successful in time period 1 would be charged a loan rate of $1 + r_L^2$ by its own bank if successful and $1 + R$ if not successful, as it cannot switch banks. Hence the expected loan rate in time period 2 is

$$\begin{aligned} E\left[1+r_L^2\right] &= \pi\left(1+r_L^2\right) + \left(1-\pi\right)\left(1+R\right) \\ &= \left(1+r_D\right) + \frac{2}{3}\frac{\overline{S}}{L} + \left(1-\pi\right)\left(1+R\right). \end{aligned}$$

Thus the profits of companies not switching and switching, respectively, are given by

$$\begin{aligned} \Pi_C &= 2\pi\left(1+R\right)L - \pi\left(1+r_L^1\right)L - \pi E\left[1+r_L^2\right]L \\ &= 2\pi\left(1+R\right)L - \frac{1}{3}\pi\overline{S} - \frac{1+\nu}{\nu}\left(1+r_D\right)L \\ \hat{\Pi}_C &= 2\pi\left(1+R\right)L - \pi\left(1+r_L^1\right)L - \pi\left(1+\hat{r}_L^2\right)L \\ &= \pi\left(3-\pi\right)\left(1+R\right)L - \frac{2-5\pi}{9}\overline{S} \\ &\quad + \frac{1+\nu\left(2-\pi\right)}{\nu}\left(1+r_D\right)L - S. \end{aligned} \tag{10.25}$$

The average profits are then given as

$$\begin{aligned} \overline{\Pi}_C &= \frac{2}{3}\Pi_C + \frac{1}{3}\left(\hat{\Pi}_C + S\right) - \frac{1}{18}\overline{S} \\ &= \frac{\pi\left(7-\pi\right)}{3}\left(1+R\right)L - \frac{9+2\pi}{54}\overline{S} \\ &\quad - \frac{3+\nu\left(4-\pi\right)}{3\nu}\left(1+r_D\right)L, \end{aligned} \tag{10.26}$$

taking into account the probability of switching banks is given by $\frac{1}{3}$ and that the expected switching costs are given by $\frac{1}{\overline{S}}\int_0^{S^*} SdS = \frac{1}{18}\overline{S}$. Comparing this expression with the company profits in the case of no information disclosure from equation (10.15), we see that unless $\overline{S}$ is prohibitively large, these profits are higher and companies that are assessed as not being creditworthy by banks relying on the disclosed information, prefer information disclosure. The low probability of success of these companies, making them not creditworthy, allows the initial bank to have a substantial informational advantage over other banks, which prevents them from competing effectively in time period 2. Disclosing information on them will benefit those companies that are assessed as being able to generate successful investments while those that are not so assessed face no detriment as they are able to secure loans from other banks in either case; this makes the disclosure of information attractive to such companies.

As the final case, we will now consider the companies that are assessed as being creditworthy based on the information provided to the credit reference agency.

Information disclosure if companies are creditworthy If we now assume that defaulting companies are still creditworthy because $\nu_D \pi (1 + R) - (1 + r_D) \geq 0$ and the expected returns from the investment exceeds the funding costs of the loans, the other banks would be willing to lend to defaulting companies. This means that the initial bank faces competition from other banks due to their own defaulting companies being able to switch banks..

Defining the threshold for of the switching costs for switching banks as derived in equation (10.3) for companies not defaulting and defaulting, respectively, as

$$S \leq S^* = \pi \left(\left(1 + r_L^2\right) - \left(1 + \hat{r}_L^2\right) \right) L, \tag{10.27}$$
$$S \leq S^{**} = \pi \left(\left(1 + r_L^{2,D}\right) - \left(1 + \hat{r}_L^{2,D}\right) \right) L,$$

we obtain the profits banks make from their own companies and those switching towards them as

$$\begin{aligned} \Pi_B^{2,A} &= \nu\pi \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) (1 - F(S^*)) \\ &\quad + \nu (1 - \pi) \left(\pi \left(1 + r_L^{2,D}\right) L - (1 + r_D) L\right) (1 - F(S^{**})), \\ \Pi_B^{2,B} &= \nu\pi \left(\pi \left(1 + \hat{r}_L^2\right) L - (1 + r_D) L\right) F(S^*) \\ &\quad + \nu (1 - \pi) \left(\pi \left(1 + \hat{r}_L^{2,D}\right) L - (1 + r_D) L\right) F(S^{**}). \end{aligned} \tag{10.28}$$

Note that we use the actual ν and not the updated beliefs of a company being able to generate successful investments, $\hat{\nu}_D$, as the bank will experience the actual quality of companies, given it lends to defaulting and non-defaulting companies and the fact they are creditworthy and therefore able to repay the loan .

These expressions are identical to the case where no information was disclosed as comparison with equations (10.4) and (10.6) shows, hence as in equation (10.8) we will get the loan rates in the second time period as

$$1 + r_L^2 = 1 + r_L^{2,D} = \frac{1 + r_D}{\pi} + \frac{2}{3}\frac{\overline{S}}{\pi L}, \tag{10.29}$$
$$1 + \hat{r}_L^2 = 1 + \hat{r}_L^{2,D} = \frac{1 + r_D}{\pi} + \frac{1}{3}\frac{\overline{S}}{\pi L}.$$

We see that whether a company defaults in the first time period or not, does not affect their loan rates in the second time period. This is due to the competition between banks for all companies.

As the bank profits of the second time periods are given by $\Pi_B^2 = \Pi_B^{2,A} + \Pi_B^{2,B} = \frac{5}{9}\nu\overline{S}L$, we get with perfect competition implying that $\Pi_B = \Pi_B^1 + \Pi_B^2 = 0$, the loan rate in the first time period as

$$1 + r_L^1 = \frac{1 + r_D}{\nu\pi} - \frac{5}{9}\frac{\overline{S}}{\pi L}. \tag{10.30}$$

Following the same steps as in previous cases, we easily get the profits of the companies not switching banks and those switching banks, respectively, as

$$\Pi_C = 2\pi\,(1 + R)\,L - \frac{1 + \nu}{\nu}\,(1 + r_D)\,L - \frac{1}{9}\overline{S}, \tag{10.31}$$
$$\hat{\Pi}_C = 2\pi\,(1 + R)\,L - \frac{1 + \nu}{\nu}\,(1 + r_D)\,L + \frac{2}{9}\overline{S} - S.$$

We see that all but the second terms are identical to the profits in the case of no information disclosure in equations (10.13) and (10.14). Analysing the second term, we see that if $\nu > \frac{1}{3}$, then this term is smaller without information disclosure. Hence, creditworthy companies would prefer no information to be disclosed as long as there is a sufficiently large fraction of companies that are able to generate successful investments. This is due to the adverse selection between banks being small enough to ensure that banks are sufficiently competitive as the initial bank has as not too large informational advantage over the other banks. The reduced adverse selection in time period 2 will increase competition between banks and thus lower loan rates for successful companies, but this is compensated for by less fierce competition for banks to provide the initial loan and obtain the information in on the company type in time period 2. This competition is less fierce, though, as the profits from unsuccessful companies in time period 2 are smaller. Given the high success rate of companies, the lower loan rates in time period 2 do not fully compensate for the lower loan rates in time period 1.

Summary With banks facing perfect competition, companies are able to extract all surplus from banks and will make a profit from their investments. We have seen that the preferences in terms of the disclosure of information differ between companies that are creditworthy and those that are not creditworthy, if assessed based on the information provided by the credit reference agency. If defaulting companies remain creditworthy, the adverse selection of other banks lending to switching companies are low, especially if combined with a sufficiently large fraction of companies being able to generate successful investments; consequently the benefits of information disclosure to companies is small as loan rates will remain low even without information disclosure. The lower level of competition to attract companies in time period 1 due to lower profits in time period 2 is aggravated by banks making less profits from unsuccessful companies, which creditworthy companies are unlikely to be. If the adverse selection is higher, though, such that companies after default would be assessed as not creditworthy, the loan rate in time period 2 would on average be higher and companies prefer information to be disclosed. The higher adverse selection will require the banks to also charge relatively high loan rates in the first time period, making information disclosure preferred by companies.

Overall therefore, high risk companies prefer information disclosure as it reduces adverse selection and opens a way of obtaining loans after default if their true qualities are known in the case that they have not defaulted. Low risk companies are worse off as those companies defaulting will suffer higher interest rates with information disclosure and they therefore prefer this information to not be disclosed. In all cases, the use of collateral in combination with information disclosure is the least preferred option for companies for the reason that the loss of collateral is not compensated sufficiently by low loan rates due to the reduced adverse selection arising from the disclosure of information.

It is thus that we should find the disclosure of information in particular in markets of high risk lending. This might include loans to small, innovative companies or highly leveraged companies. We might find disclosure of information also for individual borrowers that are seeking loans where high-risk borrowers are a common occurrence, such as mobile phone contracts or unsecured lending.

Reading Karapetyan & Stacescu (2014)

10.2 Disclosure of existing loans

Banks do not only provide credit reference agencies with information about companies repaying their loans, but typically also about them providing a loan, or even about applications for loans, even if these are not granted or the company rejects a loan offer. This information is particularly valuable in situations where the bank does not hold complete information on the financial position of a company, for example if a loan has only recently been approved. Of special concern is this information for individual borrowers who do not have to present accounts showing their financial obligations from other loans and comparable commitments.

We assume that companies can make one of two distinct investments. Both investments yield an outcome of V_H^i if successful, which happens with probability π and $V_L^i < V_H^i$ if the investment is not successful. These investments only differ in their size, not their probability of success. A small investment S requires a loan of L, while a large investment L requires a loan of $2L$ but the large investment is less efficient as we assume that $V_j^L < 2V_j^S$; thus despite requiring a loan of twice the size, compared to the small investment, the outcomes is less than twice as large.

In addition to requiring a loan for their investment, companies hold equity E, that will also be used, if necessary to repay the loan. Furthermore, the large investment carries a benefit to the company in that its outcomes includes a private benefit to the company, making up a fraction ϕ of the outcome. This private benefit is not available as a resource to repay the loan but accrues only to the company directly, thus only $(1 - \phi) V_j^i + E$ is available to repay loans. Such private benefits may include the accumulation of knowledge that may be utilised in later investments, or the build-up of a stronger market position that will allow the company to generate more profits in the future. We can interpret the fraction of private benefits ϕ as an indication of the importance of moral hazard by the company; the private benefits that may be retained even if loans are not repaid will provide an incentive to conduct large investments as the small investment does not carry this private benefit to the company.

The small investment is socially desirable in the sense that its outcome, even if the investment is not successful, will always be sufficient to cover the costs of the loan, r_L; thus we assume that $V_L^S + E > (1 + r_L) L$. On the other hand, the funds available to the company repaying the loan for the large investment, $(1 - \phi) V_i^L + E$, do not always allow to repay the loan. We assume that $2 (1 + r_L) L > (1 - \phi) V_L^L + E$ and the loan does not cover its costs if the investment is not successful; in the case where it succeeds the loan amount may or may not me covered by the outcome. This induces the moral hazard mentioned previously in that the bank would generally prefer companies to choose the safe and small investment over the large and more risky investment, while the company may well prefer the large investment to obtain the private benefits.

Each bank provides a loan of size L only and as the company is not required to disclose truthfully the type of investment it makes, the bank cannot know whether their loan is the only loan the company obtains and hence the small investment is conducted, or whether they obtain loans from two banks allowing them to conduct the large investment. If the large investment is conducted, we further assume that the bank providing the first loan obtains a more senior loan that is served first, while the bank providing the second loan, being a subordinate loan, will only be repaid if the senior loan has been fully served. Unless information on the existence of loans is disclosed, banks will have no information which loan they are providing and thus will assign an equal probability to either possibility.

We can now analyse the implications of the provision of information to credit reference agencies about the existence of loans. Having this information, would allow banks to know whether they are providing the first or only loan, or, if applicable, the second loan to the company. We commence by considering the situation in which no such information is shared.

No information disclosure If companies do consider the small investment, banks know they are providing the only loan and they are repaid either the loan amount, including interest r_L^S, or if this amount cannot be repaid, they seize the outcome of the investment as well as the equity of the company. Hence bank profits are given by

$$
\begin{aligned}
\Pi_B^S &= \pi \min\left\{\left(1+r_L^S\right)L; V_H^S + E\right\} \\
&\quad + (1-\pi)\min\left\{\left(1+r_L^S\right)L; V_L^S + E\right\} - (1+r_D)L \\
&= \left(1+r_L^S\right)L - (1+r_D)L,
\end{aligned}
\tag{10.32}
$$

where for the final equality we made use of our assumption that $V_L^S + E > \left(1+r_L^S\right)L$. If we assume that banks are competitive such that $\Pi_B^S = 0$, we get the loan rate as $r_L^S = r_D$, ensuring the assumption on the company being able to repay the loan in all circumstances is fulfilled.

If the loan demanded is to be used for the large investment, banks can be either providing the first or second loan. The bank providing the first loan obtains either the loan repayment or if the company cannot make this payment, it will seize the available assets of the company; these assets consist of the fraction of the outcome that is available to repay the loan, as well as their equity. We thus obtain that

$$
\begin{aligned}
\hat{\Pi}_B^1 &= \pi \min\left\{(1+r_L)L; (1-\phi)V_H^L + E\right\} \\
&\quad + (1-\pi)\min\left\{(1+r_L)L; (1-\phi)V_L^L + E\right\} \\
&\quad - (1+r_D)L \\
&= (1+r_L)L - (1+r_D)L,
\end{aligned}
\tag{10.33}
$$

where for the final equality we made use of our assumption that $(1-\phi)V_H^L + E > 2(1+r_D)L$. Thus the first loan is certain to be repaid and the bank faces no risk, implying that there is no differences in the bank providing the first loan for a large investment or the only loan for a small investment.

If the bank, on the other hand, provides the second loan, this loan will only be repaid if the first loan has been repaid in full, giving the bank profits of

$$
\begin{aligned}
\hat{\Pi}_B^2 &= \pi \min\left\{(1+r_L)L; (1-\phi)V_H^L + E - (1+r_L)L\right\} \\
&\quad + (1-\pi)\min\left\{(1+r_L)L; (1-\phi)V_L^L + E - (1+r_L)L\right\} \\
&\quad - (1+r_D)L \\
&= \pi \min\left\{(1+r_L)L; (1-\phi)V_H^L + E - (1+r_L)L\right\} \\
&\quad - (1+r_D)L,
\end{aligned}
\tag{10.34}
$$

where for the final equality we made use of our assumption that $(1-\phi)V_H^L + E > 2(1+r_L)L > (1-\phi)V_L^L + E$. We note that the resources available to repay the second loan have been reduced by the repayment of the first loan. As banks have no information whether their loan is the first or second loan, the loan rate they will apply must be identical. The second loan on a large investment is not guaranteed to be repaid and we can distinguish two cases, $(1+r_L)L \leq (1-\phi)V_H^L + E - (1+r_L)L$ and $(1+r_L)L > (1-\phi)V_H^L + E - (1+r_L)L$. If we define $\phi_0 = \frac{V_H^L + E - 2(1+r_L)L}{V_H^L}$, we can rewrite equation (10.34) as

$$\hat{\Pi}_B^2 = \begin{cases} \pi\left(1+r_L\right)L-\left(1+r_D\right)L & \text{if } \phi \leq \phi_0 \\ \pi\left(\left(1-\phi\right)V_H^L+E-\left(1+r_L\right)L\right) & \text{if } \phi > \phi_0 \\ -\left(1+r_D\right)L & \end{cases}. \tag{10.35}$$

With banks equally likely to provide the first and second loan, the expected profits of the bank is given by $\hat{\Pi}_B^L = \frac{1}{2}\hat{\Pi}_B^1 + \frac{1}{2}\hat{\Pi}_B^2$, where perfect competition implies that $\hat{\Pi}_B^L = 0$. Hence we have the loan rate given by

$$1+r_L = \begin{cases} 2\frac{1+r_D}{1+\pi} & \text{if } \phi \leq \phi_0 \\ \frac{2(1+r_D)L-\pi(1-\phi)V_H^L-\pi E}{(1-\pi)L} & \text{if } \phi > \phi_0 \end{cases}, \tag{10.36}$$

where after inserting this expression for r_L, we get that $\phi_0 = \frac{(1+\pi)V_H^L+(1+\pi)E-4(1+r_D)L}{(1+\pi)V_H^L}$.

We can now determine the company profits if conducting the small and large investments, respectively. If conducting the small investment, the company will retain the investment outcome and equity, which it initially invested, after repaying the loan in full; if the loan cannot be repaid, the company will lose its equity and obtain no benefits. Thus the profits are given by

$$\begin{aligned} \Pi_C^1 &= \pi \max\left\{V_H^S + E - \left(1+r_L^S\right)L; 0\right\} \\ &\quad + (1-\pi)\pi \max\left\{V_L^S + E - \left(1+r_L^S\right)L; 0\right\} - E \\ &= \overline{V}_S - \left(1+r_L^S\right)L, \end{aligned} \tag{10.37}$$

where for the final equality we made use of our assumption that $V_L^S + E > (1+r_L)L$ and define $\overline{V}_S = \pi V_H^S + (1-\pi)V_L^S$ for convenience as the expected outcome of the small investment.

The small investment will only be feasible if they are profitable, thus $\Pi_C^1 \geq 0$. This is the case if

$$\begin{cases} \pi \geq -\frac{1}{2}\frac{V_H^S}{V_H^S-V_L^S} + \sqrt{\frac{2(1+r_D)L-V_L^S}{V_H^S-V_L^S} - \frac{1}{4}\left(\frac{V_H^S}{V_H^S-V_L^S}\right)^2} & \text{if } \phi \leq \phi_0 \\ \phi \leq \frac{(1-\pi)+\pi\left(V\right)H^L+E\big)-2(1+r_D)L}{V_H^L} & \text{if } \phi > \phi_0 \end{cases} \tag{10.38}$$

It is thus that in situations where the success rates are sufficiently low and the private benefits sufficiently high, the loan rate has to increase so far to account for the potential losses from lending to companies with the large investment, that small investments are not profitable anymore. Thus the existence of large investments can crowd out all investments, including otherwise feasible small investments.

If conducting the large investment, companies obtain their private benefits, ϕV_H^j, in addition to any profits from the investment after both loans have been repaid. Hence we have

$$
\begin{aligned}
\Pi_C^2 &= \pi\left(\phi V_H^L + \max\left\{(1-\phi)V_H^L + E - 2(1+r_L)L; 0\right\}\right) \\
&\quad + (1-\pi)\left(\phi V_L^L + \max\left\{(1-\phi)V_L^L + E - 2(1+r_L)L; 0\right\}\right) \\
&\quad - E \\
&= \phi\overline{V}_L + \pi\max\left\{(1-\phi)V_H^L + E - 2(1+r_L)L; 0\right\} - E,
\end{aligned} \tag{10.39}
$$

where for the final equality we made use of our assumption that $(1-\phi)V_H^L + E > 2(1+r_L)L > (1-\phi)V_L^L + E$ and define $\overline{V}_L = \pi V_H^L + (1\pi)V_L^L$ for convenience as the expected outcome of the large investment. We can rewrite this expression as

$$
\Pi_C^2 = \begin{cases} \phi(1-\pi)V_L^L - (1-\pi)E - 2(1+r_L)L & \text{if } \phi \leq \phi_0 \\ \phi\overline{V}_L - E & \text{if } \phi > \phi_0 \end{cases}. \tag{10.40}
$$

The company would prefer the large investment over the small investment if this is more profitable, $\Pi_C^2 \geq \Pi_C^1$. Inserting the loan rate from equation (10.36) into the respective profits of equations (10.37) and (10.40), this requirement solves for

$$
\phi \geq \phi^* = \begin{cases} \frac{(1+\pi)\overline{V}_S + \left(1-\pi^2\right)E + 2(1+r_D)L}{\left(1-\pi^2\right)V_L^L} & \text{if } \phi \leq \phi_0 \\ \frac{(1-\pi)\overline{V}_S + \pi V_H^L + E - 2(1+r_D)L}{(1-\pi)V_L + \pi V_H^L} & \text{if } \phi > \phi_0 \end{cases}. \tag{10.41}
$$

If the private benefits of the large investment are sufficiently large, the company will seek this investment. The reason the private benefits need to be high is due to the large investment being less efficient and despite requiring a loan that is twice the size of the small investment, produces outcomes that are less than twice the size of the small investment. This will reduce the profits of the company from making this investment, which can only be compensated for if the private benefits of sufficient size, which they can retain regardless of the outcome of the investment, can be retained.

Of course, for companies to demand loans for such a large investment, we do not require it to be more attractive than the small investment, but the profits of this large investment have to be positive, too. Thus we require $\Pi_C^2 \geq 0$, which noting in the company profits as represented in equation (10.39) that the second term cannot be negative, easily becomes

$$
\phi \geq \phi^{**} = \begin{cases} \frac{\left(1-\pi^2\right)E + 4(1+r_D)L}{\left(1-\pi^2\right)V_L^L} & \text{if } \phi \leq \phi_0 \\ \frac{E}{\overline{V}_L} & \text{if } \phi > \phi_0 \end{cases}. \tag{10.42}
$$

Large investments with small private benefits might not be generating profits to the company as the high loan rate and substantial

Thus, if the fraction of private benefits ϕ is sufficiently high by exceeding both thresholds, ϕ^* and ϕ^{**}, companies will prefer to conduct the large investment. We can now compare this result with a situation in which banks disclose the fact the company has already applied for a loan to a credit reference agency.

With information disclosure If banks disclose the fact that a company has already obtained a loan, or has applied for a loan, the bank approached subsequently by the company knows that its loan would be the second loan and the company seeks to conduct the large investment. If no such information is available, the bank knows that the company either does not seek to conduct the large investment or is the first bank to provide a loan for a large investment.

Banks would generate the same profits regardless of whether the company seeks a small investment or it is the first bank financing a large investment as we can see from equations (10.32) and (10.33) that $\Pi_B^S = \hat{\Pi}_B^1 = \left(1 + r_L^1\right) L - (1 + r_D) L$. If banks are in perfect competition such that $\Pi_B^S = \hat{\Pi}_B^1 = 0$ the loan rate is set such that $r_L^1 = r_D$. Inserting this loan rate into the profits of the company pursuing the small investment in equation (10.37), we get the profits of the company for this small investment given by

$$\Pi_C^S = \overline{V}_S - (1 + r_D) L > 0 \tag{10.43}$$

and the small investment is always feasible.

For companies pursuing the large investment, the profits of the bank providing the second loan are given by equation (10.35) and as banks know they provide the second loan, they would seek to break even on this loan, requiring $\hat{\Pi}_B^2 = 0$ in perfect competition, which solves for

$$1 + r_L^2 = \begin{cases} \frac{1+r_D}{\pi} & \text{if } \phi \leq \phi_0 \\ \frac{1+r_D}{\pi} - \frac{(1-\phi)V_H^L+E}{L} & \text{if } \phi > \phi_0 \end{cases}, \tag{10.44}$$

where inserting r_L^2 for r_L we get that $\hat{\phi}_0 = \frac{\pi V_H^L + \pi E - 2(1+r_D)L}{\pi V_H^L}$.

The company seeking two loans for the large investment will pay different loan rates for each loan as banks know whether they are providing the first or second loan, and its profits are given similar to equation (10.39) by

$$\begin{aligned} \Pi_C^L = \phi\overline{V}_L + \pi \max \Big\{ &(1 - \phi) V_H^L + E - \left(1 + r_L^1\right) L \\ &- \left(1 + r_L^2\right) L; 0 \Big\} - E. \end{aligned} \tag{10.45}$$

Inserting the loan rate $r_L^1 = r_D$ and for r_L^2 from equation (10.44), we get these profits as

$$\Pi_C^L = \begin{cases} \phi\overline{V}_L - E & \text{if } \phi \leq \hat{\phi}_0 \\ \phi \left(V_L^L - \pi V_H^L\right) + 2\pi V_H^L + (2\pi - 1) E & \text{if } \phi > \hat{\phi}_0 \\ - (1 + \pi)(1 + r_D) L & \end{cases}, \tag{10.46}$$

where $\hat{\phi}_0 = \frac{2\pi V_H^L + 2\pi E - (1+\pi)(1+r_D)L}{2\pi V_H^L}$. In order to obtain this result, we carefully had to evaluate the cases of different loan rates for $\phi \leq \phi_0$ and $\phi > \phi_0$ as well as the cases of $(1 - \phi) V_H^L + E - \left(1 + r_L^1\right) L - \left(1 + r_L^2\right) L \geq 0$ and $(1 - \phi) V_H^L + E - \left(1 + r_L^1\right) L - \left(1 + r_L^2\right) L < 0$.

Companies will choose the large investment over the small investment if its profits are higher, $\Pi_C^L \geq \Pi_C^S$, which using equations (10.43) and (10.46) gives us

$$\phi \geq \hat{\phi}^* = \begin{cases} \frac{E}{\overline{V}_L} & \text{if } \phi \leq \hat{\phi}_0 \\ \frac{(1+\pi)(1+r_D)L-2\pi V_H^L-(2\pi-1)E}{V_L^L-\pi V_H^L} & \text{if } \phi > \hat{\phi}_0 \end{cases}. \tag{10.47}$$

Of course demand for large investments is only present if it is profitable to do so, hence we require that $\Pi_C^L \geq 0$, or

$$\phi \geq \hat{\phi}^{**} = \begin{cases} \frac{V_S+E-(1+r_D)L}{\overline{V}_L} & \text{if } \phi \leq \hat{\phi}_0 \\ \frac{((1+\pi)(1+r_D)L-2\pi V_H^L-(2\pi-1)E}{V_L^L-\pi V_H^L} & \text{if } \phi > \hat{\phi}_0 \end{cases}. \tag{10.48}$$

Again, if the fraction of private benefits, ϕ, is sufficiently high by exceeding both thresholds, $\hat{\phi}^*$ and $\hat{\phi}^{**}$, companies will prefer to conduct the large investment.

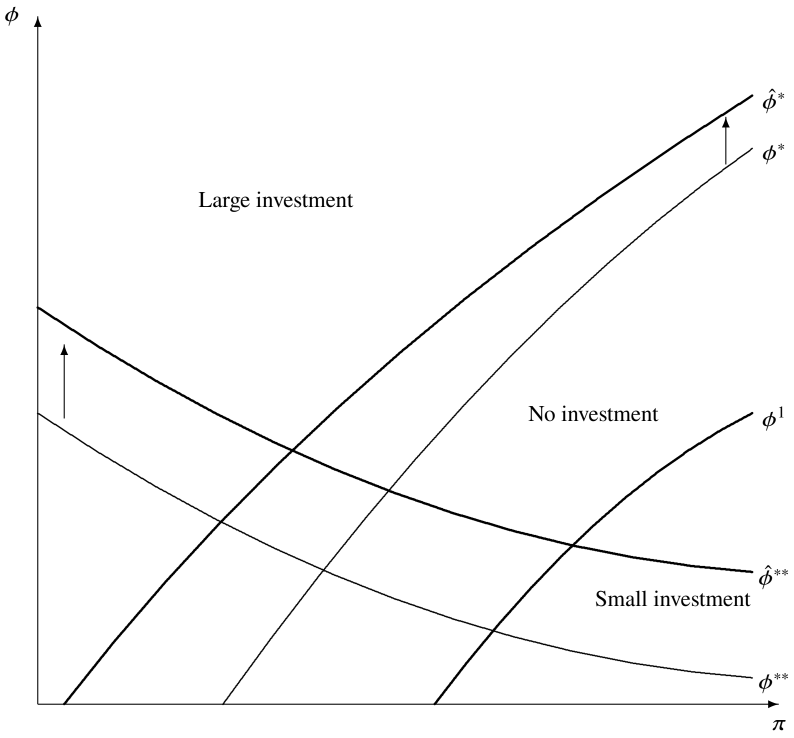


Fig. 10.1: Investment choice with and without information disclosure

We can now compare the result with and without disclosure of the fact that a loan has been granted or applied for. In figure 10.1 we illustrate the resulting constraints without information disclosure (thin lines) and with information disclosure (thick lines). We see that in most cases the constraints on the possibility of large investments become more stringent, leading the area between these lines to become unsustainable for large investments. In addition, small investments are no longer crowded out with information disclosure as the loan rate for small loan will reflect their low-risk status. The reason is that without information disclosure, the bank does not know whether its provides the first or second loan; it will thus offer a loan rate that takes into account that the loan might be the first loan and will be repaid with certainty or the second loan that might not be repaid. This will lead to a loan rate that is between that of a loan were the bank to know it is the first (or only) bank providing a loan and that it would offer if it knew it was providing the second loan. This second loan will be more expensive as it includes the risk of the loan not being repaid. This higher loan rate in the case of information disclosure makes the loan more expensive and hence less attractive than only making the small investment, but may make the large investment overall unsustainable. Hence the disclosure of information in most cases reduces the scope for the undesirable large investment.

Summary The disclosure of information on the existence of loans, or whether loans have been applied for, can reduce the moral hazard of companies choosing to secure additional loans to conduct larger but also more risky investments, that have increased benefits to the company that reduce the ability to repay the loan. Similarly, the ability of companies to conduct low-risk investments is always maintained. The ability to discriminate between loans for small and loans for large investments allows banks to charge loan rates that accurately reflect the risks these loans entail. If such information was absent, the bank would have to charge an average loan rate, making the loan too cheap for larger investments and thus encouraging this additional risk-taking by companies. At the same time, those companies that make small investments will pay a too high loan rate, giving additional incentives to conduct large investments. It is thus that information disclosure about the existence of loans allows banks to price loans more precisely in line with the risks taken and thereby reduces incentives for risk-taking.

The disclosure of information about existing loans or loan applications is particularly helpful where there are strong incentives for companies to divert funds to their own benefit and make more risky investments. This might be in situations where corporate governance structures are weakly established, such as in newly emerging industries or when informal agreements are common. Loans to individuals can also be subject to moral hazard with individuals uusing the proceeds of the loan for purposes not disclosed and thereby jeopardising the ability to repay the loan.

Reading Bennardo, Pagano, & Piccolo (2015)

10.3 Information disclosure and competition

Banks routinely provide information about the companies they provide loans to through credit reference agencies. Such information is used by other banks to assess the risks of a company and allows them to judge whether to provide a loan themselves and if so what the applicable loan rate would be to take into account any risks. The provision of information to competitors will have an impact on the informational advantage a bank might have from their interaction with a company and thus affect the competitive outcomes in providing future loans.

Let us assume that there are two types of companies, one type of companies makes a successful investment with probability π, while the other type of company will never be able make an investment successful in the sense that it generates revenue that can be used to repay the loan. We can thus interpret such a company as not creditworthy as they only have access to investments that are not sufficiently profitable to repay the loan granted.

If an investment is not successful, it generates no revenue and hence the loan L used to finance this investment cannot be repaid. The company knows their type, but even if they cannot be successful. would seek a loan as there might be other non-pecuniary benefits associated with making the investment. However, banks initially do not know the type of company and only after having lent to them, will the type be revealed to them. To other banks, this information remains unknown, unless the bank having provided the loan in the first instance, is providing them with such information.

For simplicity, companies seek loans for identical investments in two time periods, and a failure to repay the loan after the first time period is not affecting their ability to obtain a loan in the second time period. Companies are not restricted to obtain a loan in the second time period from the same bank that has provided them with a loan in the first time period and will always seek a lon from the bank that offers them the lowest loan rate.

Banks initially do not know the type of company they lend to, but after lending learn its type for the lending decision in the second time period. If the type of company is revealed as being unable to generate successful investments, the loan would never be repaid and hence the bank would not provide a further loan to this company as to avoid a certain loss. This implies that in the first time period the bank provides a loan to all companies and the loan is repaid only by those companies that can generate successful investments, ν, if they are indeed successful, π. In the second time period, this bank will only lend to the fraction ν of companies that have been revealed as being able to generate successful investments, who then repay the loan with probability π. Denoting the loan rates banks charge for each time period t by r_L^t and assuming the loan is fully financed by deposits on which interest r_D is payable, the profits of the bank in the first and second time period, respectively, are then given by

$$\Pi_B^1 = \nu\pi\left(1 + r_L^1\right)L - (1 + r_D)L, \tag{10.49}$$
$$\Pi_B^2 = \nu\left(\pi\left(1 + r_L^2\right)L - (1 + r_D)L\right).$$

The complete profits of this bank are the sum of the profits from each time period, $\Pi_B = \Pi_B^1 + \Pi_B^2$.

Those banks that have not previously lent to the company, but do so only in the second time period, will have to make inferences on the likelihood that the company is able to generate successful investments, denoted $\hat{\nu}$ and its profits are then given by

$$\hat{\Pi}_B^2 = \hat{\nu}\pi\left(1 + \hat{r}_L^2\right)L - (1 + r_D)L, \tag{10.50}$$

where these banks charge a loan rate $\hat{r}_L^2$ and we assume they face the same deposit rate.

We will only consider the company that is able to generate successful investments as the other type of company will be indifferent between all loan offers, knowing that it will not have to repay the loan. The company has identical investment opportunities in each time period, where they obtain a return R if the investment is successful and then repay the loan. Hence we have the company profits for each time period given by

$$\Pi_C^t = \pi\left((1 + R)L - \left(1 + r_L^t\right)L\right) \tag{10.51}$$

and the complete profits are the sum of the profits from both time periods, $\Pi_C = \Pi_C^1 + \Pi_C^2$.

We will now investigate the loan rates and bank profits under different degrees of information disclosure. The bank providing a loan in the first time period may not share any information about the company, share information on the type of company, or share information only about the fact that a company has not repaid their loan.

No information disclosure Let us start by assuming that banks do not share any information about the type of company after the first time period. Thus the bank not having lent to the company in time period 1 will have no opportunity make additional inferences about the likelihood of it being able to generate a successful investment, implying that $\hat{\nu} = \nu$.

We start by analysing the provision of loans in time period 2. The banks will only provide loan in time period 2 if this is profitable to do so, hence if $\Pi_B^2 \geq 0$ and $\hat{\Pi}_B^2 \geq 0$. Using equations (10.49) and (10.50), this easily becomes

$$1 + r_L^2 \geq \frac{1 + r_D}{\pi}, \tag{10.52}$$
$$1 + \hat{r}_L^2 \geq \frac{1 + r_D}{\nu\pi},$$

for the bank lending in time period 1 and a bank only providing a loan in time period 2, respectively. We see that the bank lending in time period 1 can offer a lower loan rate; this is because it now has knowledge of the type of company, knowing it is able to generate a successful investment, which reduces the risks the bank faces. The highest the bank having lent in time period 1 will offer is $\frac{1+r_D}{\nu\pi}$, as competition with the banks not having lent before would not allow this bank to provide the loan at a higher loan rate. The loan rate the company is offered is thus in the range

$$\frac{1+r_D}{\nu\pi} \geq 1+r_L^2 \geq \frac{1+r_D}{\pi}. \tag{10.53}$$

We note that the loan in the second time period will be provided by the company that provided them with a loan in time period 1 as this company can undercut any other bank and will do so marginally only to maximize its profits. The company will accept this loan only if it is profitable to do so. With its profits in time period 2 given by equation (10.51), we see that $\Pi_C^2 \geq 0$ requires that $1+r_L^2 \leq 1+R$. Combining this requirement with the condition for bank profitability in equation (10.53) and noting that the bank would charge the highest possible loan rate, we have

$$1+r_L^2 = \min\left\{\frac{1+r_D}{\nu\pi}; 1+R\right\}. \tag{10.54}$$

We can now distinguish two cases; the first case will be that $\frac{1+r_D}{\nu\pi} \geq 1+R$, or $\pi \leq \frac{1}{\nu}\frac{1+r_D}{1+R}$. In this case we have $1+r_L^2 = 1+R$ and we easily see that $\Pi_C^2 = 0$. Companies will only request a loan in time period 1 if they make profits overall, thus we require that $\Pi_C \geq 0$. Inserting for Π_C^2, we easily see that this requires that $1+r_L^1 \leq 1+R$.

The bank will make profits in time period 2 of $\Pi_B^2 = \nu\left(\pi\left(1+R\right)L - \left(1+r_D\right)L\right)$, giving us $\Pi_B = \nu\pi\left(1+r_L^2\right)L_\pi\left(1+R\right)L - \left(1+\nu\right)\left(1+r_D\right)L$. Banks will only provide loans if it is profitable to do so, thus we need to ensure that $\Pi_B \geq 0$, which solves for $1+r_L^1 = \frac{1+\nu}{\nu\pi}\left(1+r_D\right) - \left(1+R\right)$. Combining this with the requirement of companies that $1+r_L^1 \leq 1+R$, we obtain that loan rates have to be in the range of $1+R \geq 1+r_L^1 \geq \frac{1+\nu}{\nu\pi}\left(1+r_D\right) - \left(1+R\right)$. A feasible solution exists only if $1+R \geq \frac{1+\nu}{\nu\pi}\left(1+r_D\right) - \left(1+R\right)$, or $\pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$. Combining this with the initial condition for our case, we obtain that a bank loan can be provided if $\frac{1}{\nu}\frac{1+r_D}{1+R} \geq \pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$.

Competition between banks will require them to offer the lowest feasible loan rate and we have the loan rates for times periods 1 and 2, respectively, give as

$$\begin{aligned} 1+r_L^1 &= \frac{1+\nu}{\nu\pi}\left(1+r_D\right) - \left(1+R\right), \\ 1+r_L^2 &= 1+R, \end{aligned} \tag{10.55}$$

provided $\frac{1}{\nu}\frac{1+r_D}{1+R} \geq \pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$.

The second case considers that $\frac{1+r_D}{\nu\pi} < 1+R$, or $\pi > \frac{1}{\nu}\frac{1+r_D}{1+R}$. In this case we have that $1+r_L^2 = \frac{1+r_D}{\nu\pi}$ and hence $\Pi_C^2 = \pi\left(\left(1+R\right)L - \frac{1+r_D}{\nu\pi}L\right)$, such that

$\Pi_C = 2\pi(1+R)L - \pi\left(1+r_L^1\right)L - \frac{1+r_D}{\nu}L$. In order for the company to accept a loan, we need to ensure that $\Pi_C \geq 0$, with solves for $1+r_L^1 \leq 2(1+R) - \frac{1+r_D}{\nu\pi}$.

The bank makes profits of $\Pi_B^2 = (1-\nu)(1+r_D)L$ and this gives us aggregate profits of $\Pi_B = \nu\pi\left(1+r_L^1\right)L - \nu(1+r_D)L$. For banks to be willing to provide the initial loan, they need to be able to produce profits, $\Pi_B \geq 0$, which solves for $1+r_L^1 \geq \frac{1+r_D}{\pi}$. Combining this requirement with that of companies, we obtain $2(1+R) - \frac{1+r_D}{\nu\pi} \geq 1+r_L^1 \geq \frac{1+r_D}{\pi}$. This provides a feasible solution if $2(1+R) - \frac{1+r_D}{\nu\pi} \geq \frac{1+r_D}{\pi}$, or $\pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$. This condition is less strict than the restriction for this second case, $\pi > \frac{1}{\nu}\frac{1+r_D}{1+R}$ and hence does not provide an additional constraint.

Competition between banks will require banks to offer the lowest feasible loan rate and we have the loan rates for times periods 1 and 2, respectively, give as

$$\begin{aligned} 1+r_L^1 &= \frac{1+r_D}{\pi}, \\ 1+r_L^2 &= \frac{1+r_D}{\nu\pi} \end{aligned} \tag{10.56}$$

if $\pi > \frac{1}{\nu}\frac{1+r_D}{1+R}$.

Based on the results in equations (10.55) and (10.56) for the two cases, we see that for companies with $\pi < \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$ no loan is provided. In both cases, the loan rates in the first time period is below that of the second time period, despite the bank now having full knowledge of the type of company they are lending to, reducing their risk. The reason is that the bank can exploit their informational advantage and offer loan rates that are profitable to them. These profits are then used to subsidize the loan rates in the first time period; this is done to compete with other banks for this initial loan that allows them to gain the informational advantage in the second time period. Figure 10.2 illustrates the loan rates for different success rates of the company investments. We see that for $\pi < \frac{1}{\nu}1+r_D1+R$ the loan rate applied in the second time period is $1+R$ and the company makes no profits. In this case, banks without information on the type of company would have to charge a loan rate too high for the company to accept in order to be compensated for the risks they are taking. However, the informational advantage of the initial bank allows it to offer a lower loan rate that extracts all surplus from the company and still generate profits. This, of course, reduces the bank's profits from time period 2, giving it less opportunity to subsidize the loan rate in time period 1, which therefore has to increase more than would be justified by the increased risks from a falling success rate. At $\pi = \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$ the higher risk requires the bank to charge a loan rate of $1+R$ also in time period 1, making any cross-subsidies from time period 2 to time period 1 impossible.

While perfect competition between banks ensures that bank profits are eliminated, the profits of companies are positive. Inserting the loan rates from equations (10.55) and (10.56), we easily obtain that

$$\Pi_C = \begin{cases} 2\pi(1+R)L - \frac{1+\nu}{\nu}(1+r_D) & \text{if } \pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \\ 0 & \text{if } \pi < \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \end{cases}. \tag{10.57}$$

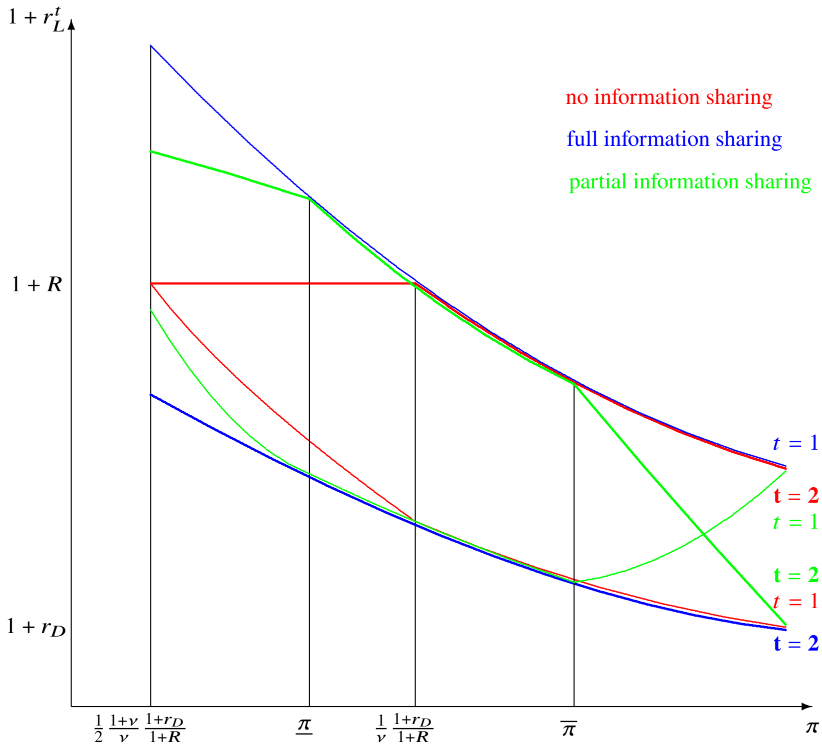


Fig. 10.2: Equilibrium loan rates with different levels of information disclosure

Having established the equilibrium without the disclosure of information, we can now proceed to evaluate the situation if the bank lending in time period 1 discloses information to their competitors through credit reference agencies. In a first step we will assume that the bank discloses the type of company, thus the full information they are holding.

Full information disclosure After having established the equilibrium loan rates if information is not disclosed, we can now consider the implications of the bank lending in time period 1 fully disclosing the information they hold. Thus they would disclose the type of company to the other banks. Thus these banks would now only lend to those companies that are able to generate successful investments. Hence their inferences about the likelihood of it being able to generate a successful investment becomes $\hat{\nu} = 1$ for time period 2.

As all banks have the same amount of information, they will face the same profits, thus $\Pi_B^2 = \hat{\Pi}_B^2 = \pi\left(1 + r_L^2\right) L - (1 + r_D) L$ and perfect competition requires that $\Pi_B^2 = \hat{\Pi}_B^2 = 0$, implying that

$$1 + r_L^2 = \frac{1 + r_D}{\pi}. \tag{10.58}$$

Companies will only take out a loan in the second time period if their profits are positive, thus we require $\Pi_C^2 = \pi \left((1 + R) L - \left(1 + r_L^2\right) L \right) = \pi (1 + R) L - (1 + r_D) L \geq 0$, from which we obtain $\pi \geq \frac{1+r_D}{1+R}$.

Similarly, banks compete in time period 1 to to attract companies. As $\Pi_B = \Pi_B^1 + \Pi_B^2 = \Pi_B^1 = 0$ due to the perfect competition in time period 2, we get

$$1 + r_L^1 = \frac{1 + r_D}{\nu \pi}, \tag{10.59}$$

recognising that banks face the additional uncertainty in time period 1 of not yet knowing the type of company they are lending to. We thus see that the loan rate in time period 1 is higher than in time period 2, accounting for this additional risk. This result is in contrast to the loan rates in the case that no information is shared between banks, where the loan rate in time period 1 was lower than in time period 2. The reason for this difference is that banks who have lent to the company in time period 1 are not able to make profits from lending in time period 2, given their informational advantage is eliminated, and they can therefore not subsidise the loan rate in period 1 to provide the first loan. It is, however, possible that loan rates in time period 1 exceed the return of investments of companies, $r_L^2 > R$, and companies thus making a loss from the initial investment. This is compensated by banks charging a lower loan rate in time period 2, allowing them to make a profits once the banks have learned their type, and thus generating profits that compensate for the losses in time period 1.

For companies to request loans in time period 1 we require that $\Pi_C \geq 0$, which after inserting the loan rates from equations (10.58) and (10.58) becomes

$$\pi \geq \frac{1}{2} \frac{1 + \nu}{\nu} \frac{1 + r_D}{1 + R}. \tag{10.60}$$

This condition is more stringent than the condition that $\pi \geq \frac{1+r_D}{1+R}$ for companies to seek a loan in time period 2, and hence a loan will be provided if this condition is fulfilled. Companies that have a success rate below this threshold will not demand loans as they cannot make profits from their investment due to the loan rate being too high to compensate banks for their risk.

In figure 10.2 this result has been included and we see that while the loan rates for higher success rates seem identical, the time periods are reversed and for lower success rates the loan rates are higher for time period 1 as no subsidy from time period 2 can be given and the loan rate in time period 2 is lower as no such subsidy needs to be charged to borrowers in that time period.

The company profits can easily be obtained by inserting from equations (10.58) and (10.59), such that we obtain

$$\Pi_C = \begin{cases} 2\pi (1 + R) L - \frac{1+\nu}{\nu} (1 + r_D) & \text{if } \pi \geq \frac{1}{2} \frac{1+\nu}{\nu} \frac{1+r_D}{1+R} \\ 0 & \text{if } \pi < \frac{1}{2} \frac{1+\nu}{\nu} \frac{1+r_D}{1+R} \end{cases}. \tag{10.61}$$

Comparing this result with the case of banks not sharing their information about the company from equation (10.56), we see that companies are indifferent between banks disclosing information about their type or not providing any information. Both instances give companies the same profits. As banks are assumed to be competitive in both cases, they make no profits and are therefore also indifferent between sharing and not sharing information on the type of company they have been lending to.

Assuming that banks disclose their full information about a company to competitors is not realistic. Banks, however, commonly disclose some information to their competitors through credit reference agencies. Most notably, they provide information on past failures of companies; such a more realistic scenario we will assess next.

Partial information disclosure Banks may not disclose the type of company they are lending to, but instead they may disclose whether a company has repaid their loan or not. If the bank reports that the company has repaid its loan, it is obvious that it is a company that is able to generate successful investments and hence the other banks can infer for the second time period that in this case $\hat{\nu}_N = 1$. Companies not repaying their loan cannot be readily assigned a type as this might be due to companies not being able to generate successful investments or they are able to generate such investments, but have not been successful this time. Bayesian learning allows banks to update their beliefs about the likelihood of the company being able to generate successful investments. Acknowledging that the prior belief of such banks on the likelihood of companies being able to generate successful investments is ν and the probability of a success being π, we obtain the new belief as

$$\hat{\nu}_D = \frac{\nu(1-\pi)}{\nu(1-\pi)+(1-\nu)}. \tag{10.62}$$

The numerator represents the likelihood that a company is able to generate successful investments, ν, but defaults, $1-\pi$, and the denominator the likelihood of observing a default, which consists of the company being able to generate successful investments but failing in the first time period, in addition to the company not being able to generate successful investments at all.

We start again by analysing the provision of loans in time period 2. The banks will only provide loan in time period 2 if this is profitable to do so, hence if $\Pi_B^2 \geq 0$ and $\hat{\Pi}_B^2 \geq 0$. Using equations (10.49) and (10.50), this easily becomes

$$\begin{aligned} 1 + r_L^2 &\geq \frac{1+r_D}{\pi}, \\ 1 + \hat{r}_L^2 &\geq \frac{1+r_D}{\hat{\nu}_D \pi}, \\ 1 + \hat{r}_L^2 &\geq \frac{1+r_D}{\hat{\nu}_N \pi}, \end{aligned} \tag{10.63}$$

for the bank lending in time period 1 and a bank only providing a loan in time period 2, after default in time period 1 and no default in time period 1, respectively. We see that the bank lending in time period 1 can offer a lower loan rate for the company has defaulted and the same loan rate for companies that do not default; this is because it now has knowledge of the type of company, thus knowing it is able to generate a successful investment and only in case a company does not default does the other bank know this. The highest the bank having lent in time period 1 will request is $\frac{1+r_D}{\hat{\nu}_i \pi}$, as competition with the banks not having lent before, would not allow this bank to provide the loan at any lower loan rate. The loan rate the company is offered is thus in the range

$$\frac{1+r_D}{\hat{\nu}_i \pi} \geq 1 + r_L^2 \geq \frac{1+r_D}{\pi}. \tag{10.64}$$

We note that the loan in the second time period will be provided by the bank that provided them with a loan in time period 1. The company will accept this loan only if it is profitable to do so. With its profits in time period 2 given by equation (10.51), we see that $\Pi_C^2 \geq 0$ requires that $1 + r_L^2 \leq 1 + R$. Combining this requirement with the condition for bank profitability in equation (10.64) and noting that the bank would charge the highest possible loan rate, we have

$$1 + r_L^2 = \begin{cases} \min\left\{\frac{1+r_D}{\pi}; 1+R\right\} & \text{if no default in t=1} \\ \min\left\{\frac{1+r_D}{\hat{\nu}_D \pi}; 1+R\right\} & \text{if default in t=1} \end{cases}. \tag{10.65}$$

As the case that $\frac{1+r_D}{\pi} > 1 + R$ is ruled out below, we can focus on comparing $\frac{1+r_D}{\hat{\nu}_D \pi}$ and $1 + R$. Let us first consider the case that $\frac{1+r_D}{\hat{\nu}_D \pi} \leq 1 + R$, implying that after inserting from equation (10.62) for $\hat{\nu}_D$, we have $1 + r_L^2 = \frac{1-\nu\pi}{\nu\pi(1-\pi)}(1 + r_D)$. If the company defaults in time period 1, this loan rate is chosen by the bank and if the company does not default in time period 1, the bank chooses $1 + r_L^2 = \frac{1+r_D}{\pi}$ as indicated in equation (10.65). Thus the expected loan rate is given by

$$\begin{aligned} E\left[1 + r_L^2\right] &= \pi \frac{1+r_D}{\pi} + (1-\pi)\frac{1-\nu\pi}{\nu\pi(1-\pi)}(1+r_D) \\ &= \frac{1+r_D}{\nu\pi}. \end{aligned} \tag{10.66}$$

Our condition for the case that $\frac{1+r_D}{\hat{\nu}_D \pi} \leq 1+R$ can be solved for $\left(\pi - \frac{1}{2}\left(1 + \frac{1+r_D}{1+R}\right)\right)^2 \leq \frac{1}{4}\left(1 + \frac{1+r_D}{1+R}\right)^2 - \frac{1}{\nu}\frac{1+r_D}{1+R}$ and hence the results are valid for $\pi \in \left[\underline{\pi}; \overline{\pi}\right]$, where $\underline{\pi}$ and $\overline{\pi}$ are given by solving the inequality as an equality.

The second case requires $\frac{1+r_D}{\hat{\nu}_D \pi} > 1 + R$ and hence from equation (10.65) we see that $1 + r_L^2 = 1 + R$. This now gives us an expected loan rate of

$$E\left[1 + r_L^2\right] = \pi \frac{1+r_D}{\pi} + (1-\pi)(1+R) = (1+r_D) + (1-\pi)(1+R). \tag{10.67}$$

The condition $\frac{1+r_D}{\hat{\nu}_D \pi} > 1 + R$ implied that this applies to $\pi \notin \left[\underline{\pi}; \overline{\pi}\right]$.

The loan rate in the second time period therefore is given by

$$E\left[1+r_L^2\right]=\begin{cases}\frac{1+r_D}{\nu\pi} & \text{if } \pi\in\left[\underline{\pi};\overline{\pi}\right]\\ \left(1+r_D\right)+\left(1-\pi\right)\left(1+R\right) & \text{if } \pi\notin\left[\underline{\pi};\overline{\pi}\right]\end{cases}. \tag{10.68}$$

The expected profits of the bank in time period 2 are from equation (10.68) given by $\Pi_B^2=\nu\left(\pi E\left[1+r_L^2\right]-\left(1+r_D\right)\right)L$, which after inserting from equation (10.68) becomes

$$\Pi_B^2=\begin{cases}\left(1-\nu\right)\left(1+r_D\right)L & \text{if } \pi\in\left[\underline{\pi};\overline{\pi}\right]\\ \nu\left(1-\pi\right)\left(\pi\left(1+R\right)-\left(1+r_D\right)\right)L & \text{if } \pi\notin\left[\underline{\pi};\overline{\pi}\right]\end{cases}. \tag{10.69}$$

The bank profits for the entire two time periods are then given by $\Pi_B=\Pi_B^1+\Pi_B^2$, which after inserting from equation (10.49) and (10.69) becomes

$$\Pi_B=\begin{cases}\nu\pi\left(1+r_L^1\right)L-\nu\left(1+r_D\right)L & \text{if } \pi\in\left[\underline{\pi};\overline{\pi}\right]\\ \nu\pi\left(\left(1+r_L^1\right)\left(1-\pi\right)\left(1+R\right)\right)L & \text{if } \pi\notin\left[\underline{\pi};\overline{\pi}\right]\\ \quad-\left(1+\nu\left(1-\pi\right)\right)\left(1+r_D\right)L & \end{cases}. \tag{10.70}$$

If banks are competitive, we require that $\Pi_B=0$, which solves for

$$1+r_L^1=\begin{cases}\frac{1+r_D}{\pi} & \text{if } \pi\in\left[\underline{\pi};\overline{\pi}\right]\\ \frac{1+\nu(1-\pi)}{\nu\pi}\left(1+r_D\right)-\left(1-\pi\right)\left(1+R\right) & \text{if } \pi\notin\left[\underline{\pi};\overline{\pi}\right]\end{cases}. \tag{10.71}$$

The profits of companies are given by equation (10.51) and after inserting from equations (10.65) and (10.71), we easily obtain that $\Pi_C=2\pi\left(1+R\right)L-\frac{1+\nu}{\nu}\left(1+r_D\right)L$. Of course, in order to demand loans, companies need to make profits, which requires $\Pi_C\geq 0$, implying $\pi\geq\frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}$. As this constraint is more binding that $\pi\geq\frac{1+r_D}{1+R}$, we see that in equation (10.65) we obtain for non-defaulting companies that $1+r_L^2=\frac{1+r_D}{\pi}$, as indicated at the time.

With company profits thus given by

$$\Pi_C=\begin{cases}2\pi\left(1+R\right)L-\frac{1+\nu}{\nu}\left(1+r_D\right) & \text{if } \pi\geq\frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}\\ 0 & \text{if } \pi<\frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}\end{cases}, \tag{10.72}$$

as in the case of no or full disclosure of information to other banks, companies are indifferent between the level of information sharing; similarly as banks are not making any profits due to perfect competition, they are also indifferent about sharing information.

Figure 10.2 illustrates the (expected) loan rates for time periods 1 and 2. We see that for all but very high or very low success rates, the loan rate in the second time period is higher than in the first time period and identical to those applied when information about the company type is shared, but with the time periods reversed. This indicates that banks compete in the first time period to obtain the loan by charging a loss-making loan rate and recovering these losses exploiting their informational advantage in time period 2, which allows them to charge a higher loan rate. For higher success rates, this informational advantage of the initial bank is becoming smaller compared to disclosing the type of company; the initial banks cannot generate sufficient profits from high loan rates in time period 2 to subsidize loan rates in time period 1 and attract companies, requiring loan rates in time period 1 to increase. The informational advantage for the initial bank is small because for high success rates, disclosing whether a company has repaid their loan, thus has been successful, is nearly identical to revealing the type of the company, it is becoming increasing unlikely that a company able to generate a successful investment, will fail.

Similarly, for low success rates the information shared with their competitors is of low value and hence even after sharing whether the company has defaulted, the informational advantage of the initial bank remains substantial. Disclosing some information in the form of whether the company defaulted or not will be very valuable to their competitors as the likelihood of companies that are able to generate successful investments, will not often success, making them undistinguishable from companies that do not generate successful investments and also fail. It is therefore that high loan rates in time period 2 cannot be sustained anymore and the lower profits to the initial bank subsequently do not allow for the subsidizing of loan rates in time period 1, requiring this loan rate to increase.

Summary Sharing information about companies through credit reference agencies with competitors is not beneficial to companies or banks, nor is it detrimental to their profits. The reason is that the perfect competition between banks will enable to company to extract all surplus from banks. While banks reduce their informational advantage after sharing the information, and hence their profits will reduce, this is offset by the level of competition to attract new companies. The sharing of information about companies will therefore mainly affect the loan rates that banks charge. If the amount of information shared is low, the bank holding this information will have a substantial advantage over its competitors and will be able to exploit this advantage by charging loan rates above their real costs, as those costs faced by competitors without the information are higher, generating profits from lending. These profits are then used to attract companies in the first place to generate the requisite information; to this effect banks will charge low loan rates to new companies. If the informational advantage of the initial bank is too low, the profits generated do not allow to reduce the initial loan rate much and the increased risk until the information is generated will dominate loan rates. For companies, these two effects exactly offset each other.

Sharing little information with competitors will allow for the initial loan rate to be lower than the loan rate in subsequent time periods, even though the bank faces less risk and covering their costs would imply a lower loan rate once the information has been learned. If companies are seeking to shift costs into later time periods or discount future costs, they would prefer banks to share less information about them, while those companies that seek to front-load costs would prefer banks to more share information. Reasons why companies might be concerned about the temporal allocation of loan rates might be found in concerns about tax planning, but also the presentation of profits in the accounts of companies and the ability to provide dividends to their owners.

Reading Padilla & Pagano (2000)

Conclusions

The sharing of information about the companies that banks are lending to erodes their informational advantage compared to other banks, who might not be able to generate such information without enjoying the same access to the company that a lending bank might have. This loss of informational advantage will necessarily increase competition between banks for providing loans to the company in the future, reducing the rent a bank is able to generate. However, the lower future rent will reduce the incentives to compete for new companies to lend to and thus compensate for the lower rent from future lending through the higher rent from new companies. Overall, companies and banks will generate the same profits, what will differ is the intertemporal allocation of profits. Without information disclosure, competition between banks in the future will be limited and future profits high, thus implying high loan rates, while in order to gain access to companies, they will compete fiercely to attract their custom, offering low loan rates as an incentive to commence borrowing from their bank. If information is shared, the incentives reverse. The lack of future profits arising from an informational advantage will reduce competition to attract companies.

While with such asymmetric information between banks, the disclosure of information only affects the degree of competition at different points of time with the effects cancelling each other out over time, it may well affect the moral hazard of a company. By sharing information on the risks the loan will be exposed to, the loan can be priced accordingly and those investment decisions that may benefit companies at the expense of banks will attract a higher loan rate. This will then discourage companies from making such investments, reducing the moral hazard.

Information provided by banks to credit reference agencies, who then allow access to this information to other banks, allows banks to gain information on the risks a company's investments impose on the bank, which can reduce adverse selection. It follows that high-risk companies prefer the disclosure of information while low-risk companies prefer that information is not disclosed. The reason is that high-risk companies have very little to lose from the disclosure of negative information but in case there is positive information will gain substantially. For low-risk companies the situation is reversed, there is not much benefit to them from positive information being disclosed, but the losses from negative information can be substantial.

Thus, information disclosure through credit reference agencies can affect the intertemporal distribution of loan costs by reducing adverse selection between banks from future lending and increase competition to attract companies in the first place. At the same time disclosing information can reduce moral hazard by companies through the ability of banks pricing loans more accurately to reflect they risks they are exposed to. It is primarily high-risk companies that benefit from the disclosure of information as they have more to gain than to lose from any information banks will provide, making the use of collateral less relevant.

Chapter 11

Relationship banking

While it is common to look at the relationship between a bank and its customers as a one-off transaction where the bank provides a loan for an investment the company is planning and after the investment has matured, the loan is repaid, which ends the transaction. For each such loan the bank would start to evaluate the company again to determine the risks of the investment they are seeking finance for. In reality, however, Banks lend repeatedly to the same company and will to a large degree depend on information accumulated from previous loans and risk assessments, rather than start afresh with their analysis. Thus banks continue to accumulate information on a companies through the repeated provision of loans, what is commonly referred to as relationship banking. In such relationship banking implicit contracts are common that anticipate a certain action, e.g. the extension of a loan or the investment of the loan proceedings in low risk projects, even if this would not be legally enforceable. In contrast to this, in transaction banking the focus is firmly only on the transaction on hand without any indication of future behaviour or the accumulation of more information than needed at this point. Thus relationship banking takes into account the experience the bank has from previous loans, but also considers not only the loan the company is seeking currently, but will take into account any future lending.

The most prominent effect of relationship banking is on loan rates as we will explore in chapter 11.1. Relationship banks will gain an informational advantage over competitors and can exploit this advantage by offering not fully competitive loan conditions, allowing banks to make excess profits from such relationships. We will also see, however, that competition to gain access to companies in the first place will limit the effect such informational advantage will have on companies. But it is not only the loan conditions that are affected by relationship banking, but the access to loans in the first place. In relationship banking the informational advantage of banks might allow them to grant loans where other banks might not be willing to offer a loan. On the other hand, relationship banks might refuse loan to companies they seem to be not creditworthy, but which other banks might happily offer a loan to. We will look into these aspects in chapter 11.2, where we will also see why companies might seek to have relationships with more than just a single bank. Finally, chapter 11.3 investigates the impact of competition on relationship banking. On the one hand, with increasing competition the profit margins of banks are reduced and this makes it more difficult to recover any costs that is not incurred in transaction banking. On the other hand, informational advantages in relationship banking cannot easily be eroded from competition and can become a major source of profits, making relationship banking ever more important for banks.

11.1 Optimal loan rates

Banks accumulate information throughout the lending relationship with a company and therefore their informational advantage over other banks continuously increases. This advantage can be exploited by banks when providing loans with conditions that are not competitive. Companies are not able to switch as either other banks do not have the same information to make a more favourable offer of the costs of companies switching their loan to another bank makes such a measure unattractive. This way companies are 'locked-in' a relationship as less well informed banks will not be able to offer better conditions.

Chapter 11.1.1 will show how the cost of obtaining information by a competitor prior to offering a loan will affect the loan rates of the initial bank, while in chapter 11.1.2 we will explore the impact the more precise information held by the current lender has on the competition with other banks not int his privileged position. The crucial feature in these models is that companies can switch banks if these offer better conditions. However, often switching banks is not without costs t the company and in chapter 11.1.3 we will see how such switching costs can affect the loan rates banks set. Surprisingly, switching costs can lead to reduced adverse selection between banks and thus enhance competition. Avoiding companies switching banks will allow loan conditions to be set such that banks can exploit their informational advantage over competitor banks even before they have obtained this informational advantage as we show in chapter 11.1.4. Thus banks can anticipate informational advantages and charge more consistent loan rates to companies.

11.1.1 Exploitation of information monopolies

Banks have to acquire information on companies before they provide loans to assess the risks the investments of the company impose on their ability to repay the loan. This acquisition and processing of information will be costly to banks, but can easily be re-used when providing future loans, with minimal costs for updating the already existing information. Thus banks that have previously provided a loan to a company have the advantage of facing lower costs when providing loans.

Let us assume that a company seeks to finance an investment fully by a loan L and can make identical investments for two consecutive time periods. This investment is successful with probability π, which allows the company to repay the loan in full; if the investment is not successful, the loan cannot be repaid at all. Companies need to obtain a loan in each time period and can switch banks prior to making the second investment. If the investment fails in the first time period, the company cannot obtaining a loan in the second time period. The initial assessment of the company when providing a loan for the first time is C and the costs for any subsequent loan is zero. These costs are incurred upfront and need to be financed by deposits, as is the loan; the bank has to pay depositors interest r_D.

If the company were to switch banks after time period 1, the new bank would charge a loan rate of r_L^2 such that their profits are given by

$$\Pi_B^2 = \pi \left(1 + r_L^2\right) L - (1 + r_D)(L + C), \tag{11.1}$$

If banks are competitive such that their profits are eliminated, $\Pi_B^2 = 0$, the loan rate in time period 2 is given by

$$1 + r_L^2 = \frac{1 + r_D}{\pi} \frac{L + C}{L}. \tag{11.2}$$

A bank granting a loan in time period 1 will not face any costs to provide a loan in time period 2 and could therefore charge a lower loan rate, but as no other bank can undercut the loan rate r_L^2, it would not charge a lower loan rate to maximize its profits.

A bank granting a loan in time period 1 will be repaid this loan with probability π and then be able to extend this loan, which in time epriod 2 is again repaid with probability π. Thus with a loan rate in the first time period of r_L^1, the banks' profits are given by

$$\begin{aligned}\Pi_B^1 = \pi &\left(\left(1 + r_L^1\right) L + \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right)\right) \\ &- (1 + r_D)(L + C).\end{aligned} \tag{11.3}$$

Banks will be competing to provide loan in time period 2 such that $\Pi_B^1 = 0$ and after inserting for r_L^2 from equation (11.2), we easily get

$$1 + r_L^1 = \frac{1 + r_D}{\pi} \frac{(1 - \pi)(L + C) + \pi L}{L} < 1 + r_L^2. \tag{11.4}$$

The loan rate in the first time period is lower than in the second time period, even though the initial bank faces the costs of acquiring and processing information, while facing no such costs in time period 2. The reason for this result is that the costs faced by any bank seeking to compete with the initial bank will face such costs and thus the initial bank can charge a loan rate in time period 2 that generates a profit to them. There is, however, competition for new companies in time period 1, and banks will incur a loss by charging a loan rate below their costs, which they then recover in time period 2.

Successful companies pay higher aggregate loan costs than they would do if markets were competitive in each time period. The loan rate for a single time period is given by equation (11.2), where when staying with the same bank in time period 2 we would have $C = 0$. The total loan costs are therefore given by

$$\left(1+\hat{r}_L^1\right)L+\left(1+\hat{r}_L^2\right)L = \frac{1+r_D}{\pi}(L+C)+\frac{1+r_D}{\pi}L \qquad (11.5)$$
$$= \frac{1+r_D}{\pi}(2L+C).$$

The loan costs here, however, are given as

$$\left(1+r_L^1\right)L+\left(1+r_L^2\right)L = \frac{1+r_D}{\pi}\left((1-\pi)(L+C)+\pi L\right) \qquad (11.6)$$
$$+\frac{1+r_D}{\pi}(L+C)$$
$$= \frac{1+r_D}{\pi}\left((2-\pi)(L+C)+\pi L\right),$$

where we insert from equations (11.2) and (11.4) for the loan rates. We easily see that the expression in equation (11.6) is higher than the expression in equation (11.5) and therefore companies are paying more interest on their loans compared to a situation where banks are competitive in each time period. The bank will not be able to make profits in the second time period off unsuccessful companies as these do not obtain any more loans; thus they will make losses from offering loan rates that do not cover the information cost, which have to be recovered from successful companies, increasing their overall loan costs.

This result can be interpreted as successful companies subsidizing unsuccessful companies. The higher loan rates in time period 2, which subsidize the lower loan rates in time period 1, are only charged to successful companies. Unsuccessful companies benefit from the lower loan rates in time period 1, which does not fully reflect the costs banks face.

We see that as banks build a relationship with a company and accumulate information, banks gain a cost advantage over competitors that allows them to charge loan rates generating them a profit. These future profits will, however be used in competitive markets to attract companies in the first instance by banks offering loan rates below their costs. This leads to attractive initial loan rates that are then increased at a later time.

Reading Freixas & Rochet (2008a, Ch. 3.6.1)

11.1.2 Exploiting informational advantage

Banks will not only seek to acquire information when first providing a loan to a company, but they will also seek to continue to acquire more information as the relationship with the company continues. Through this process banks will obtain ever more information about a company, which should reduce the uncertainty about the risks the companies are exposed to. This accumulation of information over time will give the current bank of a company a distinct informational advantage over any other bank who does not benefit from a relationship.

Let us assume that the true probability of success for the investment of a company is π, but that the assessment of the bank is not perfect and fluctuates randomly around this true value such that the observed probability of success is $\pi_i = \pi + \varepsilon_i$, where $i = 1$ represents a situation in which the bank lends to the company for the first time and $i = 2$ where the bank has lent to the company in the previous time period. The observed probabilities are unbiased in that $\mathrm{E}[\varepsilon_i] = 0$ and for the variances we assume that $\mathrm{Var}[\varepsilon_1] = \sigma_1^2 > \sigma_2^2 = \mathrm{Var}[\varepsilon_2]$, such that over time the variance reduces and the information of the bank becomes more precise. For simplicity we assume that companies only demand loans for two time periods and banks only provide a loan in the second time period if the company is successful in time period 1. The investment the company conducts in both time periods will be identical.

A bank will provide the loan if it offers the company the lower loan rate. For a bank that had no previous relationship with the company, the expected profits in time period 2 will then be given by

$$\hat{\Pi}_B^2 = \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)\left(\hat{\pi}_1\left(1 + \hat{r}_L^2\right)L - (1 + r_D)\,L\right), \tag{11.7}$$

where $\hat{r}_L^2$ denotes the loan rate offered by the new bank, r_L^2 the loan rate offered by the existing bank and $\hat{\pi}_1$ denotes the information the new bank has obtained about the company's probability of success. Maximizing this expression, we obtain the first order condition as $\frac{\partial \hat{\Pi}_B^2}{\partial\left(1+\hat{r}_L^2\right)} = 0$, which gives us the optimal loan rate of the new bank as

$$\begin{aligned} 1 + \hat{r}_L^2 &= \frac{1 + r_D}{\hat{\pi}_1} - \frac{\mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+\hat{r}_L^2\right)}} \\ &= \frac{1 + r_D}{\hat{\pi}_1} + \frac{\mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}, \end{aligned} \tag{11.8}$$

where for the last equality we used that $\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+\hat{r}_L^2\right)} = -\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)} < 0$.

We can insert this result into equation (11.7) to obtain the profits of the new bank in time period 2 as

$$\hat{\Pi}_B^2 = \frac{\mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)^2}{\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}} L. \tag{11.9}$$

Similarly, for the initial bank we obtain their profits in time period 2 as

$$\Pi_B^2 = \left(1 - \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)\right)\left(\pi_2\left(1 + r_L^2\right)L - (1 + r_D)\,L\right) \tag{11.10}$$

and the first order condition $\frac{\partial \Pi_B^2}{\partial\left(1+r_L^2\right)} = 0$ for maximum profits gives its optimal loan rate as

$$1 + r_L^2 = \frac{1 + r_D}{\pi_2} + \frac{1 - \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}. \tag{11.11}$$

The profits of the initial bank are then given after inserting this result into equation (11.10) and we obtain

$$\Pi_B^2 = \frac{\left(1 - \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)\right)^2}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}} L. \tag{11.12}$$

As from statistics we now that, as an approximation, for a random variable x we have $\mathrm{E}\left[\frac{1}{x}\right] \approx \frac{1}{\mathrm{E}[x]} + \frac{\mathrm{Var}[x]}{\mathrm{E}[x]^3}$, we easily get the expected loan rates of the new and existing bank given by

$$\mathrm{E}\left[1 + \hat{r}_L^2\right] = \frac{1 + r_D}{\pi}\left(1 + \frac{\sigma_1^2}{\pi^2}\right) + \frac{\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}, \tag{11.13}$$

$$\mathrm{E}\left[1 + r_L^2\right] = \frac{1 + r_D}{\pi}\left(1 + \frac{\sigma_2^2}{\pi^2}\right) + \frac{1 - \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}.$$

It must now be that $\mathrm{E}\left[1 + \hat{r}_L^2\right] > \mathrm{E}\left[1 + r_L^2\right]$. The first terms in equation (11.13) will be smaller for $\mathrm{E}\left[1 + r_L^2\right]$ as $\sigma_2^2 < \sigma_1^2$ and then it is consistent that $\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right) < \frac{1}{2}$, reinforcing that the average loan rate of the existing bank is lower.

The profits of banks in time period 1 are given from the loan repayments and the profits generated in time period 2, provided the bank is selected to provide the initial loan. If the bank is not selected to provide the loan in time period 1, it will only be able to make profits in time period 2 as a new bank, provided it offers the lower loan rate. With r_L^1 denoting the loan rate of the bank under consideration and $\hat{r}_L^1$ the loan rate of their competitor, we get

$$\begin{aligned}\Pi_B^1 = &\left(1 - \text{Prob}\left(\hat{r}_L^1 < r_L^1\right)\right)\left(\pi_1\left(1 + r_L^1\right)L - (1 + r_D)L + \pi_1 \Pi_B^2\right) \\ &+ \text{Prob}\left(\hat{r}_L^1 < r_L^1\right)\pi_1 \hat{\Pi}_B^2.\end{aligned} \tag{11.14}$$

We note that loan in time period 2 are only provided if the company succeeds in time period 1

If banks are competitive in time period 1, they would charge loan rates as low as possible to attract companies, implying that $\Pi_B^1 = 0$. We can now insert for Π_B^2 and $\hat{\Pi}_B^2$ from equations (11.9) and (11.12) and note that in time period 1 banks are identical due to none having superior information about the company, they will both have ex-ante the same probability of being chosen, thus $\text{Prob}\left(\hat{r}_L^1 < r_L^1\right) = \frac{1}{2}$. This then solves for the loan rate in time period 1 to be given by

$$1 + r_L^1 = \frac{1 + r_D}{\pi_1} - \frac{1 - 2\text{Prob}\left(\hat{r}_L^2 < r_L^2\right)\left(1 - \text{Prob}\left(\hat{r}_L^2 < r_L^2\right)\right)}{2\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1 + r_L^2\right)}}. \tag{11.15}$$

The expected loan rate in time period 1 is then obtained as

$$\begin{aligned} \text{E}\left[1 + r_L^1\right] = {} & \left(\frac{1}{\pi} + \frac{\sigma_1^2}{\pi^3}\right)(1 + r_D) \\ & - \frac{1 - 2\text{Prob}\left(\hat{r}_L^2 < r_L^2\right)\left(1 - \text{Prob}\left(\hat{r}_L^2 < r_L^2\right)\right)}{2\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1 + r_L^2\right)}}. \end{aligned} \tag{11.16}$$

Comparing the first terms of the expected loan rates in time periods 1 and 2 from equations (11.13) and (11.16), we see that the difference between those terms will be $\frac{1+r_D}{\pi^3}\left(\sigma_1^2 - \sigma_2^2\right) \geq 0$, but then the loan rate of the first time period is reduced by the second term and that of the second time period increased. If no information is accumulated by the existing bank, $\sigma_1^2 = \sigma_2^2$, it is obvious that $\text{E}\left[1 + r_L^1\right] < \text{E}\left[1 + r_L^2\right]$. As more an more information accumulates, reducing σ_2^2, the second term will be more sensitive to the reduction in uncertainty of the existing bank due the quadratic probability term. It is therefore that as more information is accumulated, the increase in the loan rate in time period 2 becomes more pronounced.

As more information accumulates and σ_2^2 decreases, it will allow the bank to reduce their loan rate in time period 2 slightly as they are expecting a higher profit due to the reduced uncertainty, allowing them to reduce the loan rate to increase the likelihood of obtaining this second loan, thus $\text{Prob}\left(\hat{r}_L^2 < r_L^2\right)$ reduces. This will then increase the second term of the expected loan rate in time period 1 and hence reduce the expected loan rate. This effect is stronger than the effect in reducing the expected loan rate in time period 2 as they have to balance the reduced profits against the increased likelihood of providing the loan, which is not the case in time period 1. Hence, the difference in the two loan rates is increasing the more information is accumulated.

Banks having provided the loan in the first time period, accumulate information about the company and thus gain an informational advantage over the other bank in time period 2. This allows banks to make excess profits in time period 2 by charging loan rates above their costs. In time period 1, banks will compete to provide a loan to the company, which then will enable them to gain the informational advantage and profits in time period 2. Competition to be chosen by the company to provide the loan and learn the information will lead to banks charging loan rates below their costs in time period 1, thus leading to low loan rates. The low initial loan rate is then increased in time period 2 as the informational advantage does not allow the company to change banks easily, they have to rely on the other bank making a very benign assessment of their risks to be offered better conditions, this is quite unlikely to happen.

Particularly for companies where banks can accumulate information well through an existing relationship, will we observe lower initial loan rates that are then increased once the bank has captured this information and gained an informational advantage. We might expect such a situation to emerge particularly in industries where intimate knowledge of companies and their management is required to assess the risks of their investments, or in situations where formal reporting requirements are insufficient to enable a full risk assessment of the company. The more information can be accumulated while being a lender to a company, the larger the differences in loan rates over time will be. Banks will initially compete to provide the loan and then subsequently exploit their informational advantage by recovering their losses through higher loan rates.

Reading Greenbaum, Kanatas, & Venezia (1989)

11.1.3 The impact of switching costs

It is often assumed that companies are offered loan contracts by their existing bank and if they can find a more favourable offer from another bank, they can switch their loan to this bank. In reality, however, companies will face additional costs to switch banks; such costs may include the provision of information to the new bank, negotiations on the loan terms that arise due to not having an established common ground, or any costs associated with changing current accounts to the new bank. Hence, once a relationship with a bank has been formed, such switching costs may make it costly to take up a loan at another bank, even though its conditions are more favourable.

Let us assume there are two types of companies these whose investments succeed either with a high probability, π_H, or a low probability $\pi_L < \pi_H$. Companies having a high probability of success are a fraction p of all companies, while those with a low probability of success are a fraction $1 - p$. If successful, the investment will provide the company with an outcome V, and if the investment is not successful, no outcome is generated. We can define the average probability of success as $\pi = p\pi_H + (1 - p)\,\pi_L$. With banks having to pay a deposit rate of r_D and providing a loan of L, which finances the investment in full, we assume that $\pi_H V > (1+r_D)L > \pi_L V$. This assumption implies that investments by companies with high success rates are desirable and able to cover the costs of banks to provide the loan, while companies with low success rates cannot cover the costs of their loan.

We further assume that each company has established a relationship with a bank and thus the costs of this bank to gain information about the company can be ignored. Their bank will have information on the type of company, while other banks will not have any information; the information the exiting bank holds, however, is not perfect. Let us assume that it is correct with probability ρ. Relationship banks assess a company to have a high success rate if the company actually has a high success rate, p, and the information they have received is correct, ρ, or if the company is actually having a low success rate, $1 - p$, and the information received is incorrect, $1 - \rho$. Thus the probability of a company being seen as having a high success rate is given by

$$\hat{\nu} = \rho p + (1 - \rho)\,(1 - p)\,. \tag{11.17}$$

Using Bayes' theorem, the probability of the bank obtaining the information that it has a high and low success rate, respectively, and it actually being of this type is given by

$$\hat{\pi}_H = \frac{\rho p}{\hat{\nu}}, \qquad (11.18)$$
$$\hat{\pi}_L = \frac{\rho (1-p)}{1-\hat{\nu}}.$$

The probability that the company has a high (low) success rate, p $(1-p)$, and this assessment is correct, ρ, has to be compared to the probability that the bank has assessed it as having a high (low) success rate, $\hat{\nu}$ $(1-\hat{\nu})$.

Let us now assume that we have

$$(\hat{\pi}_H \pi_H + (1-\hat{\pi}_H)\,\pi_L)\,V > (1+r_D)\,L, \qquad (11.19)$$
$$(\hat{\pi}_L \pi_L + (1-\hat{\pi}_L)\,\pi_H)\,V < (1+r_D)\,L.$$

This assumption can be interpreted as an extension of our initial assumption that $\pi_H V > (1+r_D)L > \pi_L V$. Rather than assessing the companies at their actual success rates, we assume that companies that have been assessed as having high success rates can generate profits to the bank by having expected outcomes exceeding the funding costs of their loans, while companies that have been assessed as having low success rates, will not be able to generate profits t the bank. This more strict requirements implies that companies with low success rates are not too often identified as having a high success rate and vice versa, ensuring that relationship banks can distinguish the different types of companies sufficiently well.

Companies face switching costs S if they want to take up the loan offer of another bank and would only do so if the profits they can make are higher than when staying with their current bank, even after incurring these costs. With a loan rate r_L being offered by their existing bank and $\hat{r}_L$ by a new bank, the respective profits to the company are given by

$$\Pi_C^i = \pi_i\,(V - (1+r_L)\,L)\,, \qquad (11.20)$$
$$\hat{\Pi}_C^i = \pi_i\,(V - (1+\hat{r}_L)\,L) - S.$$

Companies will take the loan offer of their existing bank if $\Pi_C^i \geq \hat{\Pi}_C^i$, or

$$1 + r_L \leq 1 + \hat{r}_L + \frac{S}{\pi_i L}, \qquad (11.21)$$

allowing exiting banks to charge higher loan rates and exploit the advantage that arises from the switching costs that companies face.

If we assume that banks are only able to offer a single loan rate for all companies, they need to ensure that this condition is fulfilled for companies with high success rates as these are the profitable companies for the bank. Furthermore, banks will not charge a loan rate that is lower than necessary, implying that the relationship in equation (11.21) will hold with equality. We thus obtain the relationship between the loan rates of the initial and any new banks as

$$1 + r_L = 1 + \hat{r}_L + \frac{S}{\pi_H L}. \tag{11.22}$$

Of course, this loan rate is only feasible if the investment of the company is profitable, $\Pi_C^i \geq 0$, thus we require that $1 + r_L \leq \frac{V}{L}$.

As lending to companies which have been assessed as having a high success rates is profitable to the existing bank due to our assumption that $(\hat{\pi}_H \pi_H + (1 - \hat{\pi}_H) \pi_L) V > (1 + r_D) L$, their bank will always make an offer to provide another loan to these companies. Companies for which information is received suggesting they have low success rates cannot be profitable to the initial bank as we had assumed that $(\hat{\pi}_L \pi_L + (1 - \hat{\pi}_L) \pi_H) V < (1 + r_D) L$. However, not all companies for which information suggests they have a low success rate are actually exhibiting a low success rate; some companies will be wrongly identified. Similarly not all companies which have been assessed as having a high success rate will actually exhibit this high success rate. Let us therefore assume that companies who do not obtain a loan offer from their existing bank switch banks with probability λ_H for companies that are actually exhibiting a high sucess rate and λ_L for those companies that actually exhibit a low success rate. Companies know their own success rate and can thus behave differently, depending on their type.

Of those companies which have high success rates, p, and have been incorrectly assessed as having low success rates and are therefore not offered a loan by their existing bank, $1 - \rho$, a fraction λ_H switches banks, as does a fraction λ_L of those companies that have low success rates, $1 - p$, and have been correctly assessed as such, ρ. Thus the fraction of companies with high success rates that are switching banks, is given by

$$\hat{p} = \frac{\lambda_H (1 - \rho) p}{\lambda_H (1 - \rho) p + \lambda_L \rho (1 - p)}. \tag{11.23}$$

For the new bank the likelihood of the loan being repaid is given by all those companies that are having high success rates π_H and switch, in addition to those that switch having low success rates π_L, such that $\hat{\pi} = \hat{p}\pi_H + (1 - \hat{p})\pi_L$. The profits of the new bank are then given by

$$\hat{\Pi}_B = \hat{\pi}(1 + \hat{r}_L)L - (1 + r_D) L. \tag{11.24}$$

If new banks are competiing for switching companies, they will erode any profits they can make, such that $\hat{\Pi}_B = 0$. Therefore they set a loan rate of

$$1 + \hat{r}_L = \frac{1 + r_D}{\hat{\pi}}. \tag{11.25}$$

Let us now assume that companies not being offered a loan are always switching banks, thus $\lambda_H = \lambda_L = 1$, from which we obtain with equation (11.23) that $\hat{p} = 1 - \hat{\pi}_L$. The company profits are then given by

$$\hat{\Pi}_C^1 = \pi_i \left(V - (1 + \hat{r}_L) L\right) - S \tag{11.26}$$
$$= \pi_i \left(V - \frac{1 + r_D}{(1 - \hat{\pi}_L) \pi_H + \hat{\pi}_L \pi_L} L\right) - S < 0,$$

where we inserted for the loan rate from equation (11.25) and for $\hat{\pi}$ from equation (11.23) with $\lambda_H = \lambda_L = 1$. Comparing this result with our assumption in equation (11.19), we see that company profits will be negative. It will thus not be optimal for all companies that are identified by their existing bank as having a low rate of success to switch banks. Some banks will have to accept that they will not obtain another loan.

In order for companies with low success rates to demand a loan from their existing bank, their profits need to be positive, thus $\Pi_C^L = \pi_L \left(V - (1 + r_L)L\right) \geq 0$. If banks are able to extract any surplus from these companies, this will fulfilled with equality and we have that $1 + r_L = \frac{V}{L}$. Inserting this result into equation (11.21), we obtain that the loan rate of the new bank is given by

$$1 + \hat{r}_L = \frac{\pi_L V - S}{\pi_L L}. \tag{11.27}$$

Comparing equations (11.25) and (11.27), we easily get

$$\hat{p} = \pi_L \frac{(1 + r_D) L - \pi_L V + S}{(\pi_H - \pi_L)(\pi_L V - S)}. \tag{11.28}$$

As for the same loan rate, companies with higher success rates will generate higher profits, they will always switch banks if companies with low success rates switch banks, hence if $\lambda_L > 0$, we have $\lambda_H = 1$. Thus equation (11.23) can be solved for the probability that companies with low success rates are switching banks:

$$\lambda_L = \frac{1 - \hat{p}}{\hat{p}} \frac{(1 - \rho)p}{\rho(1 - p)}. \tag{11.29}$$

If $\hat{p} < 1$ both types of companies that have not been offered a loan by their existing bank due to being assessed as having low success rates, are switching banks. From equation (11.28), this condition requires that

$$S < S^* = \pi_L \left(V - \frac{1 + r_D}{\pi_H} L\right). \tag{11.30}$$

Provided switching costs are not too high, both types of companies will switch banks if their initial bank assesses them as having low success rates. Inserting for $\hat{p}$ from equation (11.28) into the definition of $\hat{\pi}$, we obtain that $\hat{\pi} = \pi_L \frac{(1+r_D)L}{\pi_L V - S}$. Using equations (11.22) and (11.27), we easily get the loan rates of the initial and new bank as

$$1 + r_L = \frac{V}{L} - \frac{S}{L}\left(\frac{1}{\pi_L} - \frac{1}{\pi_H}\right), \tag{11.31}$$
$$1 + \hat{r}_L = \frac{V}{L} - \frac{S}{\pi_L L}.$$

As the switching costs S increase, the loan rate falls. We easily obtain that

$$\frac{\partial(1 + \hat{r}_L)}{\partial S} = -\frac{1}{\pi_L L} < -\frac{1}{L}\left(\frac{1}{\pi_L} - \frac{1}{\pi_H}\right) = \frac{\partial(1 + r_L)}{\partial S} < 0. \tag{11.32}$$

and the loan rate of the new bank is falling faster than the loan rate of the initial bank, ensuring that the loan rate is attractive for companies to switch with increasing costs. That the loan rate of the initial bank is falling as switching costs increase, can be explained with the observation that as switching costs increase, more and more companies with low success rates find switching banks unattractive and do not request loans at all. This increases the fraction of companies with high success rates in the market, reducing the risks to the bank and hence allowing them to reduce the loan rate. It also reduces the adverse selection between the initial and the new bank as companies the initial bank does not want to offer loans to, those they assess as having low success rates, are less and less demanding loans.

Once we have reached switching costs of S^*, no companies with low success rates will demand any loans if refused a loan by their existing bank, thus $\hat{p} = 1$ and $\hat{\pi} = \pi_H$, implying that loan rates are given by

$$1 + r_L = \frac{(1 + r_D)\, L + S}{\pi_H L}, \tag{11.33}$$
$$1 + \hat{r}_L = \frac{1 + r_D}{\pi_H}.$$

The loan rates of the initial banks are increasing while those of new banks remain constant,

$$0 = \frac{\partial(1 + \hat{r}_L)}{\partial S} < \frac{1}{\pi_H} = \frac{\partial(1 + r_L)}{\partial S}. \tag{11.34}$$

The competition between new banks will not allow them to raise loan rates beyond their costs and they know that only companies with high success rates will be remaining in the market to demand loans from new banks. Increasing switching costs will enable the initial bank to exploit their advantage by raising loan rate, which remains attractive due to the higher switching costs for their existing companies that have been assessed as having high success rates.

As switching costs increase, there will be a point at which it will become unprofitable for any company with high success rates to switch banks, even if being denied a loan by their own bank. As only companies with high success rates remain the market, we have $\hat{\pi} = \pi_H$ and hence company profits of those switching banks are given by

$$\hat{\Pi}_C^H = \pi_H\,(V - (1 + r_L)\, L) - S = \pi_H V - (1 + r_D)\, L - S \tag{11.35}$$

and requiring company profits to be positive, $\hat{\Pi}_C^H \geq 0$, is only achieved if

$$S \leq S^{**} = \pi_H V - (1 + r_D) L. \tag{11.36}$$

Once the switching costs exceed S^{**}, no company will be switching banks. This will allow the existing bank to extract all surplus from the company such that $\Pi_C^H = 0$ and hence $1 + r_L = \frac{V}{L}$.

We thus observe that initially, as switching costs are increasing, the loan rates of existing and new banks are falling as less and less companies with low success rate demand loans from new banks and thus reducing the risk to banks. The loan rate of new banks falls more to compensate for the higher switching costs their new companies face, having to compensate companies incurring these switching costs by offering lower loan rates. Once switching costs have reached a threshold of S^*, no companies with low success rates demand loans from new banks anymore and the initial bank can exploit their advantage over new banks arising from switching costs and raise loan rates as switching costs increase. This is feasible until the loan rate reaches a level S^{**} in which no more companies switch banks. This allows the initial bank to capture all their current companies without facing competition from new banks, allowing them to extract all surplus from their companies.

We thus see that increasing costs for companies to switch banks after having been refused a loan by their own bank due to being assessed as high risk, may reduce loan rates as it changes the composition of those companies seeking to switch banks. The higher the switching costs, the fewer high-risk companies will seek to switch and hence the risks to banks from offering such switching companies a loan, reduces, allowing competition between the existing and new banks to reduce loan rates. Once all high-risk companies cease to demand loans from new banks, the existing bank can exploit its advantage over new banks in that companies face ever increasing switching costs, making staying with the existing bank more attractive despite rising loan rates. With very high switching costs, existing banks obtain a monopoly position as no company will seek to switch banks, and thus can extract any surplus from companies.

It is therefore not necessarily in the interest of low-risk companies to reduce the costs of switching banks too much. They can benefit from moderately high switching costs by making switching banks unattractive to high-risk companies and thereby reduce the adverse selection of new banks, in addition to increasing the fraction of low-risk companies seeking to switch banks, which will increase competition between their existing bank and other banks they might switch to, which will then reduce loan rates. From this result follows that efforts to reduce costs associated with switching banks might actually be increasing loan rates as more high-risk companies are induced to switch banks.

Reading Vesala (2007)

11.1.4 Long-term contracts

Relationship banking can give rise to banks in that relationship having superior information to other banks who lack such a relationship with a company, giving them an informational advantage they could seek to exploit in future loan offers by quoting less than competitive loan rates. The exploitation of such informational advantages can only occur if banks are able to increase loan rates for subsequent loans. It could, however, be optimal for banks and companies to agree a contract that prevents such increases.

Let us assume there are two types of companies who only differ in their probability of succeeding with their investment. One type of company, representing a fraction p of all companies, has a high probability of success π_H and the other, representing a fraction $1 - p$ of all companies, a low probability of success $\pi_L < \pi_H$. If the investment fails, no value is generated and the loan cannot be repaid; in the case of success the return on investment will be R, allowing the loan to be repaid. This return, however, will be dependent on the size of the investment, showing a falling rate of return as the investment increases; if we assume that the investment is fully financed by a loan of size L, we thus assume that $\frac{\partial R}{\partial L} < 0$. Companies pursue identical investments over two time periods, but can obtain a loan in the second time period only if they have been successful with their initial investment.

Contracts allowing to switch banks We assume that initially the bank does not know the type of company seeking a loan. It only knows that a fraction p of the companies has a high probability of success and the remainder a low probability of success. Hence the expected rate of success in time period 1 is given by $\pi = p\pi_H + (1 - p)\pi_L$. The bank only learns the type after having lent to the company for one time period at no costs, while banks not having lent to the company in time period 1, will not be aware of their type when offering a loan in time period 2. Thus a bank not having lent to the company in time period 1 will in time period 2 make profits of

$$\Pi_B^2 = \pi\left(1 + \hat{r}_L^2\right)L - (1 + r_D)L. \tag{11.37}$$

The same profits will be made by all banks on the loan they grant in time period 1. The lowest loan rate this bank would offer is if $\Pi_B^2 = 0$ and thus

$$1 + \hat{r}_L^2 = 1 + \hat{r}_L^1 = \frac{1 + r_D}{\pi}. \tag{11.38}$$

The initial bank will have learned the type of company they are lending to and hence make profits of

$$\hat{\Pi}_B^{2,i} = \pi_i\left(1 + \hat{r}_L^{2,i}\right)L - (1 + r_D)L, \tag{11.39}$$

for lending to companies with high success rates, $i = H$ and low success rates, $i = L$, respectively. The lowest loan rate this bank would offer is if $\hat{\Pi}_B^{2,i} = 0$ and thus

$$1 + \hat{r}_L^{2,i} = \frac{1 + r_D}{\pi_i}, \tag{11.40}$$

such that $\hat{r}_L^{2,H} < \hat{r}_L^2 < \hat{r}_L^{2,L}$. If the new bank is offering competitive loan rates, they would attract all companies with low probabilities of success, but companies with high success rates could remain with the initial bank as it is able to offer a loan rate below that of the new bank while still being profitable. The existing bank would seek to exploit its informational advantage offering a loan rate of r_L^2 while the new bank, knowing it will only be able to attract companies with low rates of success, will charge $\hat{r}_L^{2,L}$, matching the offer of the existing bank.

The existing bank will make profits from providing loans to companies with high success rates of $\hat{\Pi}_B^{2,H} = \pi_H \left(1 + \hat{r}_L^2\right) L_H - (1 + r_D) L_H$. If we insert for $1 + r_D$ after solving equation (11.40) and use the definition of π, these profits become

$$\hat{\Pi}_B^{2,H} = p\frac{(1-p)(\pi_H - \pi_L)}{\pi}(1 + r_D) L_H, \tag{11.41}$$

where the initial p arises from the fact that these profits only arise for companies that are exhibiting a high success rate. As competition with new banks forces the initial bank to offer competitive loan rates for companies with low success rates, they do not obtain any profits from this lending, and $\hat{\Pi}_B^{2,H}$ gives the profits a bank makes from lending in time period 2. In time period 1, all banks have identical information and they will offer competitive loan rates set as given in equation (11.40) and will make no profits, such that the total profits of a bank is given by equation (11.41).

Contracts not allowing to switch banks Let us now consider a different form of loan contract. The loan contract above was structured such that the company could switch banks after the first time period at no costs, now however we assume that a bank offers a contract in which the company commits itself to take up another loan with the same bank, provided its investment is successful in time period 1. It does not allw the company to switch banks. The profits of the bank consist of the profits from the loan in the first time period, for which it charges a loan rate r_L^1, and then, provided this first investment is successful, the provision of a loan in the second time period, for which the bank will be allowed to charge different loan rates depending on the type of company that is identified, $r_L^{2,H}$ and $r_L^{2,L}$, respectively. Hence we have

$$\begin{aligned}\Pi_B = {} & \left(\pi\left(1 + r_L^1\right) - (1 + r_D)\right) L \\ & +\pi\left(p\left(\pi_H\left(1 + r_L^{2H}\right) - (1 + r_D)\right) L_H\right. \\ & \left. + (1-p)\left(\pi_L\left(1 + r_L^{2L}\right) - (1 + r_D)\right) L_L\right).\end{aligned} \tag{11.42}$$

Banks will offer such a loan contract if it is at least as profitable as the contract that allows companies to switch banks. Competition between banks offering this type of contract will ensure that the profits are identical, such that $\hat{\Pi}_B^{2,H} = \Pi_B$.

Assume now that the company in time period 1 does not know its own type and hence expects its success rate to be π. Companies generate a return of R and repay the loan with probability π_i if they know their type and with probability π if they are unaware of their own type. Along with the initial bank, companies will learn their type after the first time period. Their profits across the two time periods are given by the profits from the first time period and, if successful, in the second time period. With a return of R on the successful investment, we get these profits as

$$\begin{aligned}\Pi_C = {} & \pi\left((1+R)\,L - \left(1+r_L^1\right)L\right) \\ & +\pi p \pi_H\left((1+R)\,L_H - \left(1+r_L^{2,H}\right)L_H\right) \\ & +\pi\,(1-p)\,\pi_L\left((1+R)\,L_L - \left(1+r_L^{2,L}\right)L_L\right).\end{aligned} \tag{11.43}$$

The amount of the loans in time periods 1 and 2, L, L_H, and L_L, are determined by the company from maximizing their profits as given in equation (11.43), the first order conditions $\frac{\partial \Pi_C}{\partial L} = \frac{\partial \Pi_C}{\partial L_H} = \frac{\partial \Pi_C}{\partial L_L} = 0$ easily gives us

$$\begin{aligned}L &= -\frac{R - r_L^1}{\frac{\partial R}{\partial L}}, \\ L_H &= -\frac{R - r_L^{2,H}}{\frac{\partial R}{\partial L_H}}, \\ L_L &= -\frac{R - r_L^{2,L}}{\frac{\partial R}{\partial L_L}},\end{aligned} \tag{11.44}$$

where by assumption $\frac{\partial R}{\partial L} < 0$, $\frac{\partial R}{\partial L_H} < 0$, and $\frac{\partial R}{\partial L_L} < 0$ and hence all expressions are positive. Let us now assume that as banks are competitive, companies are able to extract a surplus from banks and the loan rates are set such that their profits are maximized, subject to the requirement that $\hat{\Pi}_B^{2,H} = \Pi_B$. This gives our objective function as

$$\mathcal{L} = \Pi_C + \zeta\left(\Pi_B - \hat{\Pi}_B^{2,H}\right), \tag{11.45}$$

where ζ denotes the Lagrange multiplier. The first order conditions for the optimal loan rates, $\frac{\partial \mathcal{L}}{\partial r_L^1} = \frac{\partial \mathcal{L}}{\partial r_L^{2,H}} = \frac{\partial \mathcal{L}}{\partial r_L^{2,L}} = 0$, easily solve for

$$
\begin{aligned}
1+r_L^1 &= \frac{1+r_D}{\pi} + \frac{1-\zeta}{\zeta}\frac{L}{\frac{\partial L}{\partial\left(1+r_L^1\right)}} \\
&= 1+\hat{r}_L^1 + \frac{1-\zeta}{\zeta}\frac{L}{\frac{\partial L}{\partial\left(1+r_L^1\right)}}, \\
1+r_L^{2,H} &= \frac{1+r_D}{\pi_H} + \frac{1-\zeta}{\zeta}\frac{L_H}{\frac{\partial L_H}{\partial\left(1+r_L^{2,H}\right)}} \\
&= 1+\hat{r}_L^{2,H} + \frac{1-\zeta}{\zeta}\frac{L_H}{\frac{\partial L_H}{\partial\left(1+r_L^{2,H}\right)}}, \\
1+r_L^{2,L} &= \frac{1+r_D}{\pi_L} + \frac{1-\zeta}{\zeta}\frac{L_L}{\frac{\partial L_L}{\partial\left(1+r_L^{2,L}\right)}} \\
&= 1+\hat{r}_L^{2,L} + \frac{1-\zeta}{\zeta}\frac{L_L}{\frac{\partial L_L}{\partial\left(1+r_L^{2,L}\right)}}.
\end{aligned}
\tag{11.46}
$$

Inserting these loan rates into the constraint $\hat{\Pi}_B^{2,H} = \Pi_B$, we obtain that

$$
\frac{1-\zeta}{\zeta} = \frac{p\left(1-p\right)\left(\pi_H-\pi_L\right)\left(1+r_D\right)L_H}{\pi\frac{L^2}{\frac{\partial L}{\partial\left(1+r_L^1\right)}} + \pi^2 p\frac{L_H^2}{\frac{\partial L_H}{\partial\left(1+r_L^{2,H}\right)}} + \pi^2\left(1-p\right)\frac{L_L^2}{\frac{\partial L_L}{\partial\left(1+r_L^{2,L}\right)}}} < 0,
\tag{11.47}
$$

where the final inequality arises from the easily verified fact that $\frac{\partial L_i}{\partial\left(1+r_L^{t,i}\right)} < 0$. We obtain that $\frac{\partial \Pi_C}{\partial\left(1+r_L^1\right)} = -\pi L$, $\frac{\partial \Pi_C}{\partial\left(1+r_L^{2,H}\right)} = -\pi p \pi_H L_H$ and $\frac{\partial \Pi_C}{\partial\left(1+r_L^{2,L}\right)} = -\pi\left(1-p\right)\pi_L L_L$. Taking the derivatives with respect to L, L_H, and L_L, respectively, as well as the second derivative with respect to the loan rate, allows us to apply the implicit function theorem to show the negative sign of these expressions.

This implies that the second term of the loan rates in equation (11.46) will be positive. The first term represents the competitive loan rates for the two types of companies in time periods 1 and 2, respectively, and we see that with long-term contracts the loan rates are above competitive levels, allowing banks to exploit their information advantage, but this mark-up will remain constant over time if we assume that $L \approx L_H \approx L_L$. Hence there will be no increase in loan rates in time period 1 beyond that which results from the accumulated information of the initial bank. Thus banks do not only exploit their informational advantage in time period 2, but can anticipate these profits and charge a higher loan rate in time period 1. Companies are compensated for this higher initial loan rate by not having their loan rate increased that much in time period 2. As we imposed the constraint that $\hat{\Pi}_B^{2,H} = \Pi_B$, the profits banks extract from companies are identical regardless of the contract form, and hence the company profits will be identical, too. Thus companies will be indifferent between long-term contracts that do not allow companies to switch banks, relationship banking, and those that allow for switching, transaction banking.

Summary By providing a long-term contract that doe snot allow companies to switch banks after their initial investment, banks can spread out the profits they make from their informational advantage over the entire length of the contract. This will lead to a consistent mark-up of the loan rates above their competitive levels, in contrast to a situation where companies can switch banks and the banks are limited to exploiting their informational advantage only after they have accumulated information about the company. This would lead to a significant increase in the loan rate for those companies that are assessed to have high success rates, while a long-term contract would only adjust the loan rates to take into account the new information. Engaging in such a long-term contract, the bank commits itself to not exploit their informational advantage in time period 2 as the company is locked into the relationship. In exchange, banks will charge higher loan rates in time period 1, to compensate for their commitment.

We thus see that long-term contracts result in less volatile loan rates for companies, which might be attractive to companies seeking to show consistent profits over time. In addition, banks that are guaranteed to provide loans into the future will be more willing to invest into the accumulation of information about the company and thus be able to offer more accurate loan rates in the future that fully reflect the risks the company is taking. Such more accurate loan rates would also allow a more efficient allocation of funds as companies are incentivized to take the risks they are taking into account appropriately.

Reading Sharpe (1990)

Résumé

Relationship banking gives the incumbent bank an informational advantage over their competitors; this advantage can result in banks making excess profits by extracting more surplus from companies than would have been possible if banks were competing using the same information. Knowing that once an informational advantage has been established, banks will be able to make excess profits, they will compete to attract companies in the first place. To this effect they will use the future profits they generate from companies to offer them attractive initial loan conditions, resulting in low loan rates at the start of a relationship only for these loan rates to be raised later in order to recover any losses they may have made initially. Such an informational advantage can arise from either their competitors facing additional costs of acquiring the information the relationship bank already holds or by the mere fact that the bank has privileged access to the company, which cannot be replicated by their competitors.

Competition between banks is based on the possibility of companies being able to switch banks if the conditions offered by other banks are more favourable. However, often companies can face costs for such a switch of banks. Taking into account switching costs, it may not be profitable for high-risk companies to switch banks as the conditions they are being offered would be such that it can be better to not obtain a loan in the future at all. In such a situation, the quality of the companies seeking to switch banks improves, reducing adverse selection between banks and the reduced risk to banks from providing loans, in combination with the lower risk of the banks seeking such a switch of banks, will allow the loan rate to decrease despite switching costs increasing. Only once the switching costs become sufficiently high, will there be no further reduction in adverse selection and the relationship bank can increase loan rates, exploiting the fact that companies face increasing costs to switch banks.

A common implication of relationship banking is that banks seek to attract companies by offering initially attractive loans and then once they have captured the company increase the loan rates and exploit their informational advantage. Such increases in loan rates can be avoided if companies agree beforehand to not switch banks. This will allow banks to anticipate the future profits they are going to make once they have obtained the informational advantage and charge higher loan rates from the start. Loan rates will remain consistent over time, albeit never being fully competitive.

11.2 Lending decisions by banks

While loan conditions, most notable loan rates, are important to companies, it is certainly of equal importance to be able to access loans in the first place; a loan at low costs that is not offered to a company, will be of no value to them. With relationship banks holding superior information on companies, are able to gain access to this information during the course of the relationship,will enable them to make better informed decision on whether to grant a loan in the first place. Having favourable information on a company can lead to a situation where in relationship banking a loan is granted, which in transaction banking is not given as we will see in chapter 11.2.1. The additional costs relationship banks incur will have to be weighed against the benefits of readily accessing loans. The superior information of relationship banks can also lead to additional adverse selection in that companies might not be granted a loan and hence seek a loan from alternative banks with which they do not have a relationship. Here the relationship bank might offload a company it regards as not creditworthy to other banks who, due to the lack of accurate information, will grant a loan. As we will see in chapter 11.2.2, this will keep companies surviving for longer and accumulate more losses until they are finally liquidated. Similarly, banks may engage in evergreening as discussed in chapter 11.2.3 in order to allow companies to make investments that reduces losses of the bank on outstanding loans.

It is not only adverse selection between banks that affects companies seeking loans from relationship banks. A risk the company in relationship banking faces is that their bank might be willing but not able to grant a loan; such a situation might occur if the bank faces external pressures or constraints like liquidity shortages or capital constraints. In chapter 11.2.4 we will see how companies hedge this risk by having relationships with multiple banks. Rather than resorting to unaffected banks with which they are not in a relationship, companies will be happy to incur additional costs from having multiple relationships.

11.2.1 Access to loans

If banks have better information about companies, they might be able to provide loans that other, less well informed banks, would either deny or impose unprofitable conditions. This might enable companies to continue with their investments during difficult times, such as recessions, and allows the company to emerge from such time periods in a stronger market position by being able to pursue its investments. Having information about the quality of the company and having confidence in its ability to complete investments successfully, might make the bank willing to provide a loan with reasonable conditions, that would otherwise be denied.

Let us consider companies that seek a loan L for an investment that they will successfully complete after one time period with probability π. If successful, the company generates an outcome V and unsuccessful companies generate no outcome. If this investment is not successful, there is the possibility that it might succeed in the next time period, provided the loan is extended. If the company is highly skilled, it will generate an outcome V with certainty, but if the company is low-skilled, the company will not produce any outcome. Such a scenario is realistic if we think of a company that has initially made mistakes in their investment but has the ability to learn from such mistakes and turn the investment around due to their experience and managerial skills. We assume that there is a fraction p of highly skilled companies in the market. In essence, highly-skilled companies will always succeed, provided their loan is extended after an initial failure, while companies with lower skills would only succeed with probability π, and extending the loan would not remedy any investment failure.

Relationship banks will know the type of company after investing an amount C to acquire this information, while other banks will only know that there is a fraction p of such highly skilled companies in the market. This information is only revealed to them after the first time period, but before the decision whether to extent the loan is made. Likewise, the company will only then be able to know its type.

We can now compare the lending decisions of transaction banks, who do not invest into acquiring knowledge about the type of company, and relationship banks, who make this investment.

Transaction banks Let us first consider the decision of the bank whether to extend the loan. As the bank does not know whether the company is highly skilled or not, it will expect its loan to be repaid with probability p, such that its profits from the extended loan is given by

$$\Pi_B^2 = p\left(1 + r_L^1\right)\left(1 + r_L^2\right)L - \left(1 + r_L^1\right)(1 + r_D)\,L, \tag{11.48}$$

where r_L^t denotes the loan rate of the transaction bank in time period t and r_D the deposit rate, where we assume that banks finance the loan entirely by deposits. The total loan given to the company will include the accumulated interest from the first time period. If markets are competitive such that $\Pi_B^2 = 0$, then the loan rate is given by

$$1 + r_L^2 = \frac{1 + r_D}{p}. \tag{11.49}$$

Of course the company must be able to repay the loan if it is successful. Using its outcome V, we see that this is the vase only if $V - \left(1 + r_L^2\right) L \geq 0$, or $p \geq p^* = (1 + r_D) \frac{L}{V}$, after inserting for r_L^2 from equation (11.49). In markets where there are less than a fraction p^* of highly skilled companies, the transaction bank would not extend any loans as they could not be repaid in full even if the investment is successful.

For the loan in time period 1, we know that if the loan is extended after the initial investment failed, thus $p \geq p^*$, the initial loan will be repaid with the second, extended loan. If the investment is successful, the loan is repaid in any case. If the loan is not extended as $p < p^*$, the loan will only be repaid if the initial investment is successful. Thus we have the bank profits from lending in time period 1 given as

$$\Pi_B^1 = \begin{cases} \left(1 + r_L^1\right) - (1 + r_D) L & \text{if } p \geq p^* \\ \pi \left(1 + r_L^1\right) L - (1 + r_D) L & \text{if } p < p^* \end{cases}. \tag{11.50}$$

If we assume again that banks are competitive and $\Pi_B^1 = 0$, the loan rates are then given by

$$1 + r_L^1 = \begin{cases} 1 + r_D & \text{if } p \geq p^* \\ \frac{1+r_D}{\pi} & \text{if } p < p^* \end{cases}. \tag{11.51}$$

Companies generate profits and repay their loan if they are successful in time period 1, and if they are not successful in time period 1, they are successful in time period 2 if they are highly skilled and the loan is extended. If the loan is extended, the company repays its the loan with the interest accumulated from both time period. Thus we have

$$\begin{aligned} \Pi_C &= \begin{cases} \pi \left(V - \left(1 + r_L^1\right) L\right) & \\ \quad + (1 - \pi) p \left(V - \left(1 + r_L^1\right) \left(1 + r_L^2\right) L\right) & \text{if } p \geq p^* \\ \pi \left(V - \left(1 + r_L^1\right) L\right) & \text{if } p < p^* \end{cases} \quad (11.52) \\ &= \begin{cases} (\pi + (1 - \pi) p) V & \\ \quad - \left(\pi (1 + r_D) + (1 - \pi) (1 + r_D)^2\right) L & \text{if } p \geq p^* \\ \pi V - (1 + r_D) L & \text{if } p < p^* \end{cases}, \end{aligned}$$

where the second equality uses the loan rates from equations (11.49) and (11.51).

Having established the profits of companies when dealing with transaction banks, we can now continue to assess the profits they will be making when taking a loan from a relationship bank.

Relationship bank In time period 2 the relationship bank knows the type of company and would only extend the loan if it is highly skilled. However, the bank would not charge a competitive loan rate as it seeks to maximize its profits, but would match the loan rate of the transaction bank to prevent companies from switching banks, provided the loan is extended by them. If no loan is extended, then the relationship bank faces no competition and would extract all surplus from the company such that $V - \left(1 + \hat{r}_L^2\right) L = 0$, where $\hat{r}_L^2$ denotes the loan rater offered by the relationship bank. We thus have the loan rate in the second time period given by

$$1 + \hat{r}_L^2 = \begin{cases} \frac{1+r_D}{p} & \text{if } p \geq p^* \\ \frac{V}{L} & \text{if } p < p^* \end{cases}. \tag{11.53}$$

The bank's profits across both time periods are now consisting of four elements. Firstly, if the initial investment is successful, the bank is repaid its loan, including interest $\hat{r}_L^1$. If the initial investment is not successful and the company not highly skilled, then the loan is not repaid, and if the company is highly skilled the loan is extended and repaid with certainty. Finally, relationship banks facing costs C for obtaining the information on the type of company. Thus their profits are given by

$$\begin{aligned} \hat{\Pi}_B = &\pi \left(\left(1 + \hat{r}_L^1\right) L - (1 + r_D) L\right) - (1 - \pi)(1 - p)(1 + r_D) L \qquad (11.54) \\ &+ (1 - \pi) p \left(\left(1 + \hat{r}_L^1\right)\left(1 + \hat{r}_L^2\right) L - \left(1 + \hat{r}_L^1\right)(1 + r_D) L\right) \\ &- C. \end{aligned}$$

Assuming that relationship banks are competitive such that $\hat{\Pi}_B = 0$, then we get after inserting for the loan rate in time period 2, $\hat{r}_L^2$, from equation (11.53). the loan rate in time period 1 as

$$1 + \hat{r}_L^1 = \begin{cases} \frac{(1-p(1-\pi))(1+r_D)+\frac{C}{L}}{\pi+(1-\pi)(1-p)(1+r_D)} & \text{if } p \geq p^* \\ \frac{(1-p(1-\pi))(1+r_D)+\frac{C}{L}}{\pi+(1-\pi)p\frac{V}{L}-(1-\pi)p(1+r_D)} & \text{if } p < p^* \end{cases}. \tag{11.55}$$

The profits of the company are now given by

$$\hat{\Pi}_C = \pi\left(V - \left(1 + \hat{r}_L^1\right) L\right) + (1 - \pi) p \left(V - \left(1 + \hat{r}_L^1\right)\left(1 + \hat{r}_L^2\right) L\right), \tag{11.56}$$

where the relationship bank will extend the loan if the company is highly skilled. Inserting the loan rates from equations (11.53) and (11.55) would give us an explicit expression for these profits.

With the profits of companies using relationship banks having been established, we can now analyse whether the company prefers relationship or transaction banking.

Optimal bank choice If we compare the profits the company makes when using transaction banking, equation (11.52), and relationship banking, equation (11.56) after inserting the loan rates, we can distinguish the two cases of a large fraction of highly skilled companies, $p \geq p^*$, and a small fraction of highly skilled companies, $p < p^*$. Commencing with the case of a large number of highly skilled companies, we see that for companies to prefer transaction banking we require that these profits exceed that of relationship banking, thus $\Pi_C \geq \hat{\Pi}_C$. This condition solves for

$$\frac{C}{L} \geq (1-\pi)(1-p)(1+r_D)\, r_D. \tag{11.57}$$

Hence, as long as the information costs for banks are sufficiently high, transaction banking is preferred. We notice firstly that this constraint is becoming less bonding as the fraction of highly skilled companies, p increases and approaches zero costs as all companies become highly skilled. In general, we observe that this constraint is not very binding as in realistic scenarios with high success rates for companies, π, a large fraction of highly skilled companies p, and low deposit rates r_D, this expression will be small. Thus unless information costs to banks are very low, companies will prefer transaction banking if there is a large proportion of highly skilled companies in the market.

In the case of fewer highly skilled companies in the market, $p < p^*$, relationship banking is preferred if $\hat{\Pi}_C \geq \Pi_C$, solving for the condition that

$$\frac{C}{L} < \frac{\left((1-\pi)\, p\frac{V}{L} + (1+r_D)\right)\left(\pi + (1-\pi)\, p\left(\frac{V}{L} - (1+r_D)\right)\right)}{\pi + (1-\pi)\, p\frac{V}{L}} - (1 - p(1-\pi))(1+r_D). \tag{11.58}$$

Provided the information costs of banks are not too high, relationship banking is preferred by companies if there are fewer highly-skilled companies in the market. We note that for very few highly skilled companies, $p \approx 0$, this constraint becomes very restrictive as the expression on the right-hand side approaches zero, too. It is thus that if the fraction of highly skilled companies is very low, transaction banking is preferred, but then as the fraction of highly skilled companies increases, relationship banking becomes the preferred banking form.

The reason for our findings is that with a large fraction of companies being highly skilled, the benefits of relationship banking are small, transaction banks will not charge a loan rate that is prohibitively high and thus companies can always obtain an extension if the investment is initially is not successful. This extension is available without having to cover the information costs of banks, making the reliance on loans by transaction banks preferable. As the fraction of highly skilled companies reduces, the loan by transaction banks becomes too expensive and the company would not be able to secure a loan at conditions that would make continuing with the investment profitable. Thus they turn to a relationship bank, who will always extend their loan to companies that are highly skilled. The profits relationship banks make from this loan is fully extracted from companies, but used as a subsidy in the initial loan, given that banks overall are competitive, benefitting companies from lower loan rates in the first time period. Once the fraction of highly skilled companies is reduced even further, the likelihood of a company being highly skilled and thus a loan being extended, is so low that it is more costly to cover the information costs of banks than to forego a loan extension by relationship banks, and transaction banking becomes preferable again.

Summary We have seen that in markets with a large and very small fraction of highly skilled companies, transaction banking will dominate, while in markets with an intermediate fraction of highly skilled companies, relationship banking will dominate. For market in which most companies are highly skilled or very few companies are highly skilled, there is not much informational asymmetry between relationship banks and transaction banks. This low level of adverse selection will not allow relationship banks to generate enough profits to cover their additional costs without adversely affecting companies. It will be for intermediate levels of highly skilled companies that adverse selection is highest and the benefits of relationship banks over transaction banks are such that they can recover their information costs, while still offering better conditions to companies.

In markets that are either newly developing or rapidly changing, therefore skills for turning around initially failing investments will be scarce. In markets that are well established and understood, such skills will be quite common. In both cases we would companies to engage mostly in transaction banking. It is in markets that are established but where changes require considerable skills that are not that widespread that relationship banking is most likely to be found. Alternatively we might want to look at the experience of managers; managers that are not much experienced or those that have significant experience, will prefer transaction banking, while those managers that have some experience may opt for relationship banking.

Reading Bolton, Freixas, Gambacorta, & Mistrulli (2016)

11.2.2 Lending to not-creditworthy companies

Banks are providing loans to companies they think are sufficiently likely to repay their loan and for the risk of companies not being able to do so, are requesting a loan rate over and above their funding costs, mostly deposits, as compensation. If subsequently they obtain additional negative information about the company that would induce them to call in the loan rather than extending it, the bank may encourage the company to seek a loan at another bank. In this case, the initial bank is repaid its loan and has shifted the potential default of the company to another, less well informed bank.

Let us assume there are two types of companies that a bank might provide loans to. The first type of companies makes successful investments, allowing loans to be repaid, with probability π_H, of which there is a fraction p; the other type of companies is able to repay their loans with probability $\pi_L < \pi_H$ and these make up a fraction $1 - p$ of all companies. The average probability of a loan being repaid is $\pi = p\pi_H + (1 - p)\,\pi_L$.

Investments last for two time periods, but loans are provided for only a single time period and thus need to be rolled after the first time period. The type of company is initially not known to the bank, but only revealed after one time period. If, after learning its type, a bank decides to not roll over a loan and the company cannot secure a new loan from another bank, the investment gets liquidated, which generates a fraction of the initial investment and hence the company makes no profits.

If a new loan with another bank is secured, the company repays its loan to the initial bank, and the investment continues financed by the new bank; this new bank will not be aware of the type of company as it has not lent to this company before. For simplicity we assume that no interest accrues in the first time period and hence the initial loan would be repaid at face value.

We now assume that for companies with low success rates, the expected repayment to the bank, including interest, is less than the amount the bank would achieve from liquidating the investment, making it optimal for the initial bank to liquidate the loan rather than rolling it over. As the company with a low success rate would not be able to generate a profit in this case, it would seek a new loan from another bank. A company, whether it has a low or high success rate, would also seek to switch their bank if the loan rate they are offered elsewhere is lower than what the initial bank can offer.

Let us for now assume that companies cannot switch banks after the first time period but have to remain with the same bank, who in turn cannot change the loan rate once they learn the company's type. The profits of this bank, who initially will not know the company type, will be given by the average success rate of the loan being repaid and their repayments to depositors, such that

$$\Pi_B = \pi\left(1 + r_L\right) L - \left(1 + r_D\right) L, \tag{11.59}$$

where r_L denotes the loan rate applied and r_D the deposit rate; deposits finance the entire loan. We assume that banks are competitive such that $\Pi_B = 0$, and the loan rate will then be given as

$$1 + r_L = \frac{1 + r_D}{\pi}. \tag{11.60}$$

A company that switches banks, contains information about its type. Using Bayes'theorem we can now obtain the probability that a bank switching banks has a low success rate as follows

$$\begin{aligned} 1 - \hat{p} &= \text{Prob}\left(L|\text{switch}\right) \\ &= \frac{\text{Prob}\left(\text{switch}|L\right)\text{Prob}\left(L\right)}{\text{Prob}\left(\text{switch}\right)} \\ &= \frac{1 - p}{\text{Prob}\left(\text{switch}\right)}, \end{aligned} \tag{11.61}$$

where $\text{Prob}\left(\text{switch}|L\right) = 1$ as we had indicated above that any company with a low success rate would seek to switch banks. We furthermore have

$$\begin{aligned} \text{Prob}\left(\text{switch}\right) &= \text{Prob}\left(\text{switch}|L\right)\text{Prob}\left(L\right) \\ &\quad + \text{Prob}\left(\text{switch}|H\right)\text{Prob}\left(H\right) \\ &= \left(1 - p\right) + \text{Prob}\left(r_L > \hat{r}_L\right) p \\ &= 1 - p\left(1 - \text{Prob}\left(\hat{p} > p\right)\right) \\ &= 1 - p\text{Prob}\left(\hat{p} \leq p\right). \end{aligned} \tag{11.62}$$

We note in this transformation that companies with high success rates only switch banks if the loan rate of a new bank, $\hat{r}_L$, is lower than at the initial bank. Inserting this result into equation (11.61), we get

$$p\left(1 - Prob\left(\hat{p} \leq p\right)\right) - \hat{p}\left(1 - pProb\left(\hat{p} \leq p\right)\right) = 0. \tag{11.63}$$

As p is known, we can determine Prob $(\hat{p} \leq p)$ as being either 0 or 1. If $\hat{p} > p$, then Prob $(\hat{p} \leq p) = 0$ and equation (11.63) solves for $p = \hat{p}$, violating the assumption that $\hat{p} > p$. If $\hat{p} \leq p$ and therefore Prob $(\hat{p} \leq p) = 1$, equation (11.63) implies that $\hat{p}(1-p) = 0$. As long as $p < 1$ we will observe that $\hat{p} = 0$. Hence the company type switching banks is perfectly revealed as being that of a company with slow success rates and a new bank in a competitive market would lend such that its profits are given by $\hat{\Pi}_B = \pi_L\left(1 + \hat{r}_L\right)L - \left(1 + r_D\right)L = 0$, implying a loan rate to all companies with low success rates of $1 + \hat{r}_L = \frac{1+r_D}{\pi_L}$. As obviously $\hat{r}_L > r_L$ due to $\pi_L < \pi$, companies with high success rates would not switch banks and remain with their initial bank. The new bank provides this loan to a company with a low success rates as it is profitable to do so given the higher loan rate it can charge, compared to the initial bank which we assumed cannot adjust its loan rate after learning the type of company.

Thus we have a situation in which companies that have low success rates would be liquidated by their existing bank, but they obtain a loan, at worse conditions, from another bank. The initial bank knows that companies with low success rates will switch banks and the bank profits become

$$\Pi_B^* = p\pi_H\left(1 + r_L^*\right)L + (1-p)L - \left(1 + r_D\right)L. \tag{11.64}$$

The bank will know that if the company has a high success rate π_H, which happens with probability p, it will be repaid the loan with interest if the investment is successful and if the company has a low success rate, a fraction $1-p$ of companies, it will be repaid the loan with certainty, but without interest as the company switches banks. With competitive banks such that $\Pi_B^* = 0$, we have the loan rate then given as

$$1 + r_L^* = \frac{\left(1 + r_D\right) - (1-p)}{p\pi_H}. \tag{11.65}$$

We can now see that the loan rate anticipating the switch of banks by companies with low success rates is lower than the loan rate if companies cannot switch banks, $1 + r_L^* < 1 + r_L$, if $1 + r_D < \frac{p\pi_H + (1-p)\pi_L}{\pi_L}$. Hence if the deposit rate is not too high, allowing companies to switch banks will lower the initial loan rate. This is because the bank can be sure to receive the loan back from the companies with low success rates as they switch banks. The loan rate is also lower than the loan rate provided by the new bank, $1 + r_L^* < 1 + \hat{r}_L$, if $1 + r_D < \frac{\pi_L(1-p)}{\pi_L - p\pi_H}$ assuming $\pi_L > p\pi_H$. Hence with a sufficiently low deposit rate, companies with low success rates enjoy lower initial loan rates than justified by their type.

Banks can use a strategy of not extending loans to companies once they have established that they have low rates of success and thereby ensuring the premature repayment of the loan as the result of receiving a loan from another less well informed bank. This will benefit companies with high success rates who will be offered lower loan rates than if companies were not allowed to switch banks.Hence banks will be less cautious about providing loans and incentivize the provision of loans by other, less well-informed banks, to companies that are generally not creditworthy at the loan rate offered by them.

We thus see that companies who are not seen as creditworthy by their initial bank on the terms initially agreed and would therefore be liquidated, are able to secure a loan from another bank at worse conditions and prevent their investment being liquidated. While in our model companies can be identified as having low success rates, we can easily imagine that new banks might not be able to differentiate between companies of different types that easily. Not only will companies with low success rates seek to switch banks, but companies with high success rates might want to switch banks for other reasons, such as the level of service a bank provides them. This will then induce a mix of both company types seeking a new bank, lowering the loan rate new banks can offer. This may lead to a situation where companies with low success rates are being assessed as not creditworthy by their initial bank, can obtain a loan from another bank, even if they were not creditworthy at all, even at less favourable conditions. Thus the adverse selection between banks that arises due to relationship banking can prevent the timely liquidation of companies that are not creditworthy; instead such companies are able to survive for considerable time by obtaining loans from other banks that hold less informed about them.

Reading Hu & Varas (2021)

11.2.3 Evergreening

If a company has an outstanding loan that currently cannot be repaid, the bank can liquidate the company and obtain any funds from this liquidation, usually causing them a loss. Alternatively, the bank could extend another loan to the company in the anticipation that the investment the company makes using this new loan, can repay the outstanding loan at least partially and thus reduce the losses to the bank; extending such a loan is referred to as evergreening. In this a situation an otherwise bankrupt company is kept alive by banks with the aim of them reducing their losses, but if the new investment is not successful, the losses they will face, are increased.

Let us assume a company has a loan $\hat{L}$ outstanding that it currently cannot repay. It also has a new investment opportunity that would provide them with a return R if successful and no return if not successful; the probability of success is π. The company will obtain a new loan L and have its existing loan extended at a loan rate r_L. If the investment is successful the company will have to repay the new loan as well as the outstanding loan. Thus their profits are given by

$$\Pi_C = \pi \left((1+R) L - (1+r_L) \left(L + \hat{L} \right) \right) . \tag{11.66}$$

Companies will take this loan if its expected profits are positive, $\Pi_C \geq 0$, which requires a loan rate of no more than

$$1 + r_L \leq 1 + r_L^* = (1 + R) \frac{L}{L + \hat{L}}. \tag{11.67}$$

If the bank which has provided the outstanding loan $\hat{L}$ and now provides an additional loan L gets both loans repaid in full, its profits are given by

$$\Pi_B = \pi (1 + r_L) \left(L + \hat{L}\right) - (1 + r_D) \left(L + \hat{L}\right), \tag{11.68}$$

where r_D denotes the deposit rate and we assume that loans are fully financed with deposits. If the bank does not provide the new loan, it would lose the outstanding loan and would thus provide the new loan if $\Pi_B \geq -\hat{L}$, which requires a loan rate of at least

$$1 + r_L \geq 1 + r_L^{**} = \frac{(1 + r_D) \left(L + \hat{L}\right) - \hat{L}}{\pi \left(L + \hat{L}\right)}. \tag{11.69}$$

If both loans cannot be repaid in full, the bank will obtain the entire revenue of the company, $(1 + R) L$ such that its profits are then

$$\Pi_B = \pi (1 + R) L - (1 + r_D) \left(L + \hat{L}\right) \tag{11.70}$$

and the condition to provide a loan, $\Pi_B \geq -\hat{L}$ yields

$$\hat{L} \leq \hat{L}^* = \frac{\pi (1 + R) - (1 + r_D)}{r_D} L. \tag{11.71}$$

As we can show that $1 + r_L^* \geq 1 + r_L^{**}$ if $\hat{L} \leq \hat{L}^*$, we see that the existing bank would provide a new loan, which the company is accepting, as long as the outstanding loan is not too large. With banks maximizing profits, and if the existing bank does not face competition from new banks offering loans to the company as we will introduce below, banks will charge the highest possible loan rate of $r_L = r_L^*$.

The existing bank will face competition from other banks that do not have an outstanding loan with this company. If they provide a loan, they are not concerned about the repayment of the outstanding loan directly. This new loan is repaid in full if the return of the company from their investment exceeds the funding costs of this loan as well as the repayment of the outstanding loan, thus if $(1 + R) L \geq (1 + r_L) L + \hat{L}$, or $\hat{L} \leq \hat{L}^{**} = (R - r_L) L$. In this case the bank profits of the new bank are given as

$$\hat{\Pi}_B = \pi (1 + r_L) L - (1 + r_D) L \tag{11.72}$$

and the loan is given as long as $\hat{\Pi}_B \geq 0$, or

$$1 + r_L \geq 1 + \hat{r}_L^* = \frac{1 + r_D}{\pi}. \tag{11.73}$$

As we can show that $1 + r_L^* \geq 1 + \hat{r}_L^*$ if $\hat{L} \leq \hat{L}^{**} = \frac{\pi(1+R)-(1+r_D)}{1+r_D} L$, a new bank would be willing to give a loan as long as the outstanding loan was sufficiently small.

If, on the other hand, $\hat{L} > \hat{L}^{**}$, such that the revenue of the company is not sufficient to repay both loans, the loans are repaid pro-rata using the revenue the company has produced. Thus the bank profits in this case are

$$\hat{\Pi}_B = \pi (1 + R) \frac{L}{L + \hat{L}} L - (1 + r_D) L \tag{11.74}$$

and the loan is given as long as $\hat{\Pi}_B \geq 0$, or

$$\hat{L} \leq \hat{L}^{**} = \frac{\pi (1 + R) - (1 + r_D)}{1 + r_D} L. \tag{11.75}$$

This constraint is identical to the constraint where companies willing to use a loan at loan rate $1 + r_L = \frac{1+r_D}{\pi}$ and hence does not provide a further constraint on the provision of loans.

If we assume that new banks are competitive, they will set a loan rate of $r_L = \hat{r}_L^*$ and that they make no profits. The existing bank could offer a lower loan rate than the one set by new banks, it would not do so in order to maximize its profits and match the loan rate of new banks.

We thus see that if $\hat{L} \leq \hat{L}^{**}$, new banks would provide a loan to the company and we can interpret this as a situation in which the company is seen generally as creditworthy. If $\hat{L}^{**} < \hat{L} \leq \hat{L}^*$, only the existing bank would provide the company with a new loan to reduce its losses on the outstanding loan, thus the bank evergreens the outstanding loan. For $\hat{L} > \hat{L}^*$ the company would not obtain a new loan and instead be liquidated due to not being able to repay its outstanding loan. If a new loan is extended due to evergreening, we can easily confirm that $r_L^* < \hat{r}_L^*$, implying that the loan rate the company obtains is lower than what a new bank would charge to break even. These more favourable loan conditions are offered to ensure the company accepts the loan, allowing banks to recover some of their losses.

We thus observe that evergreening occurs for an intermediate range of outstanding loans the company is unable to repay in its current situation. If the outstanding loans are sufficiently small, then any bank would be willing to extent a loan to this company and it is generally seen as creditworthy. For large outstanding loans, the recovery of the outstanding loans through new investments conducted with the help of new loans is sufficiently unlikely to be beneficial; the existing bank would expect to increase its losses and thus not provide a new loan. In an intermediate range the company is not seen as creditworthy by banks not having extended loans previously, but with a bank being exposed to an outstanding loan it will extend a new loan in order to recover some of the losses through the profits the company makes on the investment it conducts using this new loan.

Companies that are failing to repay a loan may be extended a new loan with the aim of them making an investment that recovers at least some of the losses that banks have made. In this way companies are only liquidated later, even though they are not creditworthy to an outside lender. This can prolong the process for failing companies to be recognised as such and other creditors with less knowledge about the prospects of the company might unwittingly incur additional costs, for example if extending trade credits to the company.

Reading Faria-e-Castro, Paul, & Sánchez (2024)

11.2.4 The optimal number of relationships

It is common that companies so not have relationship with a single bank, but with multiple banks. On the one hand this will allow competition between relationship banks to provide future loans and thus reduce the informational advantage of banks; this should lead to more competitive future loan rates. However, companies may also face a situation in which a bank may not be able or willing to advance further loans, despite the relationship a company has with the bank. One reason might be that the risk assessment for a suggested investment is not favourable and the bank denies a loan on these grounds. If another relationship bank comes to a different conclusion, the company would still be able to secure a loan. It might also be the case that a bank might not be able to provide a loan as it faces constraints on its liquidity, capital requirements, or restrictions on the exposure to a single company. Another relationship bank might not face the same constraints and would be able to advance the loan and the company to conduct its investment.

Let us assume that there are two types of companies, one whose investments succeed with a probability of π_H and another type of company who only succeeds with a probability of $\pi_L < \pi_H$. If there is a fraction p of companies that have a high success rate and a fraction $1 - p$ that have a low success rate, the average success rate of companies is given by $\pi = p\pi_H + (1 - p)\,\pi_L$. Each company requires a loan L that finances its investment in full, which generates an outcome V if successful and no outcome if not successful. The loan L is split equally between each of their N relationship banks, thus each bank advances an amount of $\frac{L}{N}$. The type of company is initially not known to the company itself or the bank, but it is revealed to companies after one time period and a relationship bank will learn its type at some cost C. Banks that are not a relationship bank to this company, will not learn its type. We assume that investments last for two time periods and loans need to be rolled over after the first time period.

Banks will not be able to roll over the loan with probability λ, reflecting a liquidity shortage, capital constraints or an adverse assessment of the companies future prospects. If any of the relationship banks face such a situation, the company is able to roll over the loan with any other of the remaining $\hat{N}$ relationship banks that can roll over the loan. If none of their relationship banks is able to toll over their loan, they will have to obtain a loan from other banks, transaction banks. In addition, companies that have a low success rate are not able to secure a roll over of their loan as this would not be profitable to the bank, it would have to rely on other banks, transaction banks, to continue financing their investment.

We thus see that companies with high success rates will switch from a relationship bank if none of their N relationship banks can roll over the loan, λ^N and all companies with low success rates, $1 - p$, will switch to another bank. Hence the probability of companies switching to transaction banks is given by

$$\hat{\lambda} = p\lambda^N + (1 - p). \tag{11.76}$$

An uninformed bank would have to infer the success rate of those companies that switch out of relationships as this comprises companies that are having a low success rate and those that have a high success rate but have not been able to obtain a loan from their relationship banks. The success rate of companies switching banks is given by

$$\hat{\pi} = \frac{p\pi_H\lambda^N + (1 - p)\pi_L}{p\lambda^N + (1 - p)}. \tag{11.77}$$

The numerator of this expression consists of all companies, p, with high success rates, π_H, not being able to extend their loans, λ^N, and all companies $1 - p$ with low success rates; the denominator reflects the fraction of companies switching as defined in equation (11.76).

We can now analyse this model backwards and assess the decision by banks to roll over loans.

Decisions to roll over loans After the first time period, relationship banks will have learnt the type of company they are lending to, at some costs C. The original loan L will have accumulated interest r_L^1 from the first time period, which we assume is accumulated into the loan. This the total amount the bank needs to roll over to the bank is $\left(1 + r_L^1\right) L$ and the bank is repaid the loan including interest r_L^2 if the company succeeds with its investment. With each relationship bank advancing a loan of $\frac{L}{\hat{N}}$ due to some banks not being able to roll over the loan, we get the profits of those banks able to roll over the loan as

$$\hat{\Pi}_B^{2,i} = \pi_i \left(1 + r_L^{2,i}\right)\left(1 + r_L^1\right)\frac{L}{\hat{N}} - (1 + r_D)\left(1 + r_L^1\right)\frac{L}{\hat{N}} - C, \tag{11.78}$$

where r_D denotes the deposit rate for the deposits that fully finance the loan. If banks are competitive, we will have $\hat{\Pi}_B^{2,i} = 0$ and the loan rate of relationship banks when rolling over the loan is given as

$$1 + r_L^{2,i} = \frac{1 + r_D}{\pi_i} + \frac{\hat{N}}{1 + r_L^1}\frac{C}{L}. \tag{11.79}$$

$\hat{N}$ represents the number of relationship banks that roll over the loan. Each bank has to make inferences about the number of other banks rolling over the loan to assess the size of the loan they have to provide. The bank knows itself to be lending, but for the remaining $N - 1$ relationship banks will need to assign a probability. That exactly i banks, out of the remaining $N - 1$ banks, are rolling over the loan will be those i banks facing no constraints, $(1 - \lambda)^i$, while the remaining banks will face such constraints, λ^{N-1-i}; there are a possible $\begin{pmatrix} N - 1 \\ i \end{pmatrix}$ permutations of banks for this scenario. The expected number banks rolling over the loan are thus given by

$$\hat{N} = 1 + \sum_{i=0}^{N-1} i \binom{N-1}{i} \lambda^{N-1-i} (1-\lambda)^i \tag{11.80}$$
$$= 1 + (N-1)(1-\lambda),$$

where the first term arrives from the fact that the active bank itself knows to be lending and the second equality acknowledges that this value is the expected value of a binomial distribution.

New banks providing loan to companies that switch from their relationship bank infer a success rate $\hat{\pi}$ as defined in equation (11.77) and thus its profits are given by

$$\hat{\Pi}_B^2 = \hat{\pi}(1 + \hat{r}_L^2)\left(1 + r_L^1\right) L - (1 + r_D)\left(1 + r_L^1\right) L. \tag{11.81}$$

If banks are competitive such that $\hat{Pi}_B^2 = 0$ the loan rate these banks require is given by

$$1 + \hat{r}_L^2 = \frac{1 + r_D}{\hat{\pi}}. \tag{11.82}$$

We now assume that $\left(1 + r_L^2\right)\left(1 + r_L^1\right) L > V$, meaning that the total repayment of companies having to switch banks exceed the outcome of their investment. This implies that companies switching banks will not be able to secure a loan.

As we can easily establish that $\pi_L \leq \hat{\pi} \leq \pi \leq \pi_H$, it is obvious that companies with low success rates would prefer to switch banks as for them $\hat{r}_L^2 < r_L^{2,L}$. Comparing the loan rates in equations (11.79) and (11.82) for companies with high success rates, we see that they seek to remain with their bank if

$$C \leq \frac{(1 + r_D)\left(1 + r_L^1\right)}{\hat{N}} \left(\frac{1}{\hat{\pi}} - \frac{1}{\pi_H}\right) L. \tag{11.83}$$

If the costs of relationship banks obtaining information is too high, the loan rate would have to increase so far that it outweighs the benefits of a the identified higher success rate and the company would seek to switch banks to avoid the higher loan rate.

A relationship bank would only invest into the information acquisition if this is less profitable; given we assumed banks to be competitive, this would imply that not acquiring information should be loss-making. If not acquiring information, bank will have to charge the same loan rate to companies of either type as they cannot distinguish them anymore. Such a loan rate would however, also attract companies with low success rates. To ensure banks are informed we thus need

$$\hat{\hat{\Pi}}_B = \pi\left(1 + \hat{r}_L^2\right)\left(1 + r_L^1\right)\frac{L}{\hat{N}} - (1 + r_D)\left(1 + r_L^1\right)\frac{L}{\hat{N}} < 0, \tag{11.84}$$

which after inserting from equation (11.79) for the loan rate as charged for companies with high success rates, becomes

$$C < \frac{1}{\hat{N}} \left(\frac{1}{\pi} - \frac{1}{\pi_H} \right) (1 + r_D) \left(1 + r_L^1 \right) L. \tag{11.85}$$

As $\pi > \hat{\pi}$, the constraint in equation (11.85) is more restrictive than the constraint in equation (11.83), hence if relationship banks acquire information, companies with low success rates do seek to maintain the relationship. We assume that this constraint is fulfilled.

We can now turn to the initial decision by banks to provide loans and by companies to seek an optimal number of relationship banks.

Initial lending and borrowing The loans of companies with high success rates will be rolled over, provided that bank is able to do so. Thus from a bank's perspective in the first time period, a loan is repaid after the first time period if not all of the initial banks are facing a constraints, $1 - \lambda^N$. If all relationship banks face such constraints and the company cannot roll over its loan, it will not be able to repay its loan; this is because we had assumed above that companies switching banks cannot secure a loan from a new bank. Thus the profits of a bank from lending in the first time period is given by

$$\Pi_B^1 = \left(1 - \lambda^N \right) p \left(1 + r_L^1 \right) \frac{L}{N} - (1 + r_D) \frac{L}{N} \tag{11.86}$$

If banks are competitive such that $\Pi_B^1 = 0$, the loan rate in time period 1 is given by

$$1 + r_L^1 = \frac{1 + r_D}{p \left(1 - \lambda^N \right)}. \tag{11.87}$$

As companies with low success rates never get their loans rolled over, they will never be profitable. Companies obtain their outcome and repay their loans only if they are having high success rates, p, which they are not aware of in time period 1, are successful, π_H, and at least one of their relationship banks extends the loan, $1 - \lambda^N$. Thus company profits are given by

$$\Pi_C = p\pi_H \left(1 - \lambda^N \right) \left(V - \left(1 + r_L^1 \right) \left(1 + \hat{r}_L^{2,H} \right) L \right). \tag{11.88}$$

After inserting for the loan rates $1 + r_L^1$ and $1 + \hat{r}_L^{2,H}$ from equations (11.79) and (11.87), this becomes

$$\Pi_C = p\pi_H \left(1 - \lambda^N \right) \left(V - \hat{N} C \right) - (1 + r_D)^2 L. \tag{11.89}$$

Noting the expression for $\hat{N}$ from equation (11.80), we obtain the optimal number of relationship banks to be given by the first order condition

$$\frac{\partial \Pi_C}{\partial N} = p\pi_H \left(-\lambda^N \ln \lambda \left(V - C \left((N-1)(1-\lambda) + 1 \right) \right) - \left(1 - \lambda^N \right) (1-\lambda) C \right) = 0. \tag{11.90}$$

We can now analyse the properties of the slution of this first order condition. We have, noting that $\ln \lambda < 0$, the following second order derivatives

$$\begin{aligned}
\frac{\partial^2 \Pi_C}{\partial N^2} &= -p\pi_H \left(\lambda^N (\ln \lambda)^2 \left(V - C\hat{N} \right) - \lambda^N \ln \lambda C (1-\lambda) \right. \\
&\quad \left. -\lambda^N (1-\lambda) \ln \lambda C \right) < 0, \\
\frac{\partial^2 \Pi_C}{\partial N \partial C} &= p\pi_H \left(\lambda^N \ln \lambda \hat{N} L - \left(1 - \lambda^N \right) (1-\lambda) L \right) < 0, \\
\frac{\partial^2 \Pi_C}{\partial N \partial V} &= -\lambda^N \ln \lambda > 0, \\
\frac{\partial^2 \Pi_C}{\partial N \partial \lambda} &= -\pi_H N p \lambda^{N-1} \left(C (1-\lambda) \left(\lambda^N - 1 \right) \right. \\
&\quad \left. -C \left((N-1)(1-\lambda) + 1 \right) + V \right) \ln \lambda \\
&\quad -\pi_H p \lambda^N \left(-C \left(\lambda^N - 1 \right) + CN (1-\lambda) \lambda^{N-1} \right. \\
&\quad \left. -C (1-N) \right) \ln \lambda \\
&\quad -\pi_H p \lambda^{N-1} \left(C (1-\lambda) \left(\lambda^N - 1 \right) \right. \\
&\quad \left. -C \left((N-1)(1-\lambda) + 1 \right) + V \right) > 0.
\end{aligned} \tag{11.91}$$

Using the implicit function theorem, we then have the relationships between the optimal number of relationship banks and various parameters given as

$$\begin{aligned}
\frac{\partial N}{\partial C} &= -\frac{\frac{\partial^2 \Pi_C}{\partial N \partial C}}{\frac{\partial^2 \Pi_C}{\partial N^2}} < 0, \\
\frac{\partial N}{\partial \lambda} &= -\frac{\frac{\partial^2 \Pi_C}{\partial N \partial \lambda}}{\frac{\partial^2 \Pi_C}{\partial N^2}} > 0, \\
\frac{\partial N}{\partial V} &= -\frac{\frac{\partial^2 \Pi_C}{\partial N \partial V}}{\frac{\partial^2 \Pi_C}{\partial N^2}} > 0.
\end{aligned} \tag{11.92}$$

We thus observe that the optimal number of relationship banks reduces with costs of banks acquiring information as the loan rate for the rolled over loan increases as the costs of fewer relationship banks need to be recovered, offsetting the increased risk of not having the loan rolled over. As a hedge against banks facing constraints on rolling over loans, companies seek out more banks if the probability of such an event increases. Similarly do they try to protect their higher outcome V against an early closure of their investment by having more relationship banks.

Provided the costs of information acquisition for banks are low, a numerical analysis easily shows that the optimal number of relationship banks will be small, in line with actual observations, where companies have relationships with a small number of banks only.

Summary Companies seek to establish relationships with multiple banks, typically a small number of such banks. Even though each bank will provide smaller loan to the company and thus have to recover their fixed costs through higher loan rates, these increased costs are outweighed by the company facing reduced risks of their loan not being rolled over and investment outcomes not being realised as a result. An implication of this finding is that in markets where banks are easily constrained in their ability to provide new or roll over existing loans, companies would have more relationships. This might be the case in banking systems that are under stress, such as banks having tight liquidity margins or a low capitalisation that both might require them to reduce lending to companies.

Similarly, companies that make investments that are highly profitable will want to engage in more relationships with banks as a hedge against the risk of loans not being extended and the investment not being able to be realised. Hence companies or entire industries that are highly profitable would see a larger number of relationship banks than industries or companies that are less profitable. If the collection and processing of information about companies is more costly, for example as the result of more complex businesses or the reliance on informal processes, we would observe less relationship banking.

Reading Detragiache, Garella, & Guiso (2000)

Résumé

If adverse selection between relationship banks and their competitors is sufficiently large, it will be beneficial for companies to enter into a relationship with a bank. While the loan rate offered might be higher due to their bank facing higher costs from accumulating and processing more information than other banks, the added information might well ensure that the company can obtain a loan from their bank, where other banks would judge them to be too risky to be able to offer a loan. These benefits are less prominent when adverse selection between banks is lower, for example in situations where the risks of companies to successfully complete an investment is easily assessed, and the additional costs of relationship banks to gain more precise information are not outweighed by the benefits this information generates. In these cases, transaction banking would be preferred as the costs of information gathering are significantly reduced, allowing for lower loan rates to the company.

While additional information will allow companies easier access to additional loans, assuming the information is positive, the opposite effect is also present. If a bank holds negative information about a company, it may not grant a loan and the company may seek a loan from another, less well informed bank. This poses the problem that a company which is deemed to be not creditworthy given a full set information, can well be granted a loan by a bank with less complete information. This can easily lead to a situation where companies are borrowing beyond what is desirable due to it not being profitable for banks to access all available information. This can lead to excessive borrowing and if the company eventually fails, it will cause more widespread losses to banks. With evergreening, banks do not liquidate companies in the hope that by providing them with loans at very favourable conditions, they can recover some of their losses by extracting additional profits the company might make when allowed additional investments.

Relying on access to loans through relationship banking can be optimal for companies, however, banks are not always able to provide loans. They might be prevented from doing so by a range of other concerns; for example if their loan books is sufficiently large such that their capital requirements are becoming a constraint on their ability to grant loans. companies might not be able to access loans by this bank, negating the benefits of relationship banking. Similarly, banks may face liquidity shortage and be reluctant to grant additional loans out of concern for reducing their liquidity position even further. Once a company has a sizeable amount of loans outstanding with a single bank, it might be that the exposure to this company is sufficiently large for a bank to affect its ability to meet regulatory requirements. In such situations, companies would only be able to resort to loans by other banks with which they do not have a relationship and be often granted less favourable conditions. To prevent such a scenario, it would often be optimal for companies to have relationships with multiple banks. While each bank will face additional information costs, thus increasing loan rates to cover such costs due to companies borrowing less often from them, the additional certainty of being able to access a loan from at least one of their banks will outweigh these costs.

11.3 The effect of competition

Relationship banking provides banks with an informational advantage over other banks which they can use to generate excess profits. While other sources of profits can be diminished through competition, this is does not affect their informational advantage. However, relationship banking is usually costly to banks who need to constantly collect information about the company and maintain processes to analyse this information. With other sources of profits eroding as competition increases, this might make relationship banking unviable. We will evaluate in chapter 11.3.1 how the presence of adverse selection affects the ability to sustain relationship banking and chapter 11.3.2 looks at the optimal investment banks should make into relationship banking as competition increases.

11.3.1 Adverse selection and competition

Competition between banks should reduce their ability to generate profits and extract surplus from companies. On the other hand, the informational advantage a bank has over its competitors should not be subject to these competitive forces as other banks cannot replicate the information a bank has obtained from the relationship with their company.

Let us assume that there are two types of companies. One type of companies has a high success rate for their investments, π_H, while the other type of companies has a low success, $\pi_L < \pi_H$. The company knows its type, while a bank only learn about the type once it has lent to the company and established a relationship. Other banks will only be able to establish the average success rate, which is given by $\pi = p\pi_H + (1 - p)\,\pi_L$, where p denotes the fraction of companies with high success rates.

If we assume that loans are fully financed by deposits on which interest r_D is payable, then a bank which has no relationship with the company would make profits of

$$\Pi_B = \pi\,(1 + r_L)\,L - (1 + r_D)\,L, \tag{11.93}$$

where r_L denotes the loan rate this bank charges. If we allow banks to make profits, thus not be fully competitive, the loan rate this bank charges will be given by

$$1 + r_L = \frac{1 + r_D}{\pi} + \frac{\Pi_B}{\pi L}. \tag{11.94}$$

A bank having established a relationship with this bank will face additional costs C of maintaining this relationship, and knowing the type of company they are lending to, their profits become

$$\hat{\Pi}_B^i = \pi_i\left(1 + \hat{r}_L^i\right)L - (1 + r_D)\,L - C, \tag{11.95}$$

where $\hat{r}_L^i$ denotes the loan rate the bank applies to this company. With banks able to generate profits, the loan rate would then be given by

$$1 + \hat{r}_L^i = \frac{1 + r_D}{\pi_i} + \frac{\hat{\Pi}_B^i + C}{\pi_i L}. \tag{11.96}$$

Both types of banks, those that have provided a loan to the company and those that have not provided a loan to the company, will be competing to provide the next loan. Equation (11.94) shows is the lowest loan rate a new bank can offer, given a certain level of profits are generated, and banks having established a relationship with the company, will not undercut this loan rate as they seek to charge the highest possible loan rate as long as they can provide the loan. Setting $1 + \hat{r}_L^i = 1 + r_L$, we get the profits of the current bank given as

$$\hat{\Pi}_B^i = \left(\frac{\pi_i}{\pi}\Pi_B - C\right) + \pi_i\left(\frac{1}{\pi} - \frac{1}{\pi_i}\right)(1 + r_D)\,L. \tag{11.97}$$

The informational advantage of banks already providing a loan to the company means they can generate profits. This informational advantage cannot be competed away by other banks, it is only the part of the profits that arise due to new banks not being fully competitive that can be eroded through competition. In equation (11.97) this part of the profits the existing bank generates is represented in the first term and the second term shows the profits generated from the informational advantage.

Assume that competition between banks to attract new companies is such that the initial bank can only retain a fraction $1-\theta$ of the profits not associated with their informational advantage. Thus the profits of the initial bank become

$$\hat{\hat{\Pi}}_B^i = (1-\theta)\left(\frac{\pi_i}{\pi}\Pi_B - C\right) + \pi_i\left(\frac{1}{\pi} - \frac{1}{\pi_i}\right)(1+r_D)\,L. \tag{11.98}$$

Let us assume that $\Pi_B = 0$ as there are a large number of such banks competing on equal terms. Focussing on companies with high success rates, we see that banks prefer to establish a relationship with a company if the profits generated are exceeding that of a bank offering loans without such a relationship. Hence we require

$$\hat{\hat{\Pi}}_B^H \geq \Pi_B = 0, \tag{11.99}$$

which we can solve for

$$\frac{\pi_L}{\pi_H} \leq \xi^* = \frac{(1+r_D)(1-p) - (1-\theta)\,p\frac{C}{L}}{(1-p)\left(1 + r_D + (1-\theta)\frac{C}{L}\right)}. \tag{11.100}$$

We can now interpret $\frac{\pi_L}{\pi_H}$ as the degree of asymmetric information between banks; a larger difference between companies of different types, corresponding to a lower value of this ratio, increases the value of knowing this type. Hence banks seek to enter relationships with companies if the level of asymmetric information is sufficiently large.

We easily obtain that

$$\frac{\partial\xi^*}{\partial\theta} = \frac{C}{L}\frac{1 + r_D - (1-\theta)(1-p)\frac{C}{L}}{(1-p)\left(1 + r_D + (1-\theta)\frac{C}{L}\right)^2} > 0, \tag{11.101}$$

$$\frac{\partial\xi^*}{\partial C} = -\frac{(1-\theta)(1-p)\,p\left(1 + r_D + (1-\theta)\frac{C}{L}\right) + (1-\theta)}{(1-p)\left(1 + r_D + (1-\theta)\frac{C}{L}\right)^2} < 0,$$

$$\frac{\partial\xi^*}{\partial p} = -\frac{r_D + (1-\theta)\,p}{(1-p)(1 + r_D + (1-\theta)\,p)} < 0.$$

As competition increases, θ, the adverse selection threshold at which relationship banking become feasible reduces, making its emergence more likely. The reason for this observation is that with increased competition, profits of banks are under pressure and banks can only make additional profits by gaining an informational advantage, even though this will cost them C. The larger their informational advantage, the larger the difference between the types of companies, the more profits they can generate.

Thus we find that the more competitive markets are, the more important this source of profits becomes.

Of course, increasing the costs of relationship banking will reduce its attractiveness and a higher degree of adverse selection needs to be present if the profits obtained from their informational advantage are to be recovered. A larger fraction of companies with high success rates, p, will make relationship banking less attractive as the adverse selection is reduced due to less companies with low success rates being active in the market.

If competition is perfect, $\theta = 1$, we see that relationship banking is always chosen as the condition in equation (11.100) reduces to $\xi^* \leq 1$, which is trivially fulfilled as we assumed that $\pi_L < \pi_H$. On the other hand, if $\frac{C}{L} > \frac{(1+r_D)(1-p)}{(1-\theta)p}$, the condition becomes $\xi^* < 0$ and relationship banking is never optimal. Thus if the costs of relationship banking are too high, it cannot emerge.

As a consequence, we should find relationship banking in markets where adverse selection is high, either because the differences in the risks companies are exposed to vary significantly or because low-risk companies are not very frequent. In such an environment the informational advantage is sufficiently high so that banks can generate profits that exceed the costs that relationship banking may impose on banks.

Reading Boot & Thakor (2000)

11.3.2 Investment into relationship banking

Relationship banking imposes costs on banks due to the continued need to accumulate and process information. Thus banks need to make an investment into relationship and such investment may yield diminishing returns. The more companies they provide relationship banking to, the lower the return would be as the benefits from gaining informational advantage will decrease the more companies are included.

Banks invest an amount C into relationship banking and this allows them to provide relationship loans to a fraction ρ of borrowers, where $\frac{\partial C}{\partial \rho} > 0$ and $\frac{\partial^2 C}{\partial \rho^2} < 0$. As the number of loans in relationship banking increases, the costs increasing to provide the systems that allow banks to accumulate and process the information. However, there are economies of scale and as the number of loans they provide increases, the marginal costs are reducing. The bank knows the type of company seeking a loan if they are in a relationship, while the remaining loans, $1 - \rho$, are loans that are provided as a transaction bank, without the bank knowing the type of company.

Companies invest into a project yielding on outcome V if successful and no outcome otherwise. There are two types of companies, one having a high probability of success π_H, and the other a low probability of success $\pi_L < \pi_H$. A fraction p of companies are having a high probability and we define the average probability of success as $\pi = p\pi_H + (1 - p)\,\pi_L$.

If a relationship loan is given, the bank knows the type of company they are providing a loan to. Hence their profits from this loan are given by

$$\Pi_B^{i,R} = \pi_i \left(1 + \hat{r}_L^i\right) L - (1 + r_D) L, \tag{11.102}$$

where $\hat{r}_L^i$ denotes the loan rate given to a company of this type and loans are fully financed by deposits, on which interest r_D is payable. If the market were competitive, bank would make no profits, $\Pi_B^{i,R} = 0$, and the loan rate charged would be

$$1 + \hat{r}_L^i = \frac{1 + r_D}{\pi_i}, \tag{11.103}$$

giving rise to company profits of

$$\Pi_C^{i,R} = \pi_i \left(V - \left(1 + \hat{r}_L^i\right) L\right) = \pi_i V - (1 + r_D) L. \tag{11.104}$$

This profit is the maximum profit available to companies as banks charge the lowest possible loan rate to break even. If competition is imperfect, banks will be able to extract some surplus from companies and their profits will reduce accordingly. Let us assume that companies only obtain a fraction θ of their maximum profits as defined in equation (11.104) and we can interpret θ as the level of competition in the market. The actual profits that companies will obtain are thus given by

$$\hat{\Pi}_C^{i,R} = \theta \Pi_C^{i,R}. \tag{11.105}$$

The loan rate for such relationship loans will be given such that the profits companies make, $\pi_i \left(V - \left(1 + \hat{r}_L^i\right) L\right)$, are equal to $\hat{\Pi}_C^{i,R}$. This gives us a loan rate from relationship loans of

$$1 + \hat{r}_L^i = \theta \frac{1 + r_D}{\pi_i} + (1 - \theta) \frac{V}{L}. \tag{11.106}$$

Using the bank profits as defined in equation (11.102) and inserting the loan rate from equation (11.106), we get the profits of relationship banks as

$$\hat{\Pi}_B^{i,R} = (1 - \theta) \left(\pi_i V - (1 + r_D) L\right). \tag{11.107}$$

We can now repeat the same steps for transaction loans. For a transaction bank, who does not know the type of company they are lending to, the profits for a loan of size L are given by

$$\Pi_B^T = \pi (1 + r_L) L - (1 + r_D) L, \tag{11.108}$$

where banks charge a loan rate r_L. The profits of a company receiving such a loan is given by

$$\Pi_C^{i,T} = \pi_i \left(V - (1 + r_L) L\right). \tag{11.109}$$

If banks are competitive, $\Pi_B = 0$, we have from equation (11.108) the loan rate given by

$$1 + r_L = \frac{1 + r_D}{\pi} \tag{11.110}$$

and hence company profits are

$$\Pi_C^{i,T} = \pi_i V - \frac{\pi_i}{\pi}(1 + r_D) L. \tag{11.111}$$

If we again assume that competition will be imperfect and companies can retain only a fraction θ of their profits, we obtain

$$\hat{\Pi}_C^{i,T} = \theta \Pi_C^{i,T} \tag{11.112}$$

and hence the loan rate applied by transaction banks is given as

$$1 + r_L = \theta \frac{1 + r_D}{\pi} + (1 - \theta) \frac{V}{L}. \tag{11.113}$$

Bank profits from providing such a loan are then obtained as

$$\hat{\Pi}_B^T = (1 - \theta)(\pi V - (1 + r_D) L). \tag{11.114}$$

The banks' profits from given a fraction ρ of relationship loans and $1 - \rho$ transaction loans is then given, after subtracting the sunk costs C of investing into a fraction ρ of relationship banking, by

$$\Pi_B^i = \rho \left(\hat{\Pi}_B^{i,R} - C \right) + (1 - \rho) \hat{\Pi}_B^T. \tag{11.115}$$

Thus the optimal investment into relationship banking is given from the first order condition

$$\frac{\partial \Pi_B^i}{\partial C} = \frac{\partial \rho}{\partial C} \left(\hat{\hat{\Pi}}_B^i - \hat{\Pi}_B \right) - \frac{\partial \rho}{\partial C} C - \rho = 0. \tag{11.116}$$

From inserting equations (11.107) and (11.114) we know that $\hat{\hat{\Pi}}_B^i - \hat{\Pi}_B = (1 - \theta)(\pi_i - \pi) V$. Hence differentiating the expression in expression (11.116) for θ we get

$$\frac{\partial^2 \rho}{\partial C^2} \frac{\partial C}{\partial \theta} ((1 - \theta)(\pi_i - \pi) V - C) - \frac{\partial \rho}{\partial C} (\pi_i - \pi) V - \frac{\partial \rho}{\partial C} \frac{\partial C}{\partial \theta} = 0, \tag{11.117}$$

which solves for

$$\frac{\partial C}{\partial \theta} = \frac{\frac{\partial \rho}{\partial C} (\pi_i - \pi) V}{\frac{\partial^2 \rho}{\partial C^2} ((1 - \theta)(\pi_i - \pi) V - C) - \frac{\partial \rho}{\partial C}} < 0, \tag{11.118}$$

where the last inequality arises from our assumption that the marginal cost of providing relationship banking are decreasing as the investment C increases and we only consider companies that have high success rates such that $\pi_i = \pi_H > \pi$. This can be justified through the assumption that companies with low success rates are not provided with loans if we assume that $\Pi_B^{L,R} < 0$. and hence $\pi_L V < (1 + r_D) L$. We thus find that an increase in competition between banks, θ, decreases the investment into relationship banking, C, and thus decreases the fraction of loans that are offered on the basis of relationship banking.

Hence we observe that competitive forces eroding bank profitability make relationship banking less important. The reason for this result is that increased competition allows banks to make less profits, leaving less resources available to cover the costs of relationship banking, which is therefore reduced in scope. In markets that are particularly competitive, relationship banking will be less important than in markets that are overall less competitive.

Reading Yafeh & Yosha (2001)

Résumé

The effect of competition on relationship banking is twofold. On the one hand, competition between banks increases the importance of relationship banking; banks will lose profits, but their informational advantage will allow them to retain profits arising from this source, giving them an advantage over transaction banks that makes relationship banking an ever more important source of bank profits. As competition increases, the informational advantage required to recover the costs of relationship banking can reduce as other sources of profits to cover these costs are diminished and banks rely on their informational advantage ever more. On the other hand, the reduced profits with increasing competition makes it more difficult for banks to recover the additional costs that are associated with relationship banking. Therefore, banks will reduce these costs, which will in turn reduce their capacity for relationship banking.

The strength of each of these two factors will determine the overall effect of competition on the prevalence of relationship banking. In markets where adverse selection between relationship banks and other banks is high, we can expect the high profits banks obtain from their informational advantage to dominate the effect of overall reduced profits, thus making relationship banking more important. This would be particularly the case if competition is already high and generally profits are low. In markets with low adverse selection costs and low degrees of competition between banks, increasing this competition might reduce the importance of relationship banking.

Conclusions

Relationship banking allows a bank to accumulate information over a longer period of time through repeated interactions with a company. Through these interactions, such as ongoing monitoring during the lifetime of a loan, but also the assessment of a company for many loans over time, banks are able to gain much better information about a company than would be possible at the time of a loan application alone. It is not only a problem of the time involved in making such detailed assessments, but also the costs involved. If the costs have to recovered from a single loan, such extensive information collection might prove to be too costly and the requisite loan rate might make the bank less attractive than competitors who gather less information.

If information can, however, be re-used in future lending, then these costs can be spread over multiple loans, making the bank more competitive.

Having gained an informational advantage, banks are able to exploit their improved position relative to competitors by offering higher loan rates than would be necessary given the risks they have assessed, but which other banks cannot compete with due to their inferior information. This can lead to a situation in which loan rates in relationship banking are higher than they would be if banks were competitive. With companies facing even higher loan rates at other banks, or facing additional costs when switching banks, their bank can make excess profits from the relationship. However, banks will compete to gain this relationship and would entice companies to engage with them. This can lead to banks offering very favourable introductory loan conditions, even not covering the cost of their initial loan, but financed through the excess profits they can make later once the relationship has been established. Hence we would see loan rates increased after the initial introductory phase. Such increases in loan rates can be prevented by entering long-term contracts that will prevent banks from exploiting their informational advantage once they have gained the information, but will come at the cost of higher initial loan rates, thus it will represent only a shift of costs from future loan rates to the initial loan rate.

Banks having superior information of a company can have two effects on the ability of companies access loans. If the information held is favourable, then the company will find it easier to access loans and this ability to finance investments more easily, will compensate them for any additional costs arising from the information costs banks face. But on the other hand, with negative information obtained by the bank, a company might find it more difficult to obtain a loan from their own bank. They might have to seek loans from other banks, who hold less information on them, but will charge them a higher loan rate than their own bank would. In these situations, it may well be that companies are securing loans that are inherently too risky and should not be granted a loan, causing banks larger losses in the future.

Increasing competition between banks has eroded their profits margins and this makes information about companies ever more important as this allows them to offer loan rates that are competitive, but at the same time profitable for the bank. Having an informational advantage over competitors should allow the bank to offer loan rates that are more profitable than loan rates by competitors. Hence having relationships with companies will allow banks to be profitable despite facing fierce competition; it is this informational advantage that competition cannot eliminate easily. Thus with increasing competition, the importance of relationship banking should be increasing. On the other hand, however, the lower profits margins of banks may not allow them so easily to recover the additional costs they have in relationship banking, transaction banks might offer loan rates that are below the costs of relationship banks, despite their informational advantage. Hence competition might actually hinder the importance of relationship banking. Both effects will be present and it will depend on the costs and informational advantages relationship banks can generate, which effect will dominate.

Looking for a way to overcome the effect of relationship banking by allowing companies to share information a bank holds about them with other banks, results in only low-risk companies agreeing to sharing their information. Thus only companies who would find it relatively easy to obtain a loan from another bank would agree to such a measure, limiting the positive impact open banking could have on the effect of relationship banking.

Chapter 12

Open banking

If a company has established a relationship with a bank, this bank might be able to assess the risks of the company more precisely than another bank that does not benefit from such a relationship. This puts a bank without a prior relationship at a disadvantage over the bank having established such a relationship. Even if other banks have higher abilities to assess the risks of companies, the lack of information might hold back the precision of their assessment; however, if they had access to the same amount of information, they could provide a more accurate assessment.

Companies can decide to share the information their current bank holds with any other bank in what is often referred to as open banking. In open banking, all banks are having access to the same information and thus the benefits of relationship banking are eroded significantly, being limited mainly to information that is not recorded formally. Companies can decide whether to share information their bank holds with other banks if they deem this to be profitable, but are generally not required to do so.

Let us assume that companies use loans L to make an investment that yields a return of R if successful and provides no return if unsuccessful; the probability of success is π. Banks now obtain information about the ability of the company being successful and hence repaying their loan. We denote the ability to repay the loan as the high state H and the company not able to repay its loan is referred to as the low state L. The information banks receive, is correct with probability $\nu_j > \frac{1}{2}$, such that the probability of a signal, given the true state is given by

$$\text{Prob}\left(s_j = H|H\right) = \text{Prob}\left(s_j = L|L\right) = \nu_j. \tag{12.1}$$

There are two banks competing to provide a loan, bank 1 is the current bank of the company while bank 2 has no previous relationship with the company. Thus, the signal bank 2 receives will be less precise than that of bank 1, such that $\nu_1 > \nu_2$. Banks can use Bayesian learning to obtain the probability of the loan being repaid, depending on signals H and L, respectively. We easily obtain that the probability of a state, given the signal received, is

$$\pi_j^H = \text{Prob}\left(H|s_j = H\right) = \frac{\pi\nu_j}{\pi\nu_j + (1-\pi)\left(1-\nu_j\right)}, \tag{12.2}$$
$$1-\pi_j^L = \text{Prob}\left(L|s_j = L\right) = \frac{(1-\pi)\,\nu_j}{(1-\pi)\,\nu_j + \pi\left(1-\nu_j\right)},$$

where π_j^H and π_j^L can be interpreted as the probability of success, given that signals H and L, respectively, have been observed. Using our assumption that $\nu_1 > \nu_2$, we can easily derive that $\pi_1^H > \pi_2^H$, $\pi_1^L > \pi_2^L$, and $\pi_j^H > \pi_j^L$ if $\pi > \frac{1}{2}$.

We now make the assumption that banks do not provide a loan to the company if they receive the low signal L. Thus we assume that $\pi_j^L(1+R) - (1+r_D) < 0$. The highest possible loan rate is the return R of the company in case the investment is successful and even charging this loan rate, will not cover the funding costs in form of the deposit rate r_D. We here assume in addition that bank finances the loan entirely through deposits.

We can now analyse the resulting profits to companies first without and then with open banking. By comparing these results, we can then determine under which conditions open banking is preferred by companies.

Without open banking Having two banks obtaining information, we see that both banks obtain a high signal H if the actual outcome is that the company repays its loan, π, and both banks obtain the correct information, ν_j, or alternatively, the loan is not repaid, $1-\pi$, but both banks obtain the wrong information, $1-\nu_j$. Similarly, we obtain the probabilities for bank 1 obtaining a high signal and bank 2 a low signal, bank 2 obtaining a low signal while bank 2 obtains a high signal, and both banks obtain a low signal. We can determine these probabilities easily as

$$\begin{aligned} p_{HH} &= \pi\nu_1\nu_2 + (1-\pi)(1-\nu_1)(1-\nu_2), \\ p_{HL} &= \pi\nu_1(1-\nu_2) + (1-\pi)(1-\nu_1)\nu_2, \\ p_{LH} &= \pi(1-\nu_1)\nu_2 + (1-\pi)\nu_1(1-\nu_2), \\ p_{LL} &= \pi(1-\nu_1)(1-\nu_2) + (1-\pi)\nu_1\nu_2. \end{aligned} \tag{12.3}$$

If we assume that $\pi > \frac{1}{2}$, is is straightforward to see that $p_{HL} > p_{LH}$.

Given the signals received by the two banks, we can now assess the likelihood of the company repaying their loan. Using Bayesian learning and an assumption that the signals the banks receive are independent such that $\text{Prob}(s_1 = H, s_2 = H|H) = \text{Prob}(s_1 = H|H)\,\text{Prob}(s_2 = H|H) = \nu_1\nu_2$, we obtain the success probabilities if both signals are high, only the first bank obtains a high signal, only the second bank obtains a positive signal, and both banks obtain a low signal, as

$$
\begin{aligned}
\pi_{HH} &= \frac{\nu_1 \nu_2 \pi}{p_{HH}}, \\
\pi_{HL} &= \frac{\nu_1 (1 - \nu_2) \pi}{p_{HL}}, \\
\pi_{LH} &= \frac{(1 - \nu_1) \nu_2 \pi}{p_{LH}}, \\
\pi_{LL} &= \frac{(1 - \nu_1)(1 - \nu_2) \pi}{p_{LL}}.
\end{aligned}
\tag{12.4}
$$

Let us now assume that the two banks charge the same loan rate r_L and if both offer a loan, they provide this loan with equal probability. With our assumption that banks do not provide a loan if receiving the low signal L, they will provide a loan only if either both banks obtain a high signal H and they share the provision of the loan, or only they obtain the high signal and therefore grant the loan to the company without competition. Taking into account these two scenarios, the profits of the first bank are given by

$$
\begin{aligned}
\Pi_B^1 &= \frac{1}{2} p_{HH} \left(\pi_{HH} (1 + r_L) L - (1 + r_D) L \right) \\
&\quad + p_{HL} \left(\pi_{HL} (1 + r_L) L - (1 + r_D) L \right) \\
&= \pi \nu_1 \left(1 - \frac{1}{2} \nu_2 \right) (r_L - r_D) L \\
&\quad - \frac{1}{2} (1 - \pi)(1 - \nu_1)(1 + \nu_2)(1 + r_D) L,
\end{aligned}
\tag{12.5}
$$

where the second equality arises when inserting from equations (12.3) and (12.4). Similarly we obtain for the second bank that

$$
\begin{aligned}
\Pi_B^2 &= \frac{1}{2} p_{HH} \left(\pi_{HH} (1 + r_L) L - (1 + r_D) L \right) \\
&\quad + p_{LH} \left(\pi_{LH} (1 + r_L) L - (1 + r_D) L \right) \\
&= \pi \nu_2 \left(1 - \frac{1}{2} \nu_1 \right) (r_L - r_D) L \\
&\quad - \frac{1}{2} (1 - \pi)(1 - \nu_2)(1 + \nu_1)(1 + r_D) L.
\end{aligned}
\tag{12.6}
$$

Using our assumption that $\nu_1 > \nu_2$, we can show that $\nu_1 \left(1 - \frac{1}{2}\nu_2\right) > \nu_2 \left(1 - \frac{1}{2}\nu_1\right)$ and hence the first term is larger for bank 1. We also find that $(1 - \nu_1)(1 + \nu_2) < (1 - \nu_2)(1 + \nu_1)$ making the second term smaller for bank 1. We thus see that for banks charging the same loan rate, bank 1 makes higher profits than bank 2, $\Pi_B^1 > \Pi_B^2$.

If the two banks are setting the same loan rate, a bank could undercut this loan rate and by offering the lower loan rate provide the loan for sure, increasing their profits. This competition will continue until bank 2 cannot undercut the loan rate as $\Pi_B^2 = 0$ even if they obtain the full loan in case both signal are high. Thus bank 1 could charge this loan rate, less a marginal amount, to provide all loans. If bank 2 were setting the lowest loan rate it would provide the loan regardless of the other bank's signal and this bank's profits would be given by

$$\begin{aligned}\Pi_B^2 &= p_{HH}\left(\pi_{HH}\left(1+r_L\right)L-\left(1+r_D\right)L\right) \\ &\quad +p_{LH}\left(\pi_{LH}\left(1+r_L\right)L-\left(1+r_D\right)L\right) \\ &= \pi\nu_2\left(1+r_L\right)L \\ &\quad -\left(\pi\nu_2+\left(1-\pi\right)\left(1-\nu_2\right)\right)\left(1+r_D\right)L.\end{aligned} \tag{12.7}$$

Setting $\Pi_B^2 = 0$ due to its competition with bank 1, gives us the loan rate as

$$1+r_L = \frac{\pi\nu_2+\left(1-\pi\right)\left(1-\nu_2\right)}{\pi\nu_2}\left(1+r_D\right). \tag{12.8}$$

This loan rate will be charged by bank 1 as it seeks to maximize its profits, but bank 2 is not able to undercut it as it would make a loss, where bank 1 would still remain profitable.

Companies obtain a loan if either bank receives a high signal H and pays the loan rate as determined by equation (12.8), either to bank 1 if they obtain a high signal or to bank 2 if bank 1 receives a low signal. Thus company profits are given by

$$\begin{aligned}\Pi_C &= \pi\left(p_{HH}+p_{HL}+p_{LH}\right)\left(\left(1+R\right)L-\left(1+r_L\right)L\right) \\ &= \pi\left(\pi\left(\nu_1+\nu_2-\nu_1\nu_2\right)+\left(1-\pi\right)\left(1-\nu_1\nu_2\right)\right) \\ &\quad \times\left(\left(1+R\right)-\frac{\pi\nu_2\left(1-\pi\right)\left(1-\nu_2\right)}{\pi\nu_2}\left(1+r_D\right)\right)L,\end{aligned} \tag{12.9}$$

where we used equations (12.4) and (12.8) to obtain the second equality.

We can easily show that companies will prefer banks to obtain more precise signals as $\frac{\partial\Pi_C}{\partial\nu_1} > 0$ and $\frac{\partial\Pi_C}{\partial\nu_2} > 0$. Bank 2 having more precise information allow it to compete better with bank 1 as their informational disadvantage is reduced, which lowers the loan rate and benefits the company. More precise information will also increase the probability of the company obtaining a loan as long as $\pi > \nu_j$. If the success rate is sufficiently high, banks are more likely to obtain a high signal if their information precision increases as false low signals are reduced more than false high signals; this is the result of high success rates being unlikely to be overestimated.

Thus, it is beneficial to companies if banks have better information. We can now compare these profits of companies in the absence of open banking to their profits if they agree to open banking and share information with the other bank.

With open banking If the company shares information with the other bank through open banking, we assume that this bank has a higher ability to assess the company, implying that the precision of its information is higher. Thus we assume that in this case the precision of the information increases from ν_2 to $\hat{\nu}_2$ with $\hat{\nu}_2 > \nu_1 > \nu_2$. The consideration of banks are unchanged from the above case without open banking, it is now only that the roles of the banks are reversed. Bank 1 is now the bank with the lower precision of information while bank 2 has the higher precision of information with a value of $\hat{\nu}_2$.

Replacing the variables accordingly, we easily obtain the loan rate in analogy to equation (12.8) as

$$1 + \hat{r}_L = \frac{\pi \nu_1 + (1-\pi)(1-\nu_1)}{\pi \nu_1}(1 + r_D) < 1 + r_L. \tag{12.10}$$

The more precise the information of bank 1, compared to bank 2 without open banking, allows the loan rate to be lower as the uncertainty of the bank is reduced; this lower loan rate needs to be balanced against the ability to obtain a loan, namely $p_{HH} + p_{HL} + p_{LH}$, and as argued above, provided the success rate π is sufficiently high, this probability increases in the precision of information, ν_j. However, if introducing open banking, the information precision changes such that the bank which had previously a lower information precision, has now a higher information precision. This is equivalent to changing the information precision of both banks simultaneously, having in general an ambiguous overall effect on the profits of the company.

These company profits are in analogy to equation (12.9) given by

$$\begin{aligned} \hat{\Pi}_C &= \pi (p_{HH} + p_{HL} + p_{LH})((1+R)L - (1+\hat{r}_L)L) \\ &= \pi (\pi (\hat{\nu}_2 + \nu_1 - \nu_1 \hat{\nu}_2) + (1-\pi)(1-\nu_1\hat{\nu}_2)) \\ &\quad \times \left((1+R) - \frac{\pi \nu_1 (1-\pi)(1-\nu_1)}{\pi \nu_1}(1+r_D)\right) L. \end{aligned} \tag{12.11}$$

Companies will seek to share information through open banking if it is profitable to do so, thus provided $\hat{\Pi}_C > \Pi_C$. Inserting from equations (12.9) and (12.11), we obtain that this condition is fulfilled if

$$\hat{\nu}_2 \geq \hat{\nu}_2^* = \frac{\nu_1 (\pi - \nu_2) + (1-\pi)(1-\nu_2)\frac{R-\hat{r}_L}{R-r_L} - (1 - \pi(1-\nu_1))}{\pi - \nu_1}. \tag{12.12}$$

We thus see that the improvement in the precision of the signal for bank 2 must be sufficiently precise such that open banking is beneficial to companies. An analytical analysis of the expression is difficult to conduct; what we find, however, is that open banking is mostly beneficial to companies if the differences in the precision of the initial information, ν_1 and ν_2, is not too large. In this case the two banks are nearly identical and they will be highly competitive, reducing the loan rate to a level close to perfect competition and banks extract little profits from the company. As the differences in the information precision between banks increases, competition reduces and banks can increase their profits at the expense of the company, whose profits will fall. The introduction of open banking increases the ability of bank 2 to identify companies that will not be able to repay their loan; this will reduce the chances of such companies to obtain a loan and the possibility, even if small, of making profits.

With the assessment of bank 1 unchanged, this leads to two effects, the reduction in their chance of obtaining a loan on the one hand, and a lower loan rate on the other hand. If the difference in the information precision between banks is large, the likelihood of bank 2 obtaining a different signal to bank 1 is high, and hence the chances of the company to obtain a loan is high. With open banking, bank 1 becomes the bank providing the fallback loan if the better informed bank 2 received a low signal; their more precise information, compared to the original bank 2 without open banking, makes it less likely that it will receive a high signal, reducing the likelihood of obtaining a loan. On the other hand, the higher precision of bank 2 with open banking reduces wrong low signals, making it more likely the company can obtain a loan. However, the more different banks are without open banking, the higher the information precision needs to be with open banking to compensate for this effect. As we naturally require that $\hat{\nu}_2 \leq 1$, for larger differences between banks, no feasible solution can be found and open banking is not beneficial to companies.

We also observe that companies with higher success rates, π, are more likely to benefit from open banking; the difference in information precision between the initial banks can be larger the higher the success rate of banks is. This is because the higher success rate makes incorrect assessments of companies less likely, reducing the adverse effect of having as a fallback a bank with higher information precision, but nevertheless benefitting from the reduced loan rate. Similarly, a higher the return on the investment of the company, R, or a lower the deposit rate and hence loan rate, imply that the signal precisions of the initial banks can differ more and still allow companies to benefit from open banking. In this case, companies are obtaining large profits if they are able to secure a loan and this impact will dominate, especially the ability to obtain a loan if the investment is going to succeed, which using open banking results more frequently in high signals and hence a loan being granted, while the loss of a loan that is unlikely to succeed is of less importance given its low rate of success with the high precision of information banks have.

It is thus that low-risk and highly profitable companies are most likely to benefit from open banking. Such companies will find open banking more profitable for a wider range of differences in the signal precision between banks and a lower improvement of information precision through open banking is required for any given differences. Combining these two aspects implies that the conditions for open banking being profitable to these companies is increased.

Companies with different risks We have established that companies with low risks are more likely to benefit from open banking than high-risk companies. Thus far, we implicitly have assumed that only one type of company exists in the market and banks know the probability of success, π. Let us now assume that there are two companies with different success rates, π_H and π_L, with $\pi_H > \pi_L$. While companies know their own type, banks are not aware of these, they only know that a fraction p has the high success rate π_H. Thus the expected success rate from the bank's perspective is $\pi = p\pi_H + (1 - p)\,\pi_L$. As before, banks obtain information on the success of companies and Bayesian learning allows them to update their beliefs in the same way as in equation (12.2); if they do not know the type of company, they will use π as the success rate to be updated and if they knew the success rate π_i of companies, they will use this success rate in their decision-making.

Banks might distinguish companies through their choice of choosing open banking or not, thus the choice of open banking may provide information about their type. Of course, if both companies make the same choices, the banks cannot distinguish between them and they will have to set the loan rate and grant loans based on the inferred average success rate π. However, if companies make different choices, banks can distinguish their types and use the relevant success rates if π_i for their assessment. Let us now denote by Π_C^p ($\hat{\Pi}_C^p$) the profits of the company if the bank cannot distinguish their types and the company chooses open banking (does not choose open banking). Similarly, we denote the company profits of those identified as having the high success rate by Π_C^H ($\hat{\Pi}_C^H$) and those with the low success rate by Π_C^L ($\hat{\Pi}_C^L$). In each case, we replace the generic π in equations (12.9) and (12.11) with π as defined above, π_H, or π_L, as appropriate. It is only the first π that is changed to π_H or π_L, depending on the type of company, as this signifies the success rate of the company as known by the company itself, who knows its type.

		Company H	
		Open banking	No open banking
Company L	Open banking	$\hat{\Pi}_C^p, \hat{\Pi}_C^p$	$\Pi_C^H, \hat{\Pi}_C^L$
	No open banking	$\hat{\Pi}_C^H, \Pi_C^L$	Π_C^p, Π_C^p

Fig. 12.1: Strategic interactions in choosing open banking

Figure 12.1 shows the resulting strategic interaction between the two company types. If both choose open banking or both choose no open banking, they cannot be distinguished by banks, while in those cases they make different decisions, banks can infer their type. An analytical analysis of this strategic game is again very difficult to conduct, however, we observe only two possible equilibria in this game with reasonable parameter constellations. If the differences between companies are not too substantial, thus π_H is not too much higher than π_L, then only the low-risk company, π_H, will choose open banking, while the high-risk, π_L, will not engage in open banking. If the differences between companies become more substantial, then other possible equilibria emerge, namely that only one of the companies opts for open banking, but not both banks; this could be either the low-risk company or the high-risk company.

If the differences between the success rates of companies are large, the high-risk company might want to make the same choice as the low-risk company and thus benefit from a significantly reduced loan rate, which outweighs the reduction in the chance of obtaining a loan. This, however, reduces the benefits of low-risk companies as they cannot be distinguished by banks, leading them to choose to not participate in open banking in order to distinguish themselves. By not participating in open banking, they forego the benefits of sharing information, but being distinguished from high-risk companies provides overall a higher benefit. This leads to an equilibrium in which only one of the banks participates in open banking, but it could be either companies and hence banks would not be able to determine which company it is, thus they cannot distinguish the types of companies from their choices and this equilibrium is not sustainable.

The more interesting equilibrium is the one in which only the low-risk company seeks open banking; we could expect that large differences in the risks between companies can be identified by banks. In addition, if either one of the companies would choose open banking, we can reasonably suggest that banks will not be able to infer their type as they do not know which of the two possibilities the companies have chosen, the low-risk company choosing open banking and the high-risk company not choosing open banking or vice versa. This would lead us back to a situation where companies cannot be distinguished and the choice of open banking will be solely based on their own profits, giving the same results as analysed above, namely that low-risk companies choose open banking, while high-risk companies will not choose open banking. The same result as with the strategic interaction obtained here.

For the strategic equilibrium, the differences between companies can increase the more polarised the distribution of companies becomes, thus the more one type dominates over the other. Generally, the low-risk company seeks to distinguish itself from the high-risk company as that would allow it to obtain a lower loan rate and a wrong assessment of its ability to repay the loan is reduced; the high-risk company would seek to copy the behaviour of the low-risk company to avoid being identified as being high-risk, however, a high-risk company would not want to opt for open banking, too. While their loan rate would decrease, given that banks cannot distinguish between banks of different risks if both choose open banking, the increased precision of information would make it less likely they obtain a loan, making this choice unprofitable as we have discussed above.

We also find that if the difference between the information precisions increases, open banking becomes more attractive even if companies are more different in their risks. With companies being more different, the benefits of low-risk companies being identified as such are larger. In line with our analysis above, the more different the information precision between banks is, the more difficult it is to obtain benefits from open banking. However, given the benefits for larger differences are increasing, it will be easier for open banking to generate these benefits, even if the information precision with open banking does not increase that much. It is the low-risk company being identified as such, that gives them substantial benefits and induces them to choose open banking.

It is thus that even if the choice of open banking provides information to the bank about the type of company making this decision, only low-risk companies are opting for open banking. The benefits they obtain from better information through information sharing and being identified as being low-risk outweigh the possibility of not obtaining a loan if banks obtain low signals. It is not beneficial for high-risk companies to copy this choice as their chance of not obtaining a loan due to the improved information suggesting they are not able to repay a loan outweighs the benefits of a lower loan rate.

Summary Open banking is attractive to companies that are low-risk and highly profitable, while companies that are more risky and perform less well, will find open banking not attractive. It is low-risk companies in particular that benefit from the exchange of information between banks as this allows other banks to assess their risks more accurately. Given that with low risks information will more commonly show that they are able to repay their loans, the overall effect of providing such information is positive for these companies. On the other hand, more risky companies are not able to easily benefit from open banking as their risk-assessment might become less favourable than without the additional information. Of course, opting for open banking can send a signal about the risk a company is exposed to and thus will provide information to banks, re-enforcing this mechanism.

The benefits of open banking accrue mainly to those companies that will already find it easy to obtain loans, those that are low-risk and highly profitable; the introduction of open banking might therefore make it more difficult and more expensive for more risky companies to obtain loans. It is therefore that innovative companies, which are usually associated with higher risks, are disadvantaged as their refusal to choose open banking can identify them even more clearly as being high-risk. With the reluctance of high-risk companies to choose open banking due it being not profitable, the fact that open banking has not been embraced by companies cannot be fully attributed to concerns about data security.

Reading He, Huang, & Zhou (2023)

Chapter 13

Securitization

Traditionally, banks hold loans until maturity and at this time obtains the repayments of the company, which then allows them to make payments to their depositors, who financed the loan. It has become standard practice, however, for banks to not hold their loans until maturity but sell them other investors, often hedge funds or pension funds. To achieve this sale of loans, they are transferred to a Special Purpose Vehicle, a company set up for the specific task of selling these loans. Having received these loans, the Special Purpose Vehicle then issues a bond that is sold to investors and the proceeds of this bond sale handed to the bank as payment for the sold loans. The loans, now held by the Special Purpose Vehicle, act as collateral for the repayment of the loan; given that the Special Purpose Vehicle has no other assets, the repayment of these loans will fully determine the repayment of the bond. As bonds are securities, this process is called securitization.

As the repayment of the bond fully depends on the repayment the Special Purpose Vehicle obtains from the loans, these bonds can be risky to investors. To make them more attractive, banks often apply a credit enhancement in the form of a guarantee by the bank. This guarantee consists of a promise that the bank will ensure that at least a certain fraction of the loans are repaid. Should less loans be repaid, the bank will provide payment to the Special Purpose Vehicle making up the difference.

Normally the company of a loan that is being securitized continues to make payment to the bank and the bank then transfers these payments to the Special Purpose Vehicle. However, the payments from the company will not necessarily coincide with the payments the Special Purpose Vehicle receives; the difference is often referred to as a ”service charge” and covers the cost of administration and any credit enhancements the bank provides. Securitisation thus doe snot only allow banks to sell some of their loans, but will also create a steady source of income from such a service charge.

Let us assume that companies make investments that allow them to repay their loan L, including interest r_L, if successful, which happens with probability π. With probability $1 - \pi$ the investment will not be successful and company will not be able to repay any amount of the loan. Banks finance the entire loan through deposits, on which interest r_D is payable, and in addition hold equity E that allows them to repay their depositors partly if the loan is not repaid. We finally assume that banks know the probability of the company's investment being successful, and hence the loan being repaid, but the depositor only knows this information after incurring costs C.

We will first consider the case of a loan that cannot be securitized before then considering the securitization and whether it is desirable to do so. Having deposited the amount of L with the bank, the depositor is fully repaid, including interest, if the loan the bank has granted is repaid. In the case the loan is not repaid, the depositor receives the equity the bank holds. Including the costs C of learning the probability of success of the loan, the depositor makes profits of

$$\Pi_D = \pi (1 + r_D) L + (1 - \pi) E - C - L. \tag{13.1}$$

Similar the bank, having invested its equity E, receives payment of the loan from the company if their investment was successful, and from this repays their deposits as well as retains their equity. Hence the bank profits are given by

$$\Pi_B = \pi ((1 + r_L) L - (1 + r_D) L + E) - E. \tag{13.2}$$

With banks being competitive, their profits will vanish and $\Pi_B = 0$, such that

$$1 + r_L = (1 + r)D) + (1 - \pi) \frac{E}{L}. \tag{13.3}$$

We can now compare these profits of depositors in the absence of securitization and the loan rate with those that emerge if the loan can be sold.

We assume that the bank provides a credit enhancement to the loan in the form of a guarantee for a fraction θ of the amount that is to be repaid to the buyers of this loan. The interest accruing to the buyers of the loan is r_S, and any difference to the loan rate applied in these circumstances, $\hat{r}L$, make up the service charge of the bank. The amount the bank thus guarantees, $\theta (1 + r_S) L$, cannot exceed its equity, E, hence we require that $E \geq \theta (1 + r_S) L$.

Banks now do not require depositors anymore, but buyers of the loan. Such buyers obtain their full repayment if the investment of the company is successful and it can repay its loan, and if this is not possible, the buyer will obtain the guarantee by the bank. With an initial investment of L, their profits become

$$\begin{aligned} \Pi_S &= \pi (1 + r_S) L + (1 - \pi) \theta (1 + r_S) L - L \\ &= (\pi + (1 - \pi) \theta) (1 + r_S) L - L. \end{aligned} \tag{13.4}$$

The purchasers of the loan, like depositors do not know the probability of the company investment being successful, but as we will see below are able to make such inferences form the interest r_S they obtain. Thus they do not face costs of obtaining this information.

Banks, as before, having invested their equity E, receive payment of the loan from the company if their investment was successful, and from hand on the payments to the buyers of the loan as well as retains their equity. If the loan is not repaid, the bank has to pay out its guarantee, but can retain a part of their equity. Hence, bank profits are given by

$$\begin{aligned}\hat{\Pi}_B &= \pi\left(\left(1+\hat{r}_L\right)L-\left(1+r_S\right)L+E\right) \\ &\quad+\left(1-\pi\right)\left(E-\theta\left(1+r_S\right)L\right)-E \\ &= \pi\left(1+\hat{r}_L\right)L-\left(\pi+\left(1-\pi\right)\theta\right)\left(1+r_S\right)L.\end{aligned} \tag{13.5}$$

If banks are competitive again such that $\hat{\Pi}_B = 0$, we have

$$1+\hat{r}_L = \frac{\pi+\left(1-\pi\right)\theta}{\pi}\left(1+r_S\right). \tag{13.6}$$

We see that the loan rate charged to companies exceeds the interest on the securitized loan and the bank earns a service charge for this loan.

The company is unaffected by the sale of the loan, but will be prefer the loan being sold, that is securitized, if the loan rate they obtain is smaller, this if $\hat{r_L} \geq r_L$. Using equations (13.3) and (13.6), this easily solves for

$$1+r_S \leq \frac{\pi}{\pi+\left(1-\pi\right)\theta}\left(1+r_D\right)+\frac{1-\pi}{\pi+\left(1-\pi\right)\theta}\frac{E}{L}. \tag{13.7}$$

In order to the purchase of the loan to be attractive, it has to generate at least as much profits as the providing deposits to a bank, thus we require that $\Pi_S \geq \Pi_D$, which using equations (13.1) and (13.4) becomes

$$1+r_S \geq \frac{\pi}{\pi+\left(1-\pi\right)\theta}\left(1+r_D\right)+\frac{1-\pi}{\pi+\left(1-\pi\right)\theta}\frac{E}{L}-\frac{1}{\pi+\left(1-\pi\right)\theta}\frac{C}{L}. \tag{13.8}$$

Combing these last two equations, we get

$$\begin{aligned}&\frac{\pi}{\pi+\left(1-\pi\right)\theta}\left(1+r_D\right)+\frac{1-\pi}{\pi+\left(1-\pi\right)\theta}\frac{E}{L} \geq 1+r_S \\ &\quad\geq \frac{\pi}{\pi+\left(1-\pi\right)\theta}\left(1+r_D\right)+\frac{1-\pi}{\pi+\left(1-\pi\right)\theta}\frac{E}{L}-\frac{1}{\pi+\left(1-\pi\right)\theta}\frac{C}{L}.\end{aligned} \tag{13.9}$$

If the costs of depositors becoming informed of the success rate π, C, become small the interest rate on the securitized loan would thus be determined. In general, however, with positive costs, banks will make the highest profits if the rate offered to the buyers of the loan are as low as possible, hence we set the interest rate at the lower bound such that

$$1+r_S = \frac{\pi}{\pi+\left(1-\pi\right)\theta}\left(1+r_D\right)+\frac{1-\pi}{\pi+\left(1-\pi\right)\theta}\frac{E}{L}. \tag{13.10}$$

Apart from the equity ratio $\frac{E}{L}$ and the deposit rate r_D, which are both observable, the interest offered to purchasers of the loan will depend on the level of the guarantee,

θ, and the probability of success of the investment, π, which is unknown to the purchaser. We can now easily obtain that

$$\frac{\partial\left(1+r_S\right)}{\partial\theta} = -\frac{1-\pi}{\left(\pi+(1-\pi)\,\theta\right)^2}\left(\pi\left(1+r_D\right)+(1-\pi)\,\frac{E}{L}\right) < 0, \quad (13.11)$$
$$\frac{\partial\left(1+r_S\right)}{\partial\pi} = \frac{\theta\left(1+r_D\right)-\frac{E}{L}}{\left(\pi+(1-\pi)\,\theta\right)^2} < 0.$$

We can now see when using equation (13.10) that $r_S \geq r_D$ if $E \geq \theta\left(1+r_D\right)L$ and the final inequality in equation (13.11) has the same requirement. As we had assumed that for the guarantee of the bank to be credible we require that $E \geq \theta\left(1+r_S\right)L$ and $r_S \geq r_D$, this condition is fulfilled.

Using the implicit function theorem we easily obtain that

$$\frac{\partial\theta}{\partial\pi} = -\frac{\frac{\partial(1+r_S)}{\partial\pi}}{\frac{\partial(1+r_S)}{\partial\theta}} < 0 \quad (13.12)$$

and hence when observing the credit enhancement θ, the purchaser of the loan can infer the probability of success of the investment.

It is thus that securitization is desirable for depositors and companies if both conditions in equation (13.9) are fulfilled. With company profits given by

$$\Pi_C = \pi\left(\left(1+R\right)L-\left(1+\hat{r}_L\right)L\right), \quad (13.13)$$

where R denotes the return on a successful investment, it is clear that as long as $R > \hat{r}_L$, companies would seek loans. Using equations (13.6), (13.6) and (13.10), we easily get that $1+r_L = 1+\hat{r}_L = \left(1+r_D\right)+\frac{1-\pi}{\pi}\frac{E}{L}$. As the loan rates with and without securitization are identical, companies are indifferent to securitization, where as for depositors $r_S \geq r_D$ and hence purchasing securitized loans is more attractive than deposits. Similar, if we had chosen an interest for the securitized loan at the upper bound of the constraint in equation (13.9), depositors would be indifferent to securitization, while companies would prefer banks that securitize their loans. Choosing any interest rate strictly within the constraint of equation (13.9) would see both, companies and depositors to prefer securitization. In all cases, we assumed competitive banks, who would thus be indifferent about securitizing their loans or not securitizing them.

The result that securitization is desirable arises from the use of equity by banks. If retaining the loan until maturity, the bank faces the prospect of losing this equity to repay depositors if the loan is not repaid. With securitization this loss is reduced to the guarantee the bank provides. These reduced losses are reflected in either a lower loan rate or a higher interest rate on the securitized loan, making securitization more attractive to companies.

In reality, the requirement for credit enhancement will prevent the bank from being able to sell too many loans as banks typically only hold a small amount of equity when compared to the amount of loans they provide. It will thus not be possible for banks to securitize all their loans, but they will have to retain the majority of their loans to maturity. Limits to securitization are also a possible adverse selection problem in non-competitive markets. If markets are not competitive, the interest rate charged on the securitized loan loses its role as a perfect signal for the probability of success as derived in equation (13.12) and banks could exploit the lack of knowledge by purchasers of loans to sell loans with too-low interest rates, which may lead to the collapse of the market in securitized loans.

Reading Greenbaum & Thakor (1987)

Review

Providing loans is a more complex task than anticipated. A first problem arises from banks having to establish whether a company is genuinely not able to repay a loan; facing costs of verifying the outcome of an investment, we have seen that a standard debt contract where a fixed amount is repaid at maturity is the optimal loan form as this minimizes the costs of banks verifying the outcome. With such a loan contract, banks seek information on the likelihood of the loan actually being repaid and this leads to an arms race in different banks acquiring information in order to gain an informational advantage over competitors, leading to an over-investment into information. The threat of loans not being granted after a company cannot repay a previous loan, can lead to long-term loans that allow companies to conduct multiple successive investments and the failure of one such investment will not prevent future investments from being conducted and those profits being lost. Loans may also be taken out at different seniority levels to take into account competitive advantages of banks, like different funding costs or different abilities to monitor the success of investments. Hence the way loans are constructed is driven by concerns of verifying the outcome of investments by banks, the ability of companies to continue with investments after some failures and the need for information by banks.

Despite the loan contract specified such that it optimally takes into account the need for banks to verify the outcome of investments and their ability to repay loans, such verification will not be perfect. Banks will have limited resources set aside for such monitoring and this will induce some companies to strategically default in the hope that due to limited resources this will not be detected. However, if a company defaults it will be detrimental to their ability to obtain future loans for investments, causing them losses from not being able to generate profits, which will reduce the benefits of strategic default. However, banks would not exclude companies from obtaining any future loans as this also means that companies who have genuinely defaulted on their loans would not get a loan, which means the bank foregoes future profits from companies that are creditworthy; consequently loans will not be granted for some time period and after this has elapsed, they are able to obtain loans again.

Companies are often subject to credit rationing in that they do not obtain a loan of the size they seek. Banks may not be willing to provide loans of a size that are optimal to companies as this would expose them to too high risks, affecting their profits. Banks will limit the size of loans so as to provide incentives to companies to pursue less risky investments due to them having a larger exposure to the risks themselves, but also in order to ensure that the company has the necessary resources to actually repay the loan and make defaults sufficiently unlikely.

A common feature of many loan contracts is the provision of collateral by companies. Having pledged such collateral, the bank obtains some repayment of the loan even if the investment fails, reducing the risks to the bank, who can pass on these benefits to the company through a lower loan rate. But companies providing collateral may also reveal information about the risks they are taking with investments and which the bank is not able to discern. As the collateral will be lost in case the company defaults, companies that take higher risks are less likely to offer collateral than companies that take lower risks and where the loss of the collateral is therefore less likely. This can provide information to the bank as someone offering collateral is more likely to make low-risk investments than some one who does not offer collateral. The possible loss of collateral in case of default will also affect the investment behaviour of companies, reducing the effect moral hazard and asymmetric information. Collateral does not only reduce the risk to banks from lending, but they can also use the collateral that has been pledged to them to obtain loans themselves. This would increase the value of collateral to the bank and companies providing such collateral could benefits from lower loan rates.

Given the importance of information for banks in assessing companies and providing competitive loans while avoiding to make loan offers that are not profitable, they can be expected to guard any informational advantage. However, we see that banks frequently share information about companies through credit reference agencies, eroding this very advantage. The consequence is that banks compete more to provide companies with loans, reducing future profits to banks. However, these lower future profits also means that competition to attract companies in the first place is less intense as there are less profits to be made. These two effects balance each other and banks are largely indifferent about sharing information, while companies may benefit from better loan offers.

A similar effect can be observed in relationship banking. The informational advantage a bank has over its competitors can result in large excess future profits, but then banks will compete to attract such companies, offering attractive loan rates initially, which are then later increased. But banks accumulating information on companies does not only affect loan rates, but also the willingness to provide loans. While companies with positive information might find it easier to obtain future loans, those companies with negative information will often be forced to seek loans at other, less well informed banks, often at much increased loan rates. The importance of relationship banking emerges from the ability to generate profits from their informational advantage. If competition between banks is high, this source of profit cannot be eroded and relationship banking should become more important to generate profits. On the other hand, banks commonly face higher costs in relationship banking and increased competition between banks, especially transaction banks not facing such costs, will make it more difficult to recover any such costs, eroding the position of relationship banking.

Loans may be sold by banks to other investors, allowing the bank to free up capital and provide more loans, but it also allows them to reduce the risk they are exposed to. This allows banks to increase their profits, but at the same time the amount of loans that can be sold will be limited as banks need to retain some of the risks to ensure the adverse selection with investors is not so pronounced that they would not be willing to purchase these loans.

Part III
Deposit and savings accounts

The main source of funding for banks are from deposit and savings accounts. Deposit accounts, typically used to receive payments from and make payments to other accounts, at either the same or different bank and held by different individuals and businesses, typically see a high turnover with significantly varying balances over time. Individuals use such accounts to receive their salary and pension, with additional payments occasionally obtained from other sources, and use these payments to pay household bills, living expenses, leisure activities and similar expenses. The number of transactions will be substantial. Similarly for businesses, they will use their deposit accounts to obtain payments from their customers and pay their suppliers as well as paying salaries of employees.

Savings accounts, on the other hand, have a more stable balance and are primarily opened by individuals. The balance of a savings account might be increased by regular payments or decreased through regular withdrawals, but the number of such transactions are very low, making the balance subject to only few changes. Despite not seeing as many changes to their balance as deposit accounts, the balance in savings accounts can in most cases be withdrawn partially or completely without giving any notice. holders of savings accounts can withdraw their balance to make larger purchases, such as a car or pay for home improvements, or they seek to transfer to another bank and open a savings account there. While some savings accounts require a period of notice, such arrangements affect normally only a small proportion of the overall balance held in savings accounts.

While savings accounts are typically used to maintain funds for a longer period of time than in deposit accounts, in both instances the balance can be withdrawn at any time, but also increased by receiving payments from another account or depositing cash. We therefore do not distinguish between deposit and savings accounts and refer to them jointly as deposits.

Such deposits are seen by most individuals and businesses as a safe way to invest any excess funds. The deposit contract, most notably the deposit rate, should of course nevertheless reflect any risks these deposits are exposed to, while at the same time the deposit contract should ensure that the bank does limit the amount of risk they are exposed to. Chapter 14 will discuss the deposit contract, including the deposit rate, but also the amount of risk banks expose depositors to.

With the ability to withdraw deposits instantly, banks expose themselves to the risk of having to make such repayments without having the liquid assets necessary, given that loans are usually provided for a longer period of time and can therefore not easily be liquidated. We will see in chapter 15 how such bank runs can emerge, either from a change in the expectations of how other depositors will behave or information about the risks of the bank becoming available. With depositors able to withdraw instantly and transfer their deposits to another, bank runs could easily emerge for one bank, while other banks face an influx of deposits. We will discuss in chapter 16 how banks lending to each other can alleviate the shortage of cash reserves by the bank facing the withdrawal of deposits and thus avoid a potential bank run. Interbank lending can also form a source of funding for banks, in addition to deposits. Similarly repurchase agreements, discussed in chapter 17 can serve as a funding for short-term loans the provides. Banks will not have to liquidate any assets to obtain additional cash reserves that then can be lent out, but can instead provide the asset as a collateral for an additional loan by another bank or other, mostly institutional, investors.

Despite reducing risks and the ability to alleviate short-term liquidity shortages through interbank lending, deposits are nevertheless exposed to bank runs, but also losses from banks providing loans that are not repaid. In order to eliminate the risk for depositors, in many countries deposit insurance has been established. With deposit insurance, deposits that cannot be repaid by the bank itself, will be repaid through this deposit insurance scheme, ensuring that no depositors faces any losses. In chapter 18 we discuss the consequences of such deposit insurance on bank and depositor behaviour, along with the optimal level of coverage of deposits. Most deposit insurance schemes do not cover large deposits and not deposits by all types of depositors; we will analyse why such arrangements might be optimal.

While the main focus with respect to deposits is on their role on providing a safe investment, thus addressing predominantly savings accounts, deposit accounts are an important part of the banking business. To this effect, chapter 19 will investigate the services banks provide to deposit accounts. Of particular importance to account holders is the ability make payments and access cash. Payments by individuals are more and more dominate by the use of payment cards instead of cash payments and we will see how providing access such payment forms affects competition between banks and ultimately deposit rates. The payments account holders make will also be reflected in payments that are made between banks to ensure the payment is received correctly. Payments between banks are thus of increasing relevance and as we will discuss can lead to liquidity shortages by banks, exposing banks to additional risks due to the use of transfers between accounts.

Chapter 14

Deposit contracts

A deposit is expected to be repaid by the bank, including any interest, if the depositor demands this. With banks providing loans that may not be repaid, depositors are exposed to the risk of banks not being able to meet their obligation of repaying the deposits they have taken on. This risk for depositors needs to be compensated for as we will see in chapter 14.1 and chapter 14.2 explores whether depositors would like to take on any risks are prefer deposits that are safe. However, gaining interest is not the only motivation to provide banks with deposits. Banks are offering a wide range of account services, most notably the ability to make and receive payments from other account holders at any bank, which depositors would also value. In chapter 14.3 we will see how such benefits can affect the incentives of banks to provide an insurance to depositors that their deposits are being repaid, even if the loans the bank has provided are defaulting.

The competition between banks is not limited to the provision of loans, but they will also compete for deposits. With banks taking different risks when providing loans, depositors will take into account not only the deposit rates banks are offering, but also their risks. In chapter 14.4 we will explore how these different risks banks take affect the competition between them. We include that depositors may have preferences for the account services of a specific bank to enrichen the analysis.

14.1 Deposit rate determination

Deposits can be seen as a form of investing funds and for such funds alternative, risk-less alternatives exist, such as government bonds. Taking into account that banks may fail, depositors will use banks for their investments only if the return they generate will be at least as high as this risk-less alternative. Let us assume that a bank can fail if the loans they have provided are not repaid, which happens with probability π. To become more attractive to depositors, banks may in addition insure their deposits such that if the bank fails, a fraction λ of the deposits are repaid through this insurance. As for any such insurance payout, there will be a delay in payments being made, we may include in this fraction λ an allowance for this delay, for example by discounting all such payments.

For a deposit of size D, which fully finances loans, and a deposit rate of r_D, the profits of the depositor over and above the return it would obtain from investing the amount at the risk-free rate r, are given by

$$\Pi_D = \pi (1 + r_D) D + (1 - \pi) \lambda (1 + r_D) D - (1 + r) D. \tag{14.1}$$

If we assume that depositors are competitive, then $\Pi_D = 0$ and the deposit rate is therefore given by

$$1 + r_D = \frac{1 + r}{\pi (1 - \lambda) + \lambda}. \tag{14.2}$$

Using the approximation that $\ln (1 + x) \approx x$, we can easily transform this expression into the difference between the deposit rate and the risk-free rate and obtain

$$r_D - r \approx (1 - \lambda) (1 - \pi). \tag{14.3}$$

We see immediately that in the case that banks cannot fail, $\pi = 1$, the deposit rate will be identical to the risk-free rate. Similarly, if the deposit insurance covers the deposits fully, $\lambda = 1$, the deposit rate matches the risk-free rate. In either case, the deposits are safe in the sense that they would be repaid to the depositors, including interest, for sure. It is only in the case where either the bank can fail, $\pi < 1$, and the deposit insurance is not complete, $\lambda < 1$, that the depositors face the possibility of losing the deposit. These potential losses are compensated through a higher loan rate.

In reality we often observe that deposit rates are set below comparable risk-free rates. One reason banks may be able to set deposit rates below the risk-free rate is that bank accounts offer a number of additional benefits to depositors, for example the ability to make payments, which are not given by investing into the risk-free asset. Taking into account such benefits, banks might be able to set deposit rates below the level of the risk-free rate.

Reading Cook & Spellman (1994)

14.2 Optimal risk-taking by depositors

It is common to assume that deposits may not be repaid if the loans the bank has provided are not returned. It is, however, unlikely that no repayments of loans are made and depositors will obtain some payments. Banks promising to pay higher deposit rates will be more likely face the prospects of not being able to meet these commitments, thus exposing depositors to risk. Offering a lower deposit rate, which can be paid with greater certainty, might be more attractive to depositors.

Let us assume that banks provide loans L to companies who are able to repay these loans, including interest r_L with probability π_i; loans are long-term in that they are only repaid after multiple time periods. There are two types of companies, one which has a high probability of repaying the loans, π_H, and the other type has a low probability of repaying the loans, $\pi_L < \pi_H$; we know that there is a fraction p of companies with a high repayment rate and a fraction $1 - p$ of companies with a low repayment rate. Neither banks nor depositors know which repayment rate the companies that have obtained loans applies.

A loan needs to be liquidated if depositors withdraw early and the bank requires the proceeds from the liquidation of the loan to repay these withdrawn deposits. If depositors withdraw, the bank obtains $\lambda\pi_i\,(1 + r_L)\,L < L$, depending on the type of company the loan has been granted to, while it would obtain $\pi_i\,(1 + r_L)\,L$ if deposits remain with the bank, this gives banks a net benefit of $(1 - \lambda)\,\pi_i\,(1 + r_L)\,L$ from depositors not withdrawing. To provide an incentive for depositors to retain their deposits at the bank, assume that banks are sharing a fraction α of these benefits with depositors, giving depositors a benefit of not withdrawing of $\alpha\,(1 - \lambda)\,\pi_i\,(1 + r_L)\,L$, which will be the interest they obtain on their deposits if not withdrawing; depositors withdrawing will not be paid interest. The total repayment to depositors not withdrawing will be with deposits D and the implied deposit rate r_D, will be $(1 + r_D)\,D = D + \alpha\,(1 - \lambda)\,\pi_i\,(1 + r_L)\,L$. With banks relying fully on deposits to finance their loans, thus $L = D$, these repayments to depositors will still allow banks to be profitable as we assume that $(1 + r_D)\,D \leq \pi\,(1 + r_L)\,L$, which solves for the requirements that $1 + r_L \geq \frac{1}{\pi_i(1-\alpha(1-\lambda))}$.

Let us now assume that banks promise a repayment to depositors of $\hat{D} = \pi_L\,(1 + r_L)\,L$ and the remainder of the loan is raised as equity. There is no incentive for depositors to withdraw as their deposits can always be repaid from the repayments of the loans, regardless of the type of company that has obtained the loan; we call such deposits safe. When withdrawing deposits, the loan is liquidated causing the depositor a loss that cannot occur when remaining with the bank.

In the case that $\pi_L \leq \lambda\pi_H$ liquidation would result in a certain loss to depositors, even if the repayment rate of loans is high at π_H; hence the benefits from not withdrawing, as determined above, are provide to depositors in this case. If the company with a high repayment rate has obtained the loan, which happens with probability p, the payment to depositors is $\hat{D} + \alpha\,(1 - \lambda)\,\pi_H\,(1 + r_L)\,L = (1 + \alpha\,(1 - \lambda))\,\pi_H\,(1 + r_L)\,L$ and if the company with a low repayment rate has obtained the loan, only the initial deposits $D = \pi_L\,(1 + r_L)\,L$ can be repaid as the bank has no additional resources. We thus have the profits of depositors given by

$$\begin{aligned}\Pi_D^L &= (p\,(\pi_L + \alpha\,(1 - \lambda))\,\pi_H + (1 - p)\,\pi_L)\,(1 + r_L)\,L - D \qquad (14.4)\\ &= (\pi + \alpha p\,(1 - \lambda)\,\pi_H)\,(1 + r_L)\,L - D.\end{aligned}$$

In the case that $\pi_L > \lambda\pi_H$, liquidation would not cause depositors to incur a loss if they were to withdraw in the case that the loan is given to companies with a high repayment rate. If the company with the high repayment rate obtains the loan, the bank would generate a surplus of $(\pi_H - \pi_L)\,(1 + r_L)\,L$, of which depositors would obtain a fraction α, in addition to their initial deposit, to ensure they do not withdraw. If the company with the low repayment rate obtains the loan, the bank will only able to repay its deposits and has no funds left for additional payments. Thus the profits of the depositor are

$$\begin{aligned}\Pi_D^H &= (p(\pi_L + \alpha(\pi_H - \pi_L)) + (1-p)\pi_L)(1+r_L)L - D \\ &= (\pi_L + \alpha p(\pi_H - \pi_L))(1+r_L)L - D.\end{aligned} \tag{14.5}$$

If the bank promises to repay depositors $\hat{D} = \pi_H(1+r_L)L$, it cannot guarantee this repayment; such deposits are risky. If the repayment rate on the loan is high at π_H, the depositor obtains its agreed repayment and the bank is left with no other funds to share with depositors. However, if the repayment rate is low at only π_L, the deposit cannot be repaid in full. The depositor will obtain the repayment from the loan, in addition to the benefits the bank gives to prevent the withdrawal of deposits as outlined above. This gives us depositor profits of

$$\begin{aligned}\hat{\Pi}_D &= (p\pi_H + (1-p)(1+\alpha(1-\lambda))\pi_L)(1+r_L)L - D \\ &= (\pi + \alpha(1-p)(1-\lambda)\pi_L)(1+r_L)L - D.\end{aligned} \tag{14.6}$$

We see that the safe deposits are preferred if the profits to depositors are higher than for risky deposits, $\Pi_D^L \geq \hat{\Pi}_D$ and $\Pi_D^H \geq \hat{\Pi}_D$, respectively. In the case of the low repayment rate being substantially below the high repayment rate, $\pi_L \leq \lambda\pi_H$, thus a situation in which the uncertainty on the profitability of the bank is particularly high, this condition becomes

$$\frac{\pi_H}{\pi_L} \geq \frac{1-p}{p}. \tag{14.7}$$

Hence if the differences in the risks between the two types of companies is particularly large; this is only more restrictive than the requirement that $\pi_L \leq \lambda\pi_H$ if $p \leq \frac{\lambda}{1+\lambda} \leq \frac{1}{2}$.

Similarly for the case that $\pi_L \geq \lambda\pi_H$, we obtain that $\Pi_D^H \geq \hat{\Pi}_D$ if

$$\frac{\pi_H}{\pi_L} \leq \frac{1-\alpha p-(1-p)(1+\alpha(1-\lambda))}{p(1-\alpha)}. \tag{14.8}$$

In this case the differences between the risks between the two companies must not be too large and this condition is only more restrictive than the requirement that $\pi_L > \lambda\pi_H$ if $p \leq \frac{\alpha\lambda}{\alpha(1+\lambda)-1}$.

Combining these results, we see that safe deposits are preferred if the low repayment rate π_L is sufficiently far away from $\lambda\pi_H$. If $\frac{1-p}{p} \leq \frac{\pi_H}{\pi_L} \leq \frac{1-\alpha p-(1-p)(1+\alpha(1-\lambda))}{p(1-\alpha)}$, then the safe deposit is always preferred, which is the case for $p \geq \frac{1-\alpha\lambda}{2-\alpha(1+\lambda)} \geq \frac{1}{2}$. It is thus that safe deposits are always preferred if the fraction of low-risk companies is sufficiently large, while for smaller fractions of low-risk companies risky deposits might be preferable in some situations where π_L is close to $\lambda\pi_H$.

It is intuitively clear that in case the risks from the low repayment are substantial, thus π_L is very low, the possible repayments to depositors from risky deposits are so low in the case of the low repayment of loans being realised, that the promised sharing of any benefits from depositors remaining with the bank are not able to compensate these low repayments. If the low repayment rate is sufficiently close to the high repayment rate, then the benefits from being exposed to the additional risks are low, especially as these benefits are shared with depositors; consequently, safe deposits are preferred. It is only in an intermediate range where $\pi_L \approx \lambda \pi_H$ that risky deposits might be preferred if the fraction of low-risk companies is sufficiently high such that $p < \frac{1-\alpha\lambda}{2-\alpha(1+\lambda)}$.

We have thus seen that in most cases depositors prefer safe deposits that will be repaid in full, regardless of the loan repayments the bank obtains. Central for this result was that the high repayments offered for risky deposits were unlikely to materialise and the sharing of benefits from depositors remaining with the bank were not sufficient to compensate for this risk.

Reading Diamond & Rajan (2000)

14.3 Optimal depositor protection

Bank accounts are not only a way to invest funds with banks as deposits, but they provide additional benefits to account holders, such as the ability to make and receive payments. This ability to make payments will provide depositors with additional benefits, in addition to the interest earned on any deposits. Furthermore, banks may provide depositors with additional protection against their own failure, and hence the loss of deposits, by obtaining a deposit insurance. Deposit insurance will make payments to depositors if the bank is not able to repay depositors themselves. Providing such deposit insurance can be seen as part of the deposit contract, in addition to the deposit rate.

Let us assume that banks finance their loans L entirely through deposits D such that $D = L$ and thus banks hold no equity, and promise to pay depositors interest r_D. The loans the bank gives using these deposits are repaid with a probability π and if the loans the loans are not repaid, we assume the bank does not obtain any payments from their borrowers. Banks in addition buy insurance against the default of their loans such that the deposit insurance pays the bank a fraction λ of the outstanding loan amount, which with interest r_L is an amount of $(1 + r_L) L$. The resources available to the bank to repay their depositors is now given by $\lambda (1 + r_L) D$ in the case the loans are not repaid. Depositors are due to be repaid the amount of $(1 + r_D) D$, but will only receive a fraction $\hat{\lambda}$, which is determined by setting the resources the bank has available from the insurance payout equal to the amount they pay depositors, thus

$$\lambda (1 + r_L) D = \hat{\lambda} (1 + r_D) D,$$

which then easily solves for the implied level of protection of depositors of

$$\hat{\lambda} = \lambda \frac{1 + r_L}{1 + r_D}. \tag{14.9}$$

If this level of implied depositor protection exceeds $\hat{\lambda} \geq 1$, deposits are fully repaid as depositors are never repaid more than they are entitled to. If the implied protection is imperfect, $\hat{\lambda} < 1$, depositors are making a loss.

A full repayment in case the loans are not repaid, $\hat{\lambda} \geq 1$, is given if

$$\lambda \geq \lambda^* = \frac{1 + r_D}{1 + r_L}. \tag{14.10}$$

Depositors do not only benefit from the interest on their deposits, but also from access to other services the bank offers, for example payment services. Let us assume that these services provide a benefit B to depositors. We assume however, that this value is only generated if the deposit is repaid in full. If the deposit is not or not fully repaid, the additional costs of recovering deposits from the deposit insurance, delays in insurance payouts, and changing banks, eliminate any such benefits.

The profits of depositors consist of a situation in which the loan is repaid to the bank π and the depositor is repaid, including interest, and obtains the benefits of additional services, B. If the loan is not repaid, the bank has to rely on the insurance payout to pay depositors. If $\lambda \geq \lambda^*$ deposits are fully repaid and the depositor obtains its benefits B. If, however, $\lambda < \lambda^*$, the deposit is not fully repaid, but only a fraction $\hat{\lambda}$, and they do not obtain the benefits from additional services. Neglecting that depositors could invest into a risk-free asset, their profits are given by

$$\Pi_D = \begin{cases} \pi\left(\left(1 + r_D\right) D + B\right) & \\ \quad + \left(1 - \pi\right)\left(\left(1 + r_D\right) D + B\right) - D & \text{if } \lambda \geq \lambda^* \\ \pi\left(\left(1 + r_D\right) D + B\right) & \\ \quad + \left(1 - \pi\right) \hat{\lambda} \left(1 + r_D\right) D - D & \text{if } \lambda < \lambda^* \end{cases}. \tag{14.11}$$

Let us now assume that deposit markets are competitive such that $\Pi_D = 0$. Inserting for $\hat{\lambda}$ from equation (14.9), this solves for

$$1 + r_D = \begin{cases} 1 - \frac{B}{D} & \text{if } \lambda \geq \lambda^* \\ \frac{1}{\pi} - \frac{B}{D} - \frac{1-\pi}{\pi} \lambda \left(1 + r_L\right) & \text{if } \lambda < \lambda^* \end{cases}. \tag{14.12}$$

Banks obtain insurance that covers some of their payments to depositors if the loans are not repaid. In a competitive insurance market, the insurance premium P, will be equal to the expected payments of the insurance. These payments consist of a fraction λ of the loan the banks were entitled to, $\left(1 + r_L\right) L$, which is payable only if the loan is not repaid. Thus when using that $L = D$, we obtain this insurance premium as

$$P = \left(1 - \pi\right) \lambda \left(1 + r_L\right) D. \tag{14.13}$$

We can determine the bank profits in the case deposits are fully covered as follows. If the loan is repaid, π, the bank obtain the loan and repays its depositors in full, retaining the difference; if the loan is not repaid, it receives an insurance payout of a fraction λ of the loan amount due and can repay its depositors fully, retaining the difference. If deposits are not fully covered by insurance, the bank will not obtain any profits if the loan is not repaid as all proceeds of the insurance payout will go to depositors. Of course, in both cases, the insurance premium P has to be paid. We thus obtain

$$\Pi_B = \begin{cases} \pi\left(\left(1+r_L\right)D - \left(1+r_D\right)D\right) & \\ \quad + \left(1-\pi\right)\left(\lambda\left(1+r_L\right)D - \left(1+r_D\right)D\right) & \\ \quad - P & \text{if } \lambda \geq \lambda^* \\ \pi\left(\left(1+r_L\right)D - \left(1+r_D\right)D\right) - P & \text{if } \lambda < \lambda^* \end{cases} . \tag{14.14}$$

Inserting equations (14.12) and (14.13) into equation (14.14) we obtain the profits of banks buying insurance cover as

$$\Pi_B = \begin{cases} \pi\left(1+r_L\right)D + B - D & \text{if } \lambda \geq \lambda^* \\ \pi\left(1+r_L\right)D + \pi B - D & \text{if } \lambda < \lambda^* \end{cases} . \tag{14.15}$$

We see that the bank profits are higher in the case that $\lambda \geq \lambda^*$ and hence if banks purchase deposit insurance, they will seek to ensure they fully insure their deposits.

Returning to the bank profits as represented in equation (14.14), we can see that if $\lambda < \lambda^*$, the bank profits are maximized if no deposit insurance is purchased such that $P = 0$, which using equation (14.13) implies that $\lambda = 0$ and no coverage of deposits is available. Without any deposit insurance, we easily see from equation (14.14) that

$$\Pi_B = \pi\left(\left(1+r_L\right)D - \left(1+r_D\right)D\right). \tag{14.16}$$

We can now compare the profits of a bank purchasing full insurance, $\lambda \geq \lambda^*$, from equation (14.15) with the profits of a bank purchasing no insurance, $\lambda = 0$, from equation (14.16) and we easily see that the former is giving the bank higher profits is

$$B \geq \left(1 - \pi\left(1+r_D\right)\right)D. \tag{14.17}$$

Thus, only if the benefits of holding a bank account and accessing additional services, B, are sufficiently high, will banks insure their deposits. Most notably, if $\pi\left(1+r_D\right) < 1$, banks would not seek to insure their deposits; implying that if banks provide risky loans with a low probability of success, provided the deposit rate is low, they do seek any such insurance. If the bank seeks deposit insurance, the additional benefits depositors obtain from holding an account are more valuable to depositors if their deposits are repaid fully, allowing for lower deposit rates and higher profits for banks. This lower deposit rate will allow for sufficient profits to be generated to pay the insurance premium.

An implication of our results is that banks whose depositors place a high value on the additional benefits a bank account provides them with, will seek to insure themselves against failing to repay their depositors. On the other hand, banks whose accounts provide very little added value to depositors, beyond earning interest on their funds, will not seek to insure these deposits as the lower deposit rate does not compensate for the insurance premium it needs to pay.

Reading Merton & Thakor (2019)

14.4 Competition for deposits

Banks will compete for depositors as much as they keep for companies to provide loans for. With depositors concerned about the ability by banks to repay their deposits, they will pay particular attention to the risks banks take in providing loans. In addition, depositors may have preferences for a particular bank, for example due to the range of account services that are available, and depositing their funds with another than their preferred bank, would reduce the benefits they obtain from the interest earned on their deposits. When setting deposit rates, banks will take into account these preferences of depositors, but also the banks take in providing loans and hence the risks they expose depositors to.

Let us assume that there are two banks competing for deposits, each providing loans that have different probabilities π_i to be repaid. Depositors are having preferences for one bank over the other bank, for example arising out of other accounting services. In line with the Hotelling model of spatial competition, we therefore position the two banks at the ends of a line of length 1 and potential depositors are distributed evenly along this line. Their position on this line represents the best position a bank could have and the further the distance of the bank from their position, the lower their utility. We assume that at a distance of 1, depositors lose utility c. If a bank repays depositors only if the loan they have provided is repaid, then the depositor obtains its deposit bank with probability π_i. Being a distance d_i away from the bank, their profits from depositing their funds D with bank i are thus given by

$$\Pi_D^i = \pi_i \left(1 + r_D^i\right) D - c d_i D, \tag{14.18}$$

where r_D denotes the deposit rate. A depositor prefer bank i over bank j if $\Pi_D^i \geq \Pi_D^j$. Acknowledging that $d_i + d_j = 1$ as banks are located at this distance, this condition becomes

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\pi_i \left(1 + r_D^i\right) - \pi_j \left(1 + r_D^j\right)}{2c}. \tag{14.19}$$

Hence all depositors that are having a distance from bank i of less than d_i^* will deposit their monies with this bank, all other depositors will choose bank j. Thus, d_i^* is the market share of bank i. With total deposits D available from all depositors, the bank would obtain deposits of $D_i = d_i^* D$. These deposits are now lent out at a loan rate r_L, such that the profits of the bank are given by

$$\Pi_B^i = \pi_i \left(1 + r_L\right) D_i - \left(1 + r_D^i\right) D_i, \tag{14.20}$$

where we assume that loans are fully financed by deposits. Inserting for $D_i = d_i^* D$ and for d_i^* from equation (14.19) we get the first order condition for the optimal deposit rate as

$$\frac{\partial \Pi_B^i}{\partial \left(1 + r_D^i\right)} = -D_i + \left(\pi_i \left(1 + r_L\right) - \left(1 + r_D^i\right)\right) \frac{\pi_i}{2c} D = 0, \tag{14.21}$$

which, after inserting for D_i and D solves for

$$1 + r_D^i = \frac{\pi_i^2 (1 + r_L) + \pi_j \left(1 + r_D^j\right) - c}{2\pi_i}. \tag{14.22}$$

We can easily obtain that

$$\begin{aligned} \frac{\partial \left(1 + r_D^i\right)}{\partial \pi_j} &= \frac{1 + r_D^j}{2\pi_i} > 0, \\ \frac{\partial \left(1 + r_D^i\right)}{\partial \pi_i} &= \frac{1 + r_L}{2} - \frac{\pi_j (1 + r_D) + c}{2\pi_i^2} \lesseqgtr 0. \end{aligned} \tag{14.23}$$

Thus, we see from the first equation that if the other bank is providing loans that are repaid with a higher probability, the deposit rate can be increased. This is because it will reduce the other bank's attractiveness to depositors as their deposits are less likely to be lost, increasing their market share, allowing them to increase their deposit rate to increase bank profits. In turn, the bank will be able to raise its deposit rate as well. The effect of providing loans that are more likely to be paid on its own deposit rate is ambiguous. The higher value for π_i makes the bank more attractive to depositors and hence lower deposit rates can be paid without losing them; on the other hand, this reduced deposit rate decreases the market share of the bank, who then has to increase the deposit rate in order to capture more distant depositors. Which effect dominates, will depend on the strength of each effect.

We will now establish under which conditions banks are attracting and willing to accept deposits.

Monopoly If we assume that depositors can also invest their monies into risk-free assets at an interest rate r, banks will only attract any depositors if the profits from deposits, Π_D^i exceeds that of investing into the risk-free asset, $(1 + r) D$. For a bank to be active in the market we need the condition $\Pi_D^i \geq (1 + r) D$ to be fulfilled only for a single depositor and the highest profits are given for a depositor with distance $d_i = 0$. Inserting these relationships into equation (14.18), we obtain that a bank is attracting deposits only if $\pi_i (1 + r_D) \geq 1 + r$.

A bank will only accept deposits if this is profitable for them, thus we require $\Pi_B^i \geq 0$, which using equation (14.20), becomes $\pi_i (1 + r_L) \geq (1 + r_D)$. We can combine these two conditions an attracting and accepting deposits and obtain $\pi_i (1 + r_L) \geq 1 + r_D^i \geq \frac{1+r}{\pi_i}$. A feasible solution for the deposit rate only exists if

$$\pi_i^2 \geq \frac{1 + r}{1 + r_L}. \tag{14.24}$$

If this condition is not fulfilled, the bank does not accept deposits as no deposit rate can be found that is profitable to both the bank and the depositor. In the case that this condition is violated by both banks, no deposits are taken in the economy at all.

If $\pi_i^2 \geq \frac{1+r}{1+r_L}$ and $\pi_j^2 < \frac{1+r}{1+r_L}$, only one bank, bank i, is active in the market, enjoying a monopoly.

In such a monopoly, bank i will attract all those depositors with positive profits, thus $\Pi_D^i \geq 0$, which when using equation (14.18) requires that

$$d_i \leq d_i^{**} = \frac{\pi_i \left(1 + r_D^i\right) - (1 + r)}{c}. \tag{14.25}$$

Using the market share d_i^{**} for bank i and noting that the total deposits attracted will be $D_i = d_i^{**} D$, we can insert this relationship into equation (14.20) and maximize for the optimal deposit rate set by banks maximizing their profits. This gives us the deposit rate as

$$1 + r_D^i = \frac{\pi_i^2 (1 + r_L) + (1 + r)}{2\pi_i} \tag{14.26}$$

and hence after inserting this result into the equation (14.25), the market share of bank i becomes

$$d_i^{**} = \frac{\pi_i^2 (1 + r_L) - (1 + r)}{2c}. \tag{14.27}$$

This maximization is only relevant if $d_i^{**} \leq 1$ and not all depositors will use banks, which implies that

$$\pi_i^2 \leq \frac{2c + (1 + r)}{1 + r_L}. \tag{14.28}$$

As the most that a bank can capture is the full market, a higher success rate would imply that banks capture the entire market. In this case the bank could extract any surplus of the most distant depositor ($d_i = 1$) such that $\pi_i \left(1 + r_D^i\right) - c = 1 + r$, or

$$1 + r_D^i = \frac{c + (1 + r)}{\pi_i}. \tag{14.29}$$

We illustrate these results graphically in figure 14.1. If the condition in equation (14.24) is fulfilled for both banks, they will both attract deposits and be willing to accept them. In the case that $d_i^{**} + d_j^{**} < 1$, banks will enjoy a local monopoly as there is a market that has not been served by either bank. Thus not all potential depositors will use a bank. In this case the two banks operate independently as they are not directly competing and the deposit rates are given by equation (14.26) and the respective market shares by equation (14.27). When inserting for the market shares from equation (14.25), the condition that $d_i^{**} + d_j^{**} < 1$ solves for

$$\pi_i^1 + \pi_j^2 < 2\frac{c + (1 + r)}{1 + r_L}. \tag{14.30}$$

Monopolistic competition We can now increase the probabilities of the loans banks have given to be repaid, π_i and π_j such that the condition in equation (14.30) is not fulfilled and hence $d_i^{**} + d_j^{**} \geq 1$. In this case, banks are engaged in monopolistic competition. In this case deposit rates are given by equation (14.22) and market shares by equation (14.19). Solving equation (14.22) for $1 + r_D^i$ by inserting from $1 + r_D^j$, we easily get

$$1 + r_D^i = \frac{2}{3} \frac{\pi_i^2 + \frac{1}{2}\pi_j^2}{\pi_i} (1 + r_L) - \frac{c}{\pi_i} \tag{14.31}$$

and from inserting this expression into equation (14.19), we get the market share of bank i as

$$d_i^* = \frac{\pi_i^2 + \pi_j^2}{2c} (1 + r_L) - \frac{1}{2}. \tag{14.32}$$

For both banks to be accepting deposits, we require that no bank obtains the entire market, thus $d_i < 1$, or

$$\pi_i^2 + \pi_j^2 < \frac{c}{1 + r_L}. \tag{14.33}$$

If this condition is not fulfilled, one bank, the bank with the higher success rate, will be covering the entire market and the the only active bank. This is not because this bank is a monopolist; the other bank would like to enter the market, but because of the low probability of the deposit being returned, their terms are not attractive enough to depositors and indirectly market entry by this bank is deterred.

The bank not attracting any deposits could set a deposit rate that allowed it to break even, $\Pi_B^j = 0$, from which we obtain $1 + r_D^j = \pi_j (1 + r_L)$ using equation (14.20). Inserting this deposit rate into equation (14.19) and noting that we require $d_i^* = 1$ to cover the whole market, we easily get the deposit rate applied by the bank remaining in the market as

$$1 + r_D^i = \frac{\pi_j^2 (1 + r_L) + c}{\pi_i}. \tag{14.34}$$

Summary When setting deposit rates, banks will take into account that depositors are concerned about the risks bank face. Their deposit rates will not only reflect the risks they are exposing depositors to through their provision of loans, but also that of their competitors. Banks that are taking high risks in their lending might find themselves in a situation where they are not attracting any depositors, either because they are not able to offer deposit rates that are beneficial to depositors and at the same time profitable to them, or other banks are offering depositors which are much less risky and even the lowest loan rate they could offer would not suffice to compete with these banks. If generally the risks by all banks are high, banks might not be attractive to all potential depositors and they would therefore not use banks for investing their funds, leaving banks with a smaller market.

Competition between banks for deposits can fail if large discrepancies in the risk to depositors exist. We should therefore expect to find that banks providing more risky loans have a smaller market share in the deposit market, while offering higher deposit rates to compensate for this additional risk. Such a scenario can lead to banks facing an imbalance between the deposits they can attract and the amount of loans they are able to give, causing such banks to look for alternative funding sources. Similarly banks that provide only loans with low risk may find themselves in a situation where they attract more deposits than they are able to lend out; in this case banks may seek alternative investment opportunities for their excess deposits. Such an investment might be an interbank loan to another, more risky bank seeking such additional funding, opening the way to interbank markets.

Reading Matutes & Vives (1996)

Conclusions

Depositors are compensated for the risks that banks take and which may lead to them not being able to repay deposits. This risk will be included into the deposit rate, but the effect of any deposit insurance will be accounted for. Such deposit insurance reduces the risk to depositors and will therefore reduce the deposit rate required. However, depositors in most cases would prefer to avoid taking risks and choose risk-free deposits, if these are available. Deposit insurance is offered to attract depositors who value the account services highly. It is these additional benefits to depositors that can be used to lower deposit rates and thus allow banks to make higher profits, assuming that account services are not too expensive to provide. Only banks whose services are sufficiently valued will provide deposit insurance and we would therefore expect to see that banks offering only a basic service are not seeing much value in the additional costs of deposit insurance.

Banks will naturally differ in the type of account services they offer and depositors will have their preferences. Also taking into account that banks take different risks, competition for depositors will balance these two aspects. It can well be that banks who provide more risky loans than other banks are not attractive to depositors and will be squeezed out the deposit market, having to reply on alternative funding sources for their loans.

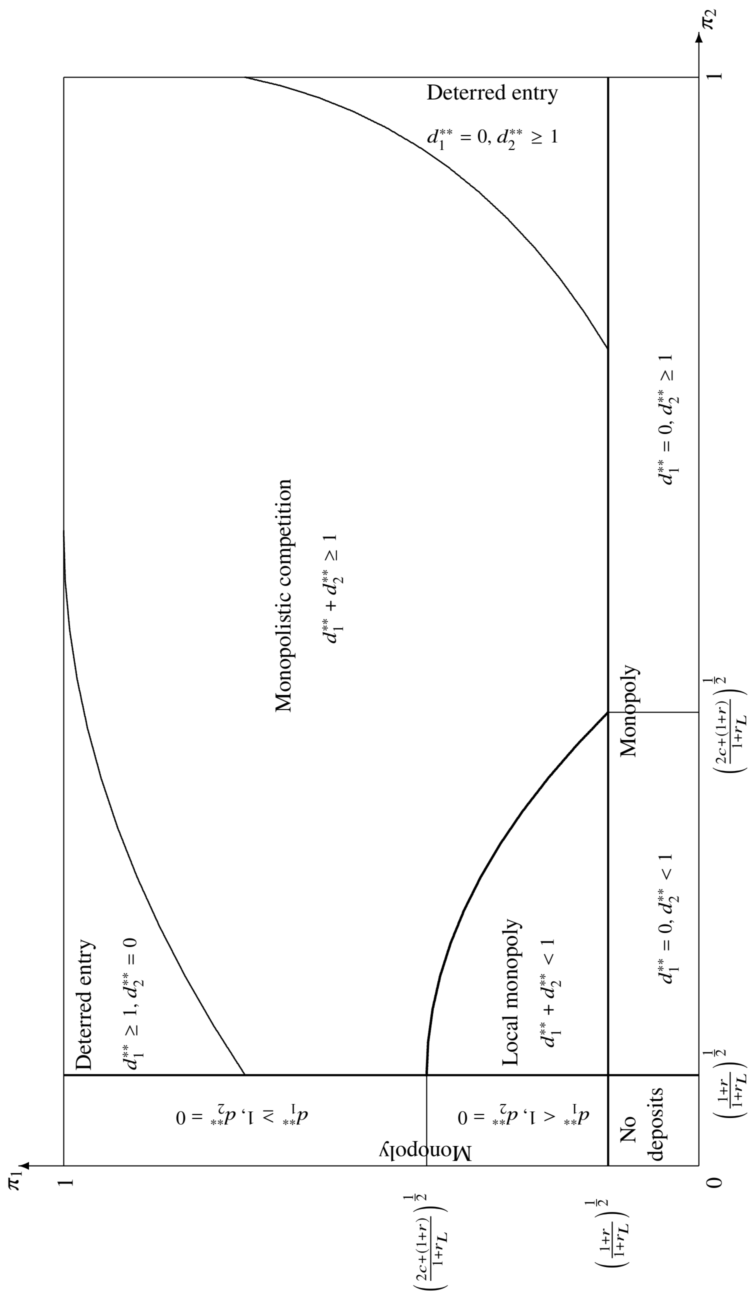


Fig. 14.1: Equilibria characteristics in the deposit market

Chapter 15

Bank runs

Deposits are provided to banks in the understanding that they can be withdrawn at any time without any restrictions. This is sometimes referred to as demand deposits to distinguish them from time deposits, who have a fixed time of maturity before which they cannot be withdrawn. Focussing on the more common demand deposits, the ability of deposits to be withdrawn instantly can be problematic for banks. While banks typically hold a certain amount of cash reserves to meet these withdrawals, the majority of deposits are invested into long-term loans; such loans cannot easily be liquidated, banks would make losses when seeking to sell them in order to raise cash meeting the demand from deposit withdrawals.

While some level of deposit withdrawals is expected and banks will account for this by holding cash reserves, the withdrawals can suddenly increase beyond this level; this is commonly referred to as a bank run. In a bank run, all or a large proportion of deposits are withdrawn suddenly from a bank, often without an apparent reason. The origin that a bank run can occur lies in the mismatch between the long-term loan that is given on the basis of short-term deposits and these loans can only be sustained if deposits are retained by the bank.

The reason for the withdrawal of deposits can broadly be classified as a sudden demand for liquidity by depositors or them receiving unfavourable information about the bank. Depositors would demand liquidity, that is withdrawing deposits from banks, if they expect their deposits not to be safely returned to them in the future. As chapter 15.1 will discuss, such concern might arise from the expectation that other depositors might withdraw and hence impose losses on banks from forcing the liquidation of loans, which would endanger the safety of the deposits that are not withdrawn. Facing such possible losses, a depositor itself would withdraw, increasing the bank run.

Bank runs may not only occur as the result of other depositors withdrawing, but also because of negative information about a bank becoming available as shown in chapter 15.2. Receiving negative information about a bank's ability to honour the repayment of deposits, may trigger depositors to withdraw as long as the bank still has the resources to make payments. Of course, such possible losses arising from bank runs should affect the value of deposits and hence the deposit rate that is required to compensate depositors for any such risk. Chapter 15.3 will show how deposit rates may change if bank runs are a possibility. The competition between banks will affect their behaviour, and hence their ability to accommodate deposit withdrawals. How competition affects banks being able to withstand banks runs is explored in chapter 15.4.

15.1 Liquidity demand

Deposits can be withdrawn at any time and doing so will exhaust the cash reserves of banks. In order to meet the demand by depositors, banks need to raise additional liquidity by either selling assets, most notably loans, or raising funds from other sources, for example the interbank market, institutional investors, or the central bank. Selling assets, especially if this has to be done quickly as a fire sale, will cause losses to the bank, which will impede their ability to have sufficient assets to repay the remaining depositors. Similarly, the higher costs associated with raising liquidity from other sources, will reduce the banks profits. In this situation, it can be profitable for depositors to withdraw in order to ensure they obtain a repayment of their deposits, rather than retain their deposits with the bank and make losses once they are eventually repaid. In chapter 15.1.1 we see how the expectations about the behaviour of other depositors can trigger a bank run and chapter 15.1.2 shows how a lack of coordination in the behaviour of depositors can lead to a bank run that is detrimental to all depositors. Chapter 15.1.3 shows that a liquidity shock by banks may nevertheless result in a bank run, even if banks could raise sufficient cash reserves.

15.1.1 The breakdown of liquidity insurance

Banks accept deposits in the understanding that these can be withdrawn at any time. As deposits are invested into loans that cannot be called in quickly, even if a small fraction os help as a cash reserve, any sudden and large withdrawal of deposits will not allow the bank to return these deposits without having to generate cash by selling the loans they made. Such loans ales are often only possible at a loss, making the ability to return all deposits uncertain. If depositors think that their deposits cannot be returned in full, they might be tempted to withdraw them early and thereby cause a bank run. Such a withdrawal will occur if the depositor anticipates that his returns from withdrawing deposits exceeds that of keeping the deposits with the bank.

Banks obtain deposits D and from this provide loan of L, such that they retain $D-L$ as cash reserves to repay any depositors with drawing their funds. This amount of cash reserves is determined such that banks can meet the demand of depositors they expect to withdraw due to them requiring these funds for consumption. Assume now that an additional amount γL is withdrawn by depositors, even though there is no need for them to do so.

In order to generate the additional cash required, banks need to liquidate loans to the value of $\hat{L}$. We assume that when liquidating loans, only a fraction λ of its value can be generated, thus we generate cash of $\lambda\hat{L}$. In order to meet the additional demand for cash reserves, we require the amount of loans to be sold meet the requirement that $\gamma L = \lambda\hat{L}$, or $\hat{L} = \frac{\gamma}{\lambda}L$, where of course we would require $\frac{\gamma}{\lambda} \leq 1$ or $\lambda \geq \gamma$ to allow the bank to raise sufficient reserves to meet the additional deposit withdrawal. The return of the remaining loans, $L - \hat{L}$, are now distributed amongst the remaining depositors, $L - \gamma L$, where the remaining deposits are L due to the reserves of $D - L$ being withdrawn by depositors for consumption.

Loans have been made at a loan rate r_L and the loans are repaid with probability π and hence the return they generate for depositors not withdrawing funds will be

$$\begin{aligned} 1 + \hat{r}_D &= \pi\left(1 + r_L\right)\frac{L - \hat{L}}{D - \gamma L} \\ &= \pi\left(1 + r_L\right)\frac{\lambda - \gamma}{\lambda\left(1 - \gamma\right)}. \end{aligned} \tag{15.1}$$

Deposits that are not withdrawn are repaid at face value such that their return is given by $1 + r_D = 1$. This return of not withdrawing deposits is higher than the return generated after the withdrawal of deposits if $\pi\left(1 + r_L\right)\frac{\lambda-\gamma}{\lambda(1-\gamma)} \geq 1$, or $\lambda \geq \frac{\gamma\pi(1+r_L)}{\pi(1+r_L)-(1-\gamma)} > \gamma$. Thus if the value from liquidating loans is sufficiently high, depositors would not withdraw.

A crucial assumption in obtaining this result was that $\frac{\gamma}{\lambda} \leq 1$ and hence sufficient cash reserves could be generated to meet the demand of all depositors withdrawing. Let us now assume that $\frac{\gamma}{\lambda} > 1$. In this case all loans are liquidated and those depositors not having withdrawn will not be able to obtain any repayment as no assets are left with the bank, hence $1 + \hat{r}_D = 0$. Of those withdrawing deposits, their entire demand cannot be met. The amount of cash available is $D - L + \lambda L = D - \left(1 - \lambda\right)L$, as all loans are sold at a discount λ and $D - L$ denotes the initial cash holding; the amount of deposits being withdrawn early is $\left(D - L\right) + \gamma L = D - \left(1 - \gamma\right)L$, consisting of the withdrawals by those consuming in time period 1, $D - L$, in addition to those withdrawing without having to consume, γL. Hence the return generated from withdrawing deposits is given by

$$1 + r_D = \frac{D - \left(1 - \lambda\right)L}{D - \left(1 - \gamma\right)L} > 0. \tag{15.2}$$

As $L \leq D$, this expression will be positive and hence be higher than the return of $1 + r_D = 0$ if not withdrawing deposits. Hence it is more profitable to withdraw deposits if $\gamma > \lambda$.

We can now combine this result with our finding that for $\lambda < \frac{\gamma\pi(1+r_L)}{\pi(1+r_L)-(1-\gamma)}$ it was also more profitable to withdraw deposits. As this term can easily be shown to be larger than γ, this is the more restrictive constraint. Re-ordering this condition for deposit withdrawals, we obtain that withdrawals are optimal if

$$\gamma > \gamma^* = \lambda\frac{\pi\left(1 + r_L\right) - 1}{\pi\left(1 + r_L\right) - \lambda}. \tag{15.3}$$

Hence if sufficient depositors withdraw, it is optimal for all depositors to withdraw.

The emerging equilibrium of withdrawing deposits is one of self-fulfilling prophecies. If depositors expect that a fraction $\gamma \leq \gamma^*$ of depositors withdraw, that is only relatively few depositors are withdrawing no depositor will withdraw as this generates a too low return to depositors, fulfilling the expectation that not enough depositors withdraw to cause remaining depositors a loss. On the other hand, if depositors expect that a relatively large fraction $\gamma > \gamma^*$ of depositors will withdraw, every depositor would withdraw in order to obtain a positive return; thus, as all depositors withdraw, the expectations of many depositors withdrawing is fulfilled. Therefore, expecting few deposit withdrawals (no bank run) will see no bank run and expecting many deposit withdrawals (a bank run) will see a bank run. This result is independent of any fundamental information about the bank, but merely based on the expectation of the bejaviour of other depositors.

The withdrawal of deposits, and hence a bank run, depends on the expectations of depositors regarding the behaviour of other depositors. If they expect a sufficiently large fraction of deposits to be withdrawn, they will withdraw deposits themselves. It is irrelevant whether actually deposits have been withdrawn, this will be based solely on expectations about the behaviour of other depositors. Thus banks are susceptible to banks runs arising from the expectation of a bank run and a bank remains unaffected by a bank run as long as depositors belief no bank run will occur.

Reading Diamond & Dybvig (1983)

15.1.2 Coordination of deposit withdrawals

Bank runs occur if deposits are withdrawn early. This might then impose losses on the bank as they need to raise the cash reserves needed to repay the withdrawn deposits, for example through the fire sale of assets, such as loan. In such a situation, the bank is unlikely to obtain the full value of these loans, making a loss that can jeopardize their ability to repay all depositors, those withdrawing early as well as those retaining their deposits with the bank. This can lead to incentives to withdraw deposits early if there are higher losses to be expected when retaining deposits with the bank, causing a bank run. In such a situation it would be beneficial for all depositors to coordinate their withdrawal decision to avoid a bank run and subsequently obtain higher profits for all of them.

Let us assume a bank has raised deposits D on which they are paying interest r_D. These deposits are invested into cash reserves, R, which pay no interest, and loans Lon which the bank charges interest r_L. We thus have that $D = L + R$. Loans are repaid with probability π and we assume that a fraction γ of deposits are withdrawn early, that is prior to the loan being due to be repaid; deposits withdrawn early do not attract any interest payments by the bank. These deposit withdrawals are financed from cash reserves R and, if necessary, the sale of loans $\hat{L}$. When selling loans, the bank cannot realise the full value of these but only obtains a fraction λ of their expected value, $\pi (1 + r_L) \hat{L}$.

Banks only need to sell loans if $R < \gamma D$ as otherwise the amount of cash reserves will be sufficient to repay the deposits withdrawn. If the withdrawal rate γ is such that loans need to be sold, the amount that needs to be sold, $\hat{L}$, is then given by

$$R + \lambda\pi\,(1 + r_L)\,\hat{L} = \gamma D, \tag{15.4}$$

where the left-hand side denotes the amount of cash reserves and the cash raised from the sale of loans and the right-hand side the amount of cash that s needed to repay the deposits withdrawn. Hence the amount of loans that need to be sold is given by

$$\hat{L} = \frac{\gamma D - R}{\lambda\pi\,(1 + r_L)}. \tag{15.5}$$

The bank will fail if, after any withdrawals, the remaining depositors cannot be repaid in full. The resources available to repay deposits that have not been withdrawn consists of the remaining loans, $L - \hat{L}$, the cash reserves R less the amount repaid to depositors withdrawing early, γD. If these resources of the bank are less than the remaining fraction of deposits that need to need to be repaid, $(1 - \gamma)\,(1 + r_D)\,D$, the bank fails. Thus we require for a failing bank that

$$\pi\,(1 + r_L)\left(L - \hat{L}\right) + R - \gamma D < (1 - \gamma)\,(1 + r_D)\,D. \tag{15.6}$$

If $\gamma \leq \frac{R}{D}$, then the reserves are sufficient to meet the demand of all depositors withdrawing early, implying that $\hat{L} = 0$. in this case equation (15.6) solves for

$$\pi < \pi^* = \frac{(1 + (1 - \gamma)\,r_D)\,D - R}{(1 + r_L)\,(D - R)}, \tag{15.7}$$

using that $L = D - R$. Hence we see that for low probabilities of the loans being repaid, the bank cannot meet its obligations to remaining depositors as the revenue generated by loans is not sufficient to repay them. As no loans are sold to repay withdrawn deposits, the bank's loan have such low probabilities of being repaid, that the bank is insolvent due to it not being able to meet its obligations to the remaining depositors. A bank run, that is am early withdrawal of deposits, would be justified by the weak fundamentals, the high risks, of the bank.

At the other extreme, the most loans that can be sold is $\hat{L} = L$, which means from equation (15.4) that a bank would fail to repay all withdrawn deposits if $R + \lambda\pi\,(1 + r_L)\,L < \gamma D$. Hence a bank would fail instantly due to their inability to repay withdrawn deposits if

$$\pi < \pi^{**} = \frac{\gamma D - R}{\lambda\,(1 + r_L)\,(D - R)}. \tag{15.8}$$

Here a too low probability of loans being repaid does not allow banks raise sufficient revenue to meet the demand of those depositors withdrawing. Banks in this situation are failing as they cannot meet the requests for deposit withdrawal and face illiquidity.

The bank run in this case is justified by the inability if the bank to meet its future obligations.

If the withdrawal rate is such that the bank can repay all withdrawn deposits by selling loans, $0 < \hat{L} < L$, we can insert equation (15.5) for the amount of loans that need to be sold into condition (15.6) for the failure of a company. A bank would fail due to not being able to repay the remaining depositors if

$$\pi < \pi^{***} = \frac{(\lambda (1 + (1 - \gamma) r_D) + \gamma) D - (1 + \lambda) R}{\lambda (1 + r_L) (D - R)}. \tag{15.9}$$

The low repayment rate of loans will see the bank left without sufficient assets to repay the remaining depositors, causing its failure. This failure only arises because depositors are seeking to withdraw deposits early, causing losses to banks from being forced to sell loans below their full value, which affects their ability to repay any of the remaining deposits. If depositors retaining their deposits with the bank are obtaining a lower pay out than when withdrawing immediately, it would be rational for these investors to withdraw their deposits early as well. As depositors withdrawing early are only foregoing the interest they are due if retaining the deposit, we can adjust equation (15.6) to represent the situation that the resources available to the bank are not sufficient to repay the remaining depositors their initial deposits: $\pi (1 + r_L) \left(L - \hat{L}\right) + R - \gamma D < (1 - \gamma) D$ and obtain

$$\pi < \hat{\pi}^{***} = \frac{(\lambda + \gamma) D - (1 + \lambda) R}{\lambda (1 + r_L) (D - R)}. \tag{15.10}$$

For such low repayment rates of the loans banks have given, a coordination problem emerges for depositors. Depositors who do not want to withdraw early are incentivized to do so by obtaining a higher repayment when withdrawing early than when retaining their deposits with the bank. It would be optimal for all depositors to not withdraw their deposits early, as this increases the payment they receive from the bank, but if depositors expect other depositors to withdraw early, they would do so too, causing a bank run. Similarly, if depositors expect other depositors to not withdraw early, they would not do so, and a bank run does not emerge. It is thus a case of self-fulfilling prophesies arising out if this problem of depositors coordinating their early withdrawals. A bank run will occur if depositors expect other depositors with withdraw, if the expectation is that other depositors are not withdrawing, no bank run should occur. If, on the other hand, $\pi \geq \hat{\pi}^{***}$, early withdrawals are not more profitable than retaining deposits with the bank. In this case, depositors should not withdraw and a bank run should not occur.

Our results are summarised in figure 15.1. We see that for low repayment rates of the loan the bank provides, banks will fail, either because they are insolvent due to not generating sufficient funds to repay their depositors at all (insolvency), or not being able to satisfy the early withdrawals of deposits as funds that can be generated from selling loans is not sufficient due to their low value. If the funds they can generate from selling loans, they might be able to repay those deposits that are withdrawn early, but then do not have sufficient funds left to repay the remaining

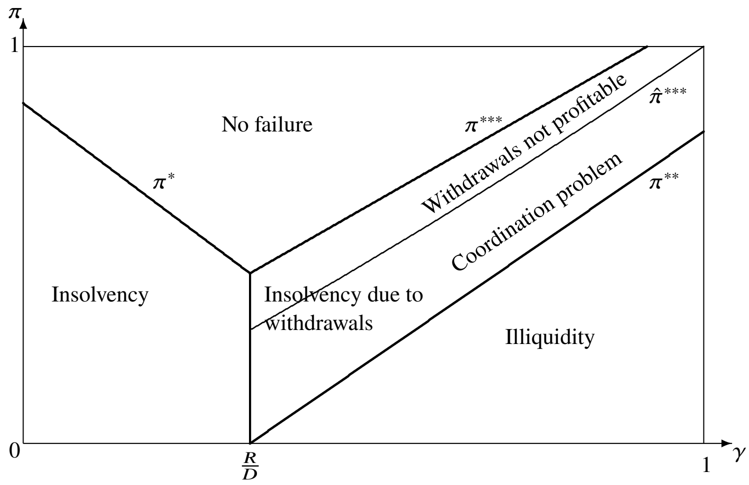


Fig. 15.1: Bank failures due to deposit withdrawals

depositors as the amount of loans that had to be sold was such that not enough funds are retained in the bank. It is in this latter situation that depositors that depositors face a coordination problem, namely if it is more profitable to withdraw deposits early than to retain them with the bank. Here we can see a bank run emerge. If the repayment of loans is sufficiently high and withdrawal rates of deposits not too high, banks will not fail, but be able to meet the demands of those depositors withdrawing early as well as those remaining with the bank.

It is thus that bank runs may occur at banks that provide reasonably risky loans and face higher withdrawal rates of deposits. Here the incentives are such that it becomes unprofitable for depositors to retain their deposits with the bank and they will rather withdraw these early along with other depositors. This leads to self-fulfilling prophecies in that expecting many early withdrawals by other depositors will cause all depositors with withdraw early, while when expecting lower early withdrawal rates, they would not with draw early themselves.

Reading Rochet & Vives (2004)

15.1.3 Liquidity shocks

If a bank faces a liquidity shock, it can raise additional funds by selling assets, most notably loan the bank has provided. Such sales will be conducted at market prices, provided a sale can be agreed. We will here investigate whether assets can be sold and thus a bank failure be avoided.

Let us assume a bank faces a liquidity shock in which as fraction λ of their deposits are withdrawn; such a liquidity shock occurs with probability p. This implies that the bank needs to repay a fraction λ of their deposits D, which can only be achieved by selling loans to outside investors at a price P_1 if sold at the time of the liquidity shock or P_0 if the bank sells loan in anticipation of a possible liquidity shock.

Solvent banks We first consider the case where banks can remain solvent after the liquidity shock, thus they will not fail due to the amount of deposits withdrawn. Selling a fraction γ_1 of the loan book after experiencing a liquidity shock at price P_1, they need to obtain sufficient funds to repay the deposits withdrawn, thus

$$\lambda D = \gamma_1 \pi \left(1 + r_L\right) L P_1, \tag{15.11}$$

where $\pi \left(1 + r_L\right) L$ denotes the expected repayments of loans L, which are repaid, including interest r_L, with probability π. This expression represents the value of the loan at maturity, of which a fraction γ_1 is sold at price P_1. If loans are fully financed by deposits such that $D = L$, then the fraction of loans to be sold is given by

$$\gamma_1 = \frac{\lambda}{\pi \left(1 + r_L\right) P_1}. \tag{15.12}$$

To ensure that the bank does not fail, we require that $\gamma_1 \leq 1$ such that banks can raise sufficient funds from selling loans to repay all withdrawn deposits; this is achieved if the deposit withdrawal is not too large. In this case, the profits of the bank are given by

$$\begin{aligned} \Pi_B^1 = {} & p \left(\left(1 - \gamma_1\right) \pi \left(1 + r_L\right) L - \left(1 - \lambda\right) \left(1 + r_D\right) D\right) \\ & + \left(1 - p\right) \left(\pi \left(1 + r_L\right) L - \left(1 + r_D\right) D\right), \end{aligned} \tag{15.13}$$

where the first term denotes the case where the liquidity shock happens and loans are sold off with the remaining deposits being returned including interest r_D, and the second term the case of no liquidity shock occurring and hence no loans being sold. We can rewrite equation (15.13) as

$$\Pi_B^1 = \pi \left(1 + r_L\right) L - \frac{\lambda p}{P_1} L - \left(1 - \lambda p\right) \left(1 + r_D\right) D, \tag{15.14}$$

using equation (15.12) to replace γ_1.

Alternatively, banks could sell loans in anticipation of the potential liquidity shortfall in the future. Similar to equation (15.12), bank would sell a fraction γ_0 of loans to cover this potential liquidity shortfall at a price P_0, such that we obtain

$$\gamma_0 = \frac{\lambda}{\pi \left(1 + r_L\right) P_0}. \tag{15.15}$$

The bank profits then consist in the case of the liquidity shock occurring of the retained loans less the not withdrawn deposits and if no liquidity shock is observed, of the retained loans, all deposits to be repaid and the funds raised from the prior loan sale. Hence we have

$$\begin{aligned}\Pi_B^0 &= p\left(\left(1-\gamma_0\right)\pi\left(1+r_L\right)L-\left(1-\lambda\right)\left(1+r_D\right)D\right) \\ &\quad +\left(1-p\right)\left(\left(1-\gamma_0\right)\pi\left(1+r_L\right)L-\left(1+r_D\right)D\right. \\ &\quad \left.+\gamma_0\pi\left(1+r_L\right)LP_0\right) \\ &= \pi\left(1+r_L\right)L-\frac{\lambda}{P_0}L-\left(1-\lambda p\right)\left(1+r_D\right)D \\ &\quad +\left(1-p\right)\lambda L,\end{aligned} \tag{15.16}$$

where the last line has been obtained by inserting for γ_0 from equation (15.15).

The relative prices P_1 and P_0 should be such that there is no preference for the bank to either sell early or later, thus we should find that $\Pi_B^1 = \Pi_B^0$, which solves for the prices prior to observing the liquidity shock and after its occurrence being related as

$$P_0^* = \frac{1}{\frac{p}{P_1}+\left(1-p\right)}. \tag{15.17}$$

The buyer of the asset would obtain a return of $\frac{1}{P_0}$ if purchasing prior to the liquidity shock occurring. If they do not obtain the asset then, they might buy it at a later point if the liquidity shock materializes for a return of $\frac{1}{P_1}$, while in the absence of a liquidity shock such buyers would get no return as their money remains uninvested. Equalling these two strategies gives us $\frac{1}{P_0} = \frac{p}{P_1} + (1-p)$, or

$$P_0^* = \frac{1}{\frac{p}{P_1}+\left(1-p\right)}, \tag{15.18}$$

the same relationship as in equation (15.17). Thus, if the bank remains solvent, buyers and the bank can agree the price P_0 and the transaction can commerce, ensuring the bank meets its obligation to depositors.

Insolvent banks Let us now assume that selling the loans at the time of the liquidity shock causes the bank to become insolvent. Hence from the first line of equation (15.13), representing the profits of the liquidity shock occurs, we have

$$\begin{aligned}&\left(1-\gamma_1\right)\pi\left(1+r_L\right)L-\left(1-\lambda\right)\left(1+r_D\right)D \\ &\quad = \pi\left(1+r_L\right)L-\frac{\lambda}{P_1}L-\left(1-\lambda\right)\left(1+r_D\right)D<0.\end{aligned} \tag{15.19}$$

The second line has been obtained by inserting for γ_1 in equation (15.12). These profits are negative as we assume that the bank becomes insolvent if the have not obtained liquidity prior to the liquidity shock. This condition requires that

$$P_1 < \frac{\lambda}{\pi\left(1+r_L\right)-\left(1-\lambda\right)\left(1+r_D\right)}. \tag{15.20}$$

Due to limited liability, however, banks do not have to cover such losses and in the absence of equity will record a profit of zero. Hence, the profits of the bank are only given for the situation in which no liquidity shock occurs, such that

$$\hat{\Pi}_B^1 = (1 - p)\left(\pi\left(1 + r_L\right)L - \left(1 + r_D\right)D\right). \tag{15.21}$$

The bank could sell their loans prior to the liquidity shock, if it was profitable to do so, $\Pi_B^0 \geq \hat{\Pi}_B^1$, where Π_B^0 is given from equation (15.16); selling loans early would ensure that the bank remains solvent. This condition then implies

$$P_0 \geq \hat{P}_0 = \frac{\lambda}{p\pi\left(1 + r_L\right) - (1 - \lambda)\, p\left(1 + r_D\right) + \lambda\,(1 - p)}. \tag{15.22}$$

Using the relationship between the prices for the loan before the liquidity shock and at the liquidity shock, P_0 and p_1 from equation (15.18), we obtain that with the condition of P_1 in equation (15.20), we require that

$$P_0 < \hat{P}_0 = \frac{\lambda}{p\pi\left(1 + r_L\right) - (1 - \lambda)\, p\left(1 + r_D\right) + \lambda\,(1 - p)}, \tag{15.23}$$

violating the requirement in equation (15.23), It is therefore not possible to find a price on which the buyer and seller would agree the sale of the loans as the liquidity shock occurs. As the bank did not sell the loans prior to the observation of the liquidity shock, no loans are sold and the bank cannot raise sufficient liquidity.

The bank thus will find itself in a situation where the limited liability, implying no need to cover losses, induces them to forego selling loans in anticipation of a liquidity shock, but then when the liquidity shock occurs, will not be able to conduct a sale. The consequence is that the bank will be insolvent and face a bank run as a result. This is despite the fact that bank hold sufficient assets to generate the liquidity needed,

Summary We have found that if the liquidity shock is such that a bank would remain solvent when selling their loans at that time, they can readily agree to raise the required liquidity. If the bank would not be solvent if they do not agree a sale of loans prior to the liquidity shock, they would not be able to sell their loans if this liquidity shock occurs as no suitable price can be agreed that retains the solvency of the bank. We have established a situation in which banks cannot sell assets to cover a liquidity shock, and hence avoid a bank run, even though they have sufficient assets that could be sold to cover this shortfall.

Hence it is not only a requirement that the bank has sufficient assets to sell whose value allows to cover any liquidity shortfall, but the price other banks are willing to pay for such assets will have to be sufficiently high. Compared to other investment opportunities banks have, purchasing the loans of a bank facing a liquidity withdrawal is not attractive and hence the price paid for these loans is not sufficient to cover the liquidity shortfall.

Reading Diamond & Rajan (2011)

Résumé

Banks runs emerge from the formation of expectations about the behaviour of other depositors. If a sufficient large proportion of other depositors is expected to withdraw their deposits early, it is rational for a depositor to also withdraw as this will increase their repayments. Retaining deposits as banks seek liquidity at substantial costs can easily lead to a situation where these depositors are not able to be repaid in full. It is then better to withdraw deposits early and obtain a higher repayment. such an equilibrium emerges even if banks are fundamentally healthy and could have repaid all deposits if no early withdrawals would have occurred. The origin of these bank runs are in the assessment of the behaviour of other depositors. Even if bank have sufficient assets to raise cash reserves that cover any deposit withdrawals, we have seen that it might be optimal for banks to not do so at an early stage and then will find themselves unable to agree a sale of assets as a liquidity shock emerges.

15.2 Information-based bank runs

Bank runs do not only emerge as the result of depositors forming expectations about the behaviour of other depositors and acting in an attempt to pre-empt their decisions by withdrawing deposits and securing a higher repayment. There can well be information become available that suggests that the bank will struggle to repay the deposits in the future and it would be beneficial to withdraw them early for as long as existing cash reserves by banks allow them to make full repayments, or they can raise additional funds that can accommodate the first depositors that seek to withdraw. Chapter 15.2.1 looks at how a bank facing a liquidity event that reduces their cash reserves can lead to sequential withdrawals of deposits as information becomes available to depositors, resulting in a slow bank run. Adverse information on a bank can cause depositors to withdraw, in particular if they take into account that other depositors will have obtained similar information, increasing losses due to their withdrawals, as chapter 15.2.2 will show. Meanwhile, chapter 15.2.3 will show how that the lack of information about the risk of a bank may lead to unjustified bank runs or that bank runs that are justified would not occur.

Using loan guarantees, for example provided by governments, reduces the losses banks make when providing loans and should therefore affect the decisions of depositors to withdraw. In chapter 15.2.4 we will therefore explore how such loan guarantees affect the emergence of bank runs.

15.2.1 Sequential deposit withdrawals

Often bank runs are not very sudden withdrawals by all depositors, but deposits are withdrawn slowly over a number of days or even weeks. It is as information about a bank having liquidity problems is spreading that deposits are withdrawn sequentially by depositors until the cash reserves of banks are depleted and the bank fails.

Let us assume that a bank has initially all required cash reserves to meet deposit withdrawals, but at some time t^* a liquidity event occurs that reduces the cash reserves R to a level below that of the deposits D, that are expected to be withdrawn in the normal matter of business, $R < D$. We can assume that these deposits include any accrued interest. This implies that not all of these deposits that are ordinarily withdrawn could be repaid fully. The liquidity event the bank faces could be the withdrawal of deposits by large institutional investors or expected new deposits by such investors not materialising, but also a general and wide spread withdrawal of deposits from the bank, for example due to a recession. Whether such a liquidity event has occurred is not directly observable; it is, however, known that in each time period a liquidity even occurs with probability γ.

Timing of the liquidity event We can now determine the probability that a liquidity event occurs in a time period ranging from t to $t + \Delta t^*$, denoted by θ_t. If we denote by $F_\gamma(t)$ the probability that a liquidity event occurs before time period t, we know that such an event occurs in the interval from t to $t + \Delta t^*$ with probability $F_\gamma(t + \Delta t^*) - F_\gamma(t)$, thus the probability that the liquidity event occurs prior to time $t + \Delta t^*$ less the probability of this even happening prior to time t. Such a liquidity event only occurs once, hence it must not have occurred before time period t. We thus have

$$\begin{aligned}
\theta_t^\gamma &= \frac{F_\gamma(t + \Delta t^*) - F_\gamma(t)}{1 - F_\gamma(t)} \\
&= \frac{\left(1 - F_\gamma(t)\right) - \left(1 - F_\gamma(t + \Delta t^*)\right)}{1 - F_\gamma(t)} \\
&= \gamma \Delta t^*.
\end{aligned} \tag{15.24}$$

With the last equality we make the assumption that this probability is proportional to the length of the time interval Δt^* in which this liquidity even could occur. Taking the limit $\Delta t^* \to 0$ for short time periods, we get

$$\frac{d\left(1 - F_\gamma(t)\right)}{1 - F_\gamma(t)} = \gamma dt,$$

which is a differential equation that can be solved for

$$F_\gamma(t) = 1 - e^{-\gamma t}. \tag{15.25}$$

If a liquidity event occurs at time t^*, this is not immediately observed, but information about this liquidity event might be revealed during a time interval $[t^*; t^* + \Delta t^*]$. In each time period there is a probability p that the information becomes available to a depositor i. Following the same steps as above, we obtain the probability that information on the liquidity even is obtained before time t_i, given the liquidity event occurred at time t^*, as

$$F_p(t_i | t^*) = 1 - e^{-p(t_i - t^*)}. \tag{15.26}$$

We are now interested in the time a depositor i infers the liquidity event occurred, given that the information was obtained at time t_i. If the information about the liquidity even has been received at time t_i, then the earliest the liquidity event could have happened is $t_i - \Delta t^*$ and the latest at t_i. Denoting $f_\gamma(t^*) = \frac{\partial F_\gamma(t^*)}{\partial t^*}$ and $f_p(t_i|t^*) = \frac{\partial F_p(t_i|t^*)}{\partial t}$ as the density functions, we then have

$$f_p(t_i|t^*) f_\gamma(t^*) = \gamma p e^{-(\gamma-p)t^*} e^{-pt_i}, \tag{15.27}$$

$$\int_{t_i-\Delta t^*}^{t_i} f_p(t_i|s) f_\gamma(s)\, ds = \gamma p e^{-pt_i} \int_{t_i-\Delta t^*}^{t_i} e^{-(\gamma-p)s} ds$$

$$= \frac{\gamma p e^{-pt_i}}{\gamma - p} e^{-(\gamma-p)t_i} \left(e^{(\gamma-p)\Delta t^*} - 1\right).$$

Using Bayesian learning, we can now determine the density of the liquidity event happening at time t^*, given the information was received at time t_i as

$$f(t^*|t_i) = \frac{f_p(t_i|t^*) f_\gamma(t^*)}{\int_{t_i-\Delta t^*}^{t_i} f_p(t_i|s) f_\gamma(s)\, ds} = \frac{\gamma - p}{e^{(\gamma-p)\Delta t^*} - 1} e^{(\gamma-p)(t_i-t^*)}, \tag{15.28}$$

where the second equality uses the results of equation (15.27). The probability that the liquidity happened before time t^*, given the information was received at t_i, is then given by

$$F(t^*|t_i) = \int_{t_i-\Delta t^*}^{t_i} f(s|t_i)\, ds = \frac{(\gamma - p)\, e^{(\gamma-p)t_i}}{e^{(\gamma-p)\Delta t^*} - 1} \int_{t_i-\Delta t^*}^{t_i} e^{-(\gamma-p)sds} = \frac{e^{(\gamma-p)\Delta t^*} - e^{(\gamma-p)(t_i-t^*)}}{e^{(\gamma-p)\Delta t^*} - 1}. \tag{15.29}$$

Using this inference, we can now determine the probability of the bank failing at a given time beyond a depositor receiving information about the liquidity event.

Probability of the bank failing Let us now propose that the bank will fail T time periods after a liquidity event as occurred, thus it will fail at $t^* + T$, where T will be determined endogenously. Hence the bank will fail if the current time, $t_0 = t^* + T$. Let us further define τ such that the current time t_0 is exactly τ time periods after information has been received, thus $t_0 = t_i + \tau$. Setting these two equal, we get that the bank fails if $t^* = t_i + \tau - T$.

With $f(t^*|t_i)$ representing the probability that a bank would fail at t_0, which is τ time periods after the information has been obtained, we get the probability that the bank fails, given it had not failed earlier, as

$$h_i(\tau) = \frac{f(t_i + \tau - T|t_i)}{1 - F(t_i + \tau - T|t_i)} = \frac{(\gamma - p)\, e^{(\gamma - p)(T-\tau)}}{e^{(\gamma - p)(T-\tau)} - 1}, \tag{15.30}$$

using equations (15.28) and (15.29) for the second equality. We observe two effects causing a bank to survive for longer. Firstly, the longer a bank has already survived after the information was received, the less likely it is to fail in the future, given its past successes; hence the hazard rate is reducing over time. On the other hand, the long time elapsed since the information was obtained may indicate that due to cumulative withdrawals of deposits, which we address below, the bank gets ever closer to failure, increasing the hazard rate over time. The latter effect dominates

$$\frac{\partial h_i(\tau)}{\partial \tau} = \frac{(\gamma - p)^2\, e^{(\gamma - p)(T-\tau)}}{\left(e^{(\gamma - p)(T-\tau)} - 1\right)^2} > 0. \tag{15.31}$$

Depositor withdrawals If banks fail we assume that they have to liquidate their assets and obtain a fraction λ of the deposits that can be distributed. Thus if a bank fails, the depositor will λD after the bank has been liquidated, but gives up their original deposits, that will have accumulated interest over time. If we denote the value of deposits τ time periods after learning the information on a liquidity event at the bank by $V(\tau)$, then the change in value experienced by the depositor is $\lambda D - V(\tau)$ if the bank fails. The value of deposits will also change as the time they have waited increases, $\frac{\partial V(\tau)}{\partial \tau}$.

We are seeking to maximize the value of our deposits and the first order condition of this maximum would be obtained if the total change in the value of the deposits is equal to zero, thus

$$h_i(\tau)(\lambda D - V(\tau)) + \frac{\partial V(\tau)}{\partial \tau} = 0. \tag{15.32}$$

If the bank fails, the depositor received λD. We know that T time period after the liquidity event, the bank will fail, hence if the depositor waits until $\tau = T$, we have $V(T) = \lambda D$. Using this condition as a boundary condition for the first order differential equation with variable coefficients in equation (15.32), we obtain after inserting for the probability of a bank failing from equation (15.30), that the value of deposits is given by

$$V(\tau) = \lambda D \tag{15.33}$$

Depositors retaining deposits obtain interest r_D on the value $V(\tau)$. On the other hand they are exposed to the bank failing with probability $h(\tau)$ and losing a fraction $1 - \lambda$ of their deposits if the bank cannot repay all deposits due to the liquidity event and the need to sell assets at a loss when seeking to repay deposits. We assume that a failing bank does not pay interest on their deposits for simplicity. Thus the profits made by depositors after waiting τ time periods is given by

$$d\Pi_D^i(\tau) = r_D V(\tau)\,dt - h_i(\tau)(1-\lambda) D dt, \tag{15.34}$$

where $V(\tau)$ and $h(\tau)$ are given by equations (15.33) and (15.30), respectively. Given that $V(\tau) - \lambda D$ and $h(\tau)$ is increasing, we see that this expression is decreasing in the waiting time.

If $d\Pi_D^i(T) > 0$, the depositor is making profits from retaining their deposits in the bank. On the other hand, if $d\Pi_D^i(0) < 0$, then depositors make losses and would be better to withdraw. Thus, as long as $d\Pi_D^i(T) > 0$, depositors will remain with the bank. This condition, after inserting for $h_i(\tau)$ from equation (15.30) becomes

$$\tau \leq \tau^* = T - \frac{1}{\gamma - p} \ln \frac{\frac{\lambda}{1-\lambda} r_D}{\frac{\lambda}{1-\lambda} r_D - (\gamma - p)}, \tag{15.35}$$

where we assumed that $0 < \gamma - p < \frac{\lambda}{1-\lambda} r_D$.

Each depositor obtains the information at a different time, and will withdraw their deposits D after τ time periods. This information will be obtained between the occurrence of the liquidity event, t^*, and τ time periods prior to the time the bank will fail, $t^* + T$, taking into account that depositors wait τ time periods before they withdraw deposits. Using equation (15.26), we know that with each deposit D we have at the time the bank fails due to running out of cash reserves

$$\begin{aligned} D \int_{t^*}^{t^*+(T-\tau)} f_p(s|t^*)\,ds &= D \int_{t^*}^{t^*+(T-\tau)} p e^{-p(s-t^*)} ds \\ &= \left(1 - e^{-p(T-\tau)}\right) D \\ &\leq R \end{aligned} \tag{15.36}$$

The final inequality defines the condition that a bank does not run out of cash reserves and avoids failure This can easily be solved for

$$\tau \geq \tau^{**} = T + \frac{\ln\left(1 - \frac{R}{D}\right)}{p}. \tag{15.37}$$

If we now combine the two conditions in equations (15.35) and (15.37), we get that depositors do not withdraw and banks do not fail if $\tau^{**} \leq \tau \leq \tau^*$, which has a viable solution if

$$\frac{R}{D} \geq 1 - \left(\frac{\frac{\lambda}{1-\lambda} r_D}{\frac{\lambda}{1-\lambda} r_D - (\gamma - p)} \right)^{\frac{p}{\gamma - p}}. \tag{15.38}$$

Thus if the remaining cash reserves after the liquidity event are sufficiently large, the bank would not fail on average. On the other hand, if the remaining cash reserves are too low, they would quickly be exhausted by depositors withdrawing and the bank would fail on average. As the information about the liquidity event reaches depositors at random times, it cannot be excluded that the bank will fail; in the case that many depositors become informed soon after the liquidity event, their withdrawals could cause the bank to fail, even if the condition in equation (15.39) is fulfilled. This possibility is also the reason that depositors continue to withdraw, despite the bank on average not failing.

However, the bank would not fail instantly after the liquidity event, but the withdrawal of deposits would be gradual and the failure of the bank delayed. The reason is on the one hand that even if depositors were to withdraw immediately after obtaining the information, $\tau = 0$, we see from equation (15.37) that the time from a liquidity event until the failure of the bank is $T = -\frac{\ln\left(1-\frac{R}{D}\right)}{p} > 0$. The reason for this observation is that information about the liquidity event reaches depositors only sequentially.

If depositors do not withdraw instantly, this time period is until the bank fails is extended. Depositors would not withdraw instantly as with a withdrawal they would forego to earn any interest on their deposits. Depositors will balance this ability to earn interest against the risk of the bank failing. The optimal time to withdraw their deposits is when waiting $\tau = \tau^*$ time periods. We have $\tau^* > 0$ and thus a positive waiting time, if

$$\lambda > \lambda^* = \frac{(\gamma - p)\, e^{(\gamma-p)T}}{(\gamma - p)\, e^{(\gamma-p)T} + r_D \left(e^{(\gamma-p)T} - 1\right)}. \tag{15.39}$$

As long as the losses to depositors imposed by a failing bank are not too high, depositors will wait to gain some interest on their deposits. Hence, deposit withdrawals are occurring over a period of time as information about a liquidity event becomes available to depositors and these withdrawals might not be happening instantly after any such information has become available to a depositor, delaying any bank failures further.

Summary We have seen that bank runs will occur over time as depositors become aware of banks facing a liquidity event and withdraw their deposits. Hence bank runs will happen slower and over time as such information arrives with depositors. A regulator learning about the liquidity event before the first depositors opens a time window to seek measures that can avoid the collapse of the bank, such as the injection of liquidity by the central bank. If the information that is spreading about a liquidity is not correct, it also gives times to counter any such unfounded rumours and re-build the trust into the bank.

Reading He & Manela (2016)

15.2.2 Deposit withdrawals after bad information

Depositors have to rely on banks obtaining repayments on the loans they provide suing their deposits. If they obtain information that reduces the repayments, for example because loans are more risky than originally expected, their assessment in the profitability of retaining deposits with the bank might change. The lower return on deposits left with the bank, could make it more attractive to withdraw deposits early and obtain the certainty of a small return on their deposits.

Assume a depositor will want to consume either in time period 1 or in time period 2; with probability p he seeks to consume in time period 1 and with probability $1-p$ in time period 2. Those depositors consuming in time period 1 will withdraw their deposits, while those depositors seeking to consume in time period 2 and not withdrawing their deposits, will share the proceeds the bank generates. These proceeds consist of the loans that have been repaid with interest r_L and we assume that the probability of loans being repaid is π_i, depending on the information the depositor has obtained.

As we also allow depositors seeking to consume in time period 2 to withdraw if they wish so, the total fraction of deposits with drawn will be $p_i = p + (1-p)\lambda_i$, where depositor i assesses that a fraction λ_i of those depositors seeking to consume in time period 2, also withdraw. With a fraction p_i of depositors seeking to withdraw their deposits D including interest r_D, the remaining deposits are $D - p_i(1+r_D)D$. This amount is now lent out by banks and they obtain a return of $\pi_i(1+r_L)$ on these loans. These proceeds are then shared between the remaining fraction of $1-p_i$ depositors. Hence the payment to the remaining depositors is

$$D_i = \frac{D - p_i(1+r_D)D}{1-p_i}\pi_i(1+r_L). \tag{15.40}$$

Let us now assume that the true probability with which loans are repaid to the bank, π, is not known by depositors. They obtain a noisy signal $\pi_i = \pi + \varepsilon_i$, where the noise term ε_i has a mean of zero and a known distribution $F(\varepsilon_i)$. The benefits of retaining the deposits with the bank are now given by the difference between the payments received in this case, $\hat{D}_i$ and the payment received if the deposit is withdrawn, $(1+r_D)D$. Taking into account that the true probability with which the loans are repaid to the bank is not known to depositors, we thus have the value of the depositor not withdrawing given by

$$\Pi_D^i = \int_{-\infty}^{+\infty} \frac{D - p_i(1+r_D)D}{1-p_i}\pi_i(1+r_L)\,dF(\varepsilon) - (1+r_D)D. \tag{15.41}$$

If $\pi_i = 0$, we see immediately that the first term in equation (15.41) is zero, making the entire expression negative and depositors would withdraw. If, on the other hand, $\pi_i = 1$, then the entire expression is positive as long as $p_i(1+r_D) < 1$, thus deposit rates are not too high. As obviously, equation (15.41) is increasing in π_i, there will exist a π^* such that $\Pi_D^i = 0$ and depositors withdraw early if $\pi_i < \pi^*$, while retaining their deposits if $\pi_i \geq \pi^*$. We will therefore see a partial bank run to the extent that a fraction of depositors receiving a low signal on the repayment of loans withdraws, even though they are only seeking to consume in time period 2.

The fraction of depositors inferred to be withdrawing early, λ_i will now be the fraction of depositors who are inferred to obtain a signal below this threshold π^*. If the true repayment rate of loans, π, decreases, the signals depositors receive will on average decrease, implying that more depositors will obtain a signal below their threshold π^* and withdraw. Thus we see that for a given threshold π^*, $\frac{\partial \lambda_i}{\partial \pi} < 0$. As we can easily show that $\frac{\partial \Pi_D^i}{\partial p_i} < 0$ and $\frac{\partial p_i}{\partial \lambda_i} > 0$, we have $\frac{\partial \Pi_D^i}{\partial \pi} = \frac{\partial \Pi_D^i}{\partial p_i}\frac{\partial p_i}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial \pi} > 0$ and hence the threshold π^* below which depositors withdraw, will increase, $\frac{\partial \pi^*}{\partial \pi} > 0$. This gives an effect that increases the extent of early withdrawals as firstly the signals received will be lower, increasing the number of depositors receiving a signal below π^*, and secondly, the threshold itself is lowered, offsetting this effect at least partially.

Let us consider the benefits of depositors remaining with the banks, Π_D^i, at the level the threshold π^*, where $\Pi_D^i = 0$. Using the implicit function theorem, we obtain that

$$\frac{\partial \lambda_i}{\partial \pi} = \frac{\partial \lambda_i}{\partial p_i} \frac{\partial p_i}{\partial \Pi_D^i} \frac{\partial \Pi_D^i}{\partial \pi} < 0. \tag{15.42}$$

The final equality arises as from above $\frac{\partial \lambda_i}{\partial p_i} > 0$, $\frac{\partial p_i}{\partial \Pi_D^i} < 0$, $\frac{\partial \Pi_D^i}{\partial \pi} > 0$. It is thus that the effect of more depositors receiving low signals on the ability of the bank to repay loans dominates the effect of a lower threshold for withdrawal and the rate of early withdrawal in increasing as the repayment of loans becomes less likely.

We have thus seen that depositors will withdraw early if they receive sufficiently bad information on the ability of the bank to repay their deposits. Based on this information some depositors will withdraw and cause a partial bank run. Furthermore, if the overall expectations on the ability of the bank to repay deposits are lowered, for example in a recession where higher default rates on loans are expected, this will increase the extent of the bank run. Depositors will seek not only obtain worse information, but will also increase their threshold below which they seek to withdraw their deposits. In order to prevent an increased withdrawals of deposits, the banks would have to re-assure depositors that the loans they have provided have not increased in risk.

Reading Goldstein & Pauzner (2005)

15.2.3 Efficient bank runs

Depositors will seek to assess the risks banks take when providing loans to evaluate the risks this imposes on their deposits. If the risks are too high to be compensated by an adequate deposit rate, depositors will withdraw and cause a bank run. Whether a bank run is justified or not, will depend on the information depositors hold, holding incomplete information might lead to bank runs when none are justified or justified bank runs might not occur.

In each of two time periods, banks have used their deposits D to finance loans on which they charge a loan rate r_L and which are repaid with probability π_t. On these deposits, banks have committed to pay a deposit rate r_D. Hence, deposits can be repaid in full after time period 1 if $\pi_1 (1 + r_L) D \geq (1 + r_D) D$, or

$$\pi_1 \geq \pi_1^* = \frac{1 + r_D}{1 + r_L}. \tag{15.43}$$

If the deposit cannot be repaid fully, this $\pi_1 < \pi_1^*$, the bank pays their depositors the funds available to them, $\pi_1 (1 + r_L) D$. Hence the fraction of deposits repaid is given by

$$\lambda_1 = \begin{cases} \pi_1 \frac{1+r_L}{1+r_D} & \text{if } \pi_1 < \pi_1^* \\ 1 & \text{if } \pi_1 \geq \pi_1^* \end{cases}. \tag{15.44}$$

If deposits are not withdrawn after time period 1 and loans extended such that both interest is accumulated, banks can repay deposits in time period 2 if $\pi_2 (1 + r_L)^2 D \geq (1 + r_D)^2 D$, or

$$\pi_2 \geq \pi_2^* = \left(\frac{1 + r_D}{1 + r_L}\right)^2 = \pi_1^{*\,2}. \tag{15.45}$$

If the loan is not repaid in time period 2, depositors obtain the funds the bank has obtained, $\pi_2 (1 + r_L)^2 D$, such that the fraction of their deposits they obtain is given by

$$\lambda_2 = \begin{cases} \pi_2 \left(\frac{1+r_L}{1+r_D}\right)^2 & \text{if } \pi_2 < \pi_2^* \\ 1 & \text{if } \pi_2 \geq \pi_2^* \end{cases}. \tag{15.46}$$

We can now investigate the decision by depositors with withdraw after time period 1, considering the cases where the risks banks have taken in time period 1 is known, thus π_1 is known, and subsequently the case where this risk is not perfectly known.

Known bank risk Let us now assume that the probability with which banks are repaid their loans in time period 2, π_2, is not known with certainty. What is known is that the difference of this probability to its long-term average, $\overline{\pi}$, persistent with a factor θ, and subject to a random fluctuation ε_2, which has a mean of zero and a distribution $F(\cdot)$. We thus have

$$\pi_2 - \overline{\pi} = \theta (\pi_1 - \overline{\pi}) + \varepsilon_2. \tag{15.47}$$

Let us now define the error term ε_2 that will lead to the probability of loan repayment, π_2, being equal to the threshold for repaying all deposits, π_2^*. This gives us $\varepsilon_2^* = \pi_2^* - \theta\pi_1 - (1 - \theta)\overline{\pi}$.

At the beginning of time period 2 we assume that depositors know the risks banks have been taking in the previous time period, £π_1. Using this information, they know the fraction of deposits they would have returned if they withdraw instantly, λ_1, but the risks the banks are taking in the future is not perfectly known and they can only form expectations about the fraction of deposits they obtain after time period 2, $\mathrm{E}[\lambda_2|\pi_1]$. We obtain after inserting equation (15.47) into equation (15.46) that

$$\begin{aligned}\mathrm{E}\left[\lambda_2|\pi_1\right] &= \int_{-\infty}^{\varepsilon_2^*} \left(\theta\pi_1 + (1-\theta)\,\overline{\pi} + \varepsilon_2\right)\left(\frac{1+r_L}{1+r_D}\right)^2 dF\left(\varepsilon_2\right) \\ &\quad + \int_{\varepsilon_2^*}^{+\infty} dF\left(\varepsilon_2\right) \\ &= 1 - F\left(\varepsilon_2^*\right) + \frac{\theta\pi_1 + (1-\theta)\,\overline{\pi}}{\pi_2^*} F\left(\varepsilon_2^*\right) \\ &\quad + \frac{1}{\pi_2^*}\int_{-\infty}^{\varepsilon_2^*} \varepsilon_2 dF\left(\varepsilon_2\right).\end{aligned} \tag{15.48}$$

From this we obtain using the Leibniz integral rule that

$$\frac{\partial \mathrm{E}\left[\lambda_2|\pi_1\right]}{\partial \pi_1} = \frac{\theta}{\pi_2^*} F\left(\varepsilon_2^*\right). \tag{15.49}$$

If depositors withdraw after time period 1, they will obtain a payment of $\lambda_1\left(1+r_D\right)D$ and if they retain their deposits, they expect to obtain after time period 2 the amount of $\mathrm{E}\left[\lambda_2|\pi_1\right]\left(1+r_D\right)^2 D$. Depositors will withdraw if $\lambda_1\left(1+r_D\right)D \geq \mathrm{E}\left[\lambda_2|\pi_1\right]\left(1+r_D\right)^2 D$, which solves for

$$1 + r_D \leq 1 + r_D^* = \frac{\lambda_1}{\mathrm{E}\left[\lambda_2|\pi_1\right]}. \tag{15.50}$$

Thus we see that if deposit rates are not too high, deposits will be withdrawn, which we interpret as a bank run. We note here that the threshold deposit rate is only implicitly defined as λ_1 and λ_2 itself are dependent on the deposit rate $1 + r_D$.

As we assumed that depositors know the risks banks have taken, π_1, they can apply equation (15.44) to obtain the fraction of deposits repaid, λ_1. If $\pi_1 \geq \pi_1^*$, we know that $\lambda_1 = 1$ and using equation (15.49), we see that the critical deposit rate for a withdrawal is decreasing in the repayment rate of the bank as we have

$$\frac{\partial\left(1+r_D^*\right)}{\partial \pi_1} = -\frac{1}{\mathrm{E}\left[\lambda_2|\pi_1\right]^2}\frac{\theta}{\pi_2^*} F\left(\varepsilon_2^*\right) < 0. \tag{15.51}$$

In the case that $\pi_1 < \pi_1^*$, this becomes

$$\begin{aligned}\frac{\partial\left(1+r_D^*\right)}{\partial \pi_1} &= \frac{1}{\pi_1^*}\frac{1 - F\left(\varepsilon_2^*\right) + (1-\theta)\,\frac{\overline{\pi}}{\pi_2^*} F\left(\varepsilon_2^*\right) + \frac{1}{\pi_2^*}\int_{-\infty}^{\varepsilon_2^*} \varepsilon_2 dF\left(\varepsilon_2\right)}{\mathrm{E}\left[\lambda_2|\pi_1\right]^2} \\ &> 0.\end{aligned} \tag{15.52}$$

We thus see that using that the further the repayment rate of loans is away from the critical threshold for a full repayment of deposits, π_1^*, the lower the deposit rate can be without triggering a bank run. It is that if the current repayment rate is high, it is expected to remain high and thus the risk to depositors is reduced. On the other hand, a low current repayment rate and thus a high risk to depositors will impose losses on depositors withdrawing, making a bank run less attractive and with the repayment rate expected to revert back towards its long-term average, remaining with the bank becomes more attractive.

Depositors balance the future returns they can obtain from retaining their deposits against the risks of the deposits not being repaid in full. As this risk reverts slowly towards its long-term average, it may well worth to not withdraw deposits at a loss and consider the likelihood that risks will reduce in the future, reducing any such losses. With banks taking low risks, such an increase in the risk as it reverts to its long-term average might not be of substantial concern as the risk is unlikely to increase such that losses are incurred.

Unknown bank risk Banks usually do not disclose the risks they taking to depositors, hence banks will have no information regarding the value of π_1. However, depositors know the long term average to be $\overline{\pi}$, hence we assume that their believe is $\pi_1 = \overline{\pi} + \varepsilon_1$, where ε_1 is a random term with a mean of zero. If we define now ε_1^* as the random term for which the probability of loan repayments is at its threshold for being able to repay deposits fully, π_1^*, we have $\varepsilon_1^* = \pi_1^* - \overline{\pi}$. Assuming that ε_1 and ε_2 have the same distribution $F(\cdot)$, The expected fraction of deposits repaid is then given as

$$\begin{aligned} \mathrm{E}[\lambda_1] &= \int_{-\infty}^{\varepsilon_1^*} \pi_1 \frac{1+r_L}{1+r_D} dF(\varepsilon_1) + \int_{\varepsilon_1^*}^{+\infty} dF(\varepsilon_1) \\ &= 1 - F(\varepsilon_1^*) + \frac{\overline{\pi}}{\pi_1^*} F(\varepsilon_1^*) + \frac{1}{\pi_1^*} \int_{-\infty}^{\varepsilon_1^*} \varepsilon_1 dF(\varepsilon_1), \end{aligned} \tag{15.53}$$

where we made use the definition of π_1^* in equation (15.43). Using from the definition of ε_1^* that $\frac{\partial \varepsilon_1^*}{\partial \overline{\pi}} = -1$, we get using the Leibniz integration rule that

$$\frac{\partial \mathrm{E}[\lambda_1]}{\partial \overline{\pi}} = \frac{F(\varepsilon_1^*)}{\pi_1^*}. \tag{15.54}$$

As depositors do now know the value of π_1, their expectation of the fraction of deposits repaid in time period 2, will be $\mathrm{E}[\lambda_2] = \mathrm{E}[\mathrm{E}[\lambda_2|\pi_1]]$, where $\mathrm{E}[\lambda_2|\pi_1]$ is given from equation (15.48). It is thus

$$\begin{aligned} \mathrm{E}[\lambda_2] &= \int_{-\infty}^{+\infty} \mathrm{E}[\lambda_2|\pi_1]\, dF(\varepsilon_1) \\ &= 1 - F(\varepsilon_2^*) + \frac{\overline{\pi}}{\pi_2^*} F(\varepsilon_2^*) + \frac{1}{\pi_2^*} \int_{-\infty}^{\varepsilon_2^*} \varepsilon_2 dF(\varepsilon_2), \end{aligned} \tag{15.55}$$

having used that when inserting for $\pi_1 = \overline{\pi} + \varepsilon_1$ and taking expectations we have $\varepsilon_2^* = \pi_2^* - \overline{\pi}$, as well as $\mathrm{E}[\varepsilon_1] = 0$. From this expression we easily obtain that

$$\frac{\partial \mathrm{E}\left[\lambda_2\right]}{\partial \overline{\pi}} = \frac{F\left(\varepsilon_2^*\right)}{\pi_2^*}. \tag{15.56}$$

Depositors will withdraw if their expected payments from withdrawing exceed the expected payments from remaining with the bank, $\mathrm{E}\left[\lambda_1\right]\left(1+r_D\right)D \geq \mathrm{E}\left[\lambda_2\right]\left(1+r_D\right)^2 D$, or

$$1+r_D \leq 1+r_D^{**} = \frac{\mathrm{E}\left[\lambda_1\right]}{\mathrm{E}\left[\lambda_2\right]}. \tag{15.57}$$

We note here that the threshold deposit rate is only implicitly defined as λ_1 and λ_2 itself are dependent on the deposit rate $1 + r_D$. Using the expected values for the expected fraction of deposits repaid from equations (15.53) and (15.56), as well as nothing that $\pi_2^* = \pi_1^{*\,2}$, we obtain

$$\begin{aligned}\frac{\partial\left(1+r_D^{**}\right)}{\partial\overline{\pi}} &= \frac{\pi_1^* F\left(\varepsilon_1^*\right) - F\left(\varepsilon_2^*\right)}{\pi_2^*} + \frac{\pi_1^*-1}{\pi_2^*}F\left(\varepsilon_1^*\right)F\left(\varepsilon_2^*\right) \\ &\quad + \frac{F\left(\varepsilon_1^*\right)\int_{-\infty}^{\varepsilon_2^*}\varepsilon_2 dF\left(\varepsilon_2\right) - F\left(\varepsilon_2^*\right)\int_{-\infty}^{\varepsilon_1^*}\varepsilon_1 dF\left(\varepsilon_1\right)}{\pi_1^*\pi_2^*} \\ &< 0.\end{aligned} \tag{15.58}$$

We now note that $\varepsilon_2^* < \varepsilon_1^*$ and hence the numerator final term, rewritten as $\int_{-\infty}^{\varepsilon_2^*}\int_{-\infty}^{\varepsilon_1^*}\left(\varepsilon_2-\varepsilon_1\right)dF\left(\varepsilon_1\right)df\left(\varepsilon_2\right)$ is negative, given that the values for ε_1 can become larger. The first term is also negative as multiplying all values up to ε_2^* with $\pi_1^* < 1$ makes the expression of the first integral smaller; this is not compensated for by having a higher upper integration bound as these higher bounds multiplied by $\pi_1^* < 1$ would still be smaller. The second term is obviously negative and hence, the entire expression is negative.

We therefore see that the higher the expected repayment rate of the loans is, thus the lower the bank risk, the lower the deposit rate can be to avoid a bank run. A higher repayment rate of loans, reduces the risk to depositors in time period 1 and time period 2, given that no information on the actual repayment rate is available, and depositors know that there is some persistence in these repayment rates; this makes the withdrawal of deposits less attractive to depositors as the repayment rate increases. Depositors have no information on the actual risks and can therefore not consider changes towards its long-term average, they only can consider the long-term average risk.

We can now compare the results if depositors know the risks bank take with the case of such information not being available. Figure 15.2 depicts the areas in which bank runs occur in both cases. We label a bank run as efficient if it is profitable to withdraw deposits from the bank due to the risk the bank is taking and if this assessment is based on information about the actual risk that banks have taken, thus it will be area below $1 + r_D^*$. If this information is not available and the depositor can only infer the average risk the bank chooses, deposit withdrawals and hence a bank run occurs if the deposit rate is below $1 + r_D^{**}$. We see from the figure that for low

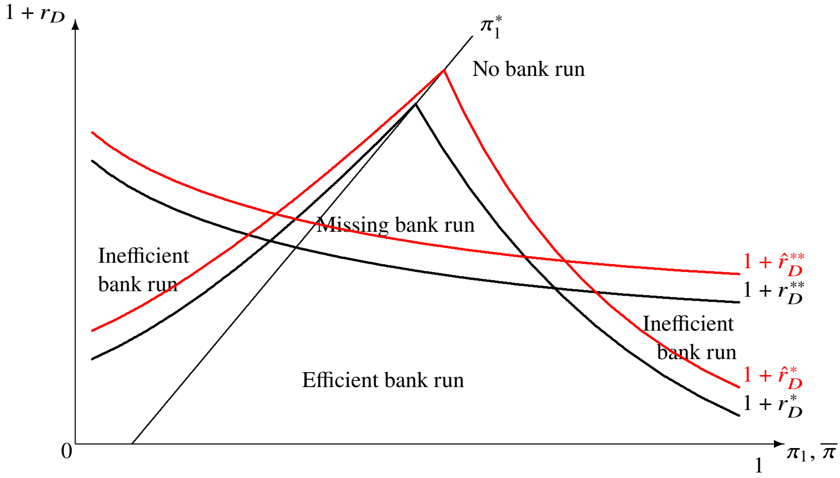


Fig. 15.2: Efficient and inefficient bank runs

and high risks, this threshold in the deposit rate to cause a bank run is lower if no precise information on the risk of the bank is available. This will lead to bank runs that would not occur with information on the actual risks the bank takes. The reason is that with very low risks, depositors do not foresee any losses from withdrawing deposits, while taking into account the possibility of future losses if they remain with the bank. This gives them an incentive to withdraw deposits early. On the other hand if the bank risks are high, depositors do not a good prospect of recovering any losses they may make from withdrawing early by retaining their deposits with the bank, making en early withdrawal profitable. If they had information on the actual risk taken, the tendency of the risk to revert to the long-term mean would have induced an incentive to retain the deposits with the bank for high risks as the risk is likely to reduce and the losses they are currently facing are more likely to reduce. If the risks are low, although the risks are likely to increase, they are unlikely to do so significantly enough to cause losses to depositors, giving an incentive to retain deposits.

For intermediate risks, those around the threshold at which depositors face losses in time period 1, bank runs might not occur often enough as the threshold for the deposit rate is lower if depositors have no information on the risk the bank takes. Being close the threshold for deposits not being repaid fully, depositors having this information will anticipate more losses in the future and would therefore prefer to either withdraw making only small losses, or withdraw to prevent future losses. If depositors do not have this information, they cannot act in such a situation and will therefore not withdraw deposits, avoiding a bank run to occur.

An increase in the deposit rate will also increase the threshold for depositors to be repaid in full, π_i^*, causing the range in which bank runs occur to increase, as the red line in figure 15.2 shows, where a higher deposit rate is applied, changing the thresholds to $1 + \hat{r}_D^*$ and $1 + \hat{r}_D^{**}$, respectively. The higher deposit rates increases the payments depositors are due and hence with the same risk and loan rates, the repayment rates would reduce. Thus increasing the deposit rate might not alleviate the problem of a bank run, especially not in the intermediate range of risks around the threshold of π_1^* if depositors hold information in the bank's risk.

Summary If depositors do not have sufficient information on the risks banks take, the bank runs that occur, are not always efficient. For high and low bank risks, bank runs are occurring too often, while for intermediate risks, the risk of bank runs is too low compared to the situation in which depositors have full information on the risks the bank is currently taking. These inefficiencies in having too many or too few bank runs are arising because of the incomplete information depositors have.

It looks as if bank runs can be prevented through raising deposit rates, making it more attractive to depositors to remain with the bank, but this will not only affect the profitability of banks, but also the threshold at which they are able to repay depositors in full. Hence raising deposit rates to avert a bank run might be a viable option in some instances where risks are either particularly high or low, but for intermediate ranges in particular, this would not be feasible. Inefficient bank runs could be prevented by providing depositors with reliable information on the risks that banks are taking, but this might on the other hand increase the risk of bank runs for banks taking intermediate risks, those that are close to the threshold of not being able to meet the withdrawal demand of depositors.

Reading Gorton (1985)

15.2.4 The effect of loan guarantees

One reason for bank runs to occur is that depositors are concerned about the ability of banks to repay their deposits and withdraw these. Such a concern could be based on possible defaults of loans the bank has provided; with loan guarantees given to the bank, for example by government agencies or commercial providers, the potential losses of banks from loan defaults are reduced as at least some of these losses will be compensated. Thus, depositors should receive a higher repayment of their deposits, reducing the threat of bank runs. However, having been provided with loan guarantees also affects the behaviour of banks when granting loans and these may counter the risk reduction to depositors.

Let us assume that banks provide loans L with interest r_L to companies, who repay these loans with probability π_i. The probability of success is π_H with probability p, but with probability $1 - p$ a shock occurs to the company and the probability of success is reduced to $\pi_L = \alpha\pi_H$, where $\alpha < 1$. Whether this shock occurs is unknown to the bank at the time it provides the loan but will be revealed later to the bank as well as depositors. Upon learning the probability of the loan being repaid, depositors decide whether to withdraw and cause a bank run.

Banks can monitor those companies they provide loans to facing costs C. Through monitoring, the bank can influence the probability with which loans are repaid, where we assume that these costs are given by $C = \frac{1}{2}c\pi_H^2 L$ and hence increasing in the probability of success. These costs are incurred in the anticipation that companies do not face a shock that reduces their probability of repaying the loan. In addition, banks have been provided with a loan guarantee that pays banks a fraction λ of the outstanding loan if the company is not able to repay. Assuming that loans are fully financed by deposits commanding interest r_D, the profits of the bank is given by

$$\begin{aligned}\Pi_B = {} & p\left(\pi_H + (1-\pi_H)\lambda\right)(1+r_L)L \\ & + (1-p)\left(\alpha\pi_H + (1-\alpha\pi_H)\lambda\right)(1+r_L)L \\ & - (1+r_D)L - C.\end{aligned} \tag{15.59}$$

Banks will maximise their profits by choosing their monitoring optimally such that $\frac{\partial \Pi_B}{\partial \pi_H} = 0$, which after inserting for C easily solves for

$$\pi_H^* = \frac{p + (1-p)\alpha}{c}(1+r_L)(1-\lambda). \tag{15.60}$$

If the loan guarantee pays a larger fraction λ of the loan in the case the company defaults, the efforts of the bank in monitoring the company are also reduced; this reduces the likelihood of the loan being repaid. The payment from the loan guarantee induces a moral hazard in that banks will reduce their monitoring effort and instead rely on the loan guarantee to improve their profits rather than their own efforts. With reduced monitoring, the monitoring costs are also reduced, benefitting the bank in addition to the payments from the loan guarantee, and we easily get these costs as $C = \frac{1}{2}\frac{(p+(1-p)\alpha)^2}{c}(1+r_L)^2(1-\lambda)^2 L$ from inserting for π_H from (15.59).

After learning whether a shock has occurred to the company, depositors have to decide whether they want to withdraw their deposits. Depositors would withdraw their deposits if their claim of $(1+r_D)L$ cannot be met by the proceeds the bank obtains, less their monitoring costs. If we assume that no shock occurs to the company and the probability of success is π_H, this condition, equivalent to $\Pi_B \geq 0$, becomes $(\pi_H + (1-\pi_H)\lambda)(1+r_L)L - C \geq (1+r_D)L$, which, when inserting for π_H from equation (15.60) and for C from the expression above, solves for

$$(p + (1-p)\alpha)\left(1 - \frac{1}{2}(p + (1-p)\alpha)\right) \geq c\frac{(1+r_D) - \lambda(1+r_L)}{(1+r_L)^2(1-\lambda)^2}. \tag{15.61}$$

Similarly, if the company faces a shock and the probability of success reduces to $\pi_L = \alpha\pi_H$, we easily get by replacing π_H with π_L in the condition for depositors to not withdraw that

$$(p + (1-p)\alpha)\left(\alpha - \frac{1}{2}(p + (1-p)\alpha)\right) \geq c\frac{(1+r_D) - \lambda(1+r_L)}{(1+r_L)^2(1-\lambda)^2}. \tag{15.62}$$

It is obvious that the left-hand side in this case is smaller as $\alpha < 1$ and thus this condition is more restrictive. Not surprisingly, if the company's ability to repay the loan is adversely affected by the shock, the condition for a bank run is more easily fulfilled.

If the right-hand side of equations (15.61) and (15.62) is decreasing, the constraint on avoiding a bank run becomes less binding and hence bank runs are less likely observed. We easily get that the first derivative with respect to the loan guarantee λ of this expression is negative if

$$\lambda > \lambda^* = 2\frac{1 + r_D}{1 + r_L} - 1. \tag{15.63}$$

Thus, by increasing the extent of the loan guarantee, λ, bank runs are only becoming less likely if the loan guarantee is already sufficiently high. Using realistic values for the loan and deposit rates, this threshold for reducing bank runs will require a high level of protection from the loan guarantee. Thus for loan guarantees that do not cover most of the loan, increasing this loan guarantee would actually increase the likelihood of a bank run as the restriction in equations (15.61) and (15.62) become more binding. The reason for this observation is that banks will reduce their monitoring efforts, which will increase the risks companies take, and this effect is stronger than the increased payments from the loan guarantee and the reduced monitoring costs.

We can also compare the right-hand side of the constraint in equations (15.61) and (15.62) with a situation in which no loan guarantee is provided, $\lambda = 0$, we then easily see that this expression is smaller in the presence of a loan guarantee only if

$$\lambda > \lambda^{**} = 2 - \frac{1 + r_L}{1 + r_D}. \tag{15.64}$$

Again, for realistic loan and deposit rates, this would again require a loan guarantee that covers a substantial amount of the loan for the same reasons as outlined above.

We have thus seen that loan guarantees, while providing additional payments to banks, and indirectly depositors, in many instances will make bank runs more likely. This result is driven by the incentives of the bank to reduce their monitoring efforts as the result of the loan guarantee, increasing the risks companies are taking. This increased risk of companies will have to be balanced against the increased loan repayments due to the loan guarantee and the reduced monitoring costs of banks. Unless loan guarantees are extensive, the effect from the increased risks companies take, will dominate.

Providing, or increasing, loan guarantees as a measure to instill confidence into the banking system and reduce the threat of bank runs may in fact be counterproductive. Unless the loan guarantee introduced is very high, the moral hazard of banks providing loans to companies with higher risks will dominate any positive effects depositors will obtain from the loan guarantee. Thus, using loan guarantees to prevent bank runs is not an effective policy measure.

Reading Carletti, Leonello, & Marquez (2023)

Résumé

Depositors are exposed to the risks a bank takes when providing loans as the proceeds from these loans are used to repay depositors. Negative information about a bank's ability to repay depositors provides incentives to withdraw deposits as long as any losses have either not materialised or are expected to increase in the future. Such information about the state of a bank may affect the quality of their assets, mainly the risks of loans, but also any adverse effect on their liquidity position. With deposits constantly withdrawn for consumption, banks hold a certain amount of cash reserves to meet such demands. If unexpected outflows of deposits, or the absence of expected inflows of deposits, cause cash reserves to reduce, this might put the ability of the bank to repay depositors into question. If they have to raise additional cash reserves at a loss, they might not be able to repay depositors in the future. Hence, depositors would withdraw early in order to avoid any such losses, causing a bank run. They might not withdraw their deposits instantly, but might wait to obtain additional interest, while limiting their exposure to risk.

Information can also be contagious in that having obtained negative information, even if this in itself does not justify the withdrawal of deposits, is likely to be received by many depositors. Expecting other depositors to have obtained similar information, it might be beneficial to withdraw deposits. This benefit might only be arising due to expecting other depositors to obtain similar information and act in the same way. Often banks react to the outflow of deposits, a slowly emerging bank run, with increasing deposit rates, but such a measure may well be unsuccessful. The higher deposit rate will make it more attractive to retain deposits with the bank, but on the other hand, the bank also has to be able to pay such higher deposit rates. If a bank faces a bank run as there are concerns about their ability to repay depositors, promising higher future payments will often not alleviate the concerns if depositors believe that these payments cannot be made.

We also observe that loan guarantees provided by governments and covering a some of the losses banks are making from defaults does not necessarily reduce banks runs. The loan guarantee induces moral hazard in the bank's incentives to exert effort. With banks reducing their effort, risks will increase and the possible losses to depositors might be higher, despite the loan guarantee provided, making bank runs more likely. This will be the case where loan guarantees are not sufficiently large to compensate for this moral hazard.

15.3 Deposit rates in the presence of bank runs

Deposit rates commonly reflect the risks depositors are exposed to in the form of the bank not being able to repay deposits. This is commonly assumed to be the consequence of bank loans not being repaid, but it will also take into account any deposit insurance schemes that lower such risk. However, losses do not only emerge from loans not being repaid, but also if deposits are withdrawn early and leaving the bank with less funds to provide loans, which will then not allow them to generate sufficient funds to repay depositors fully if any interest is to be paid. This risk of a bank run needs to be taken into account when determining deposit rates.

Depositors provide banks with a deposit D, which they may withdraw at any time. Depositors are unaware of when they might need their deposits returned, but know this will happen with probability γ after one time period and with probability $1-\gamma$ after two time periods. Banks use these deposits by providing loans L over two time periods, on which they earn interest r_L and which are repaid with probability π. Not all of the deposits banks obtain are invested into loans, however, but an amount of R is held as a cash reserve, such that $L = D - R$. For depositors withdrawing after one time period the bank will pay interest r_D^1 and for those remaining with the bank until time period 2, the bank pays a deposit rate of r_D^2, covering both time periods.

We will now explore the deposit rates if deposits can only be withdrawn by those requiring their return, for example to finance consumption. Afterwards we will explore the case where depositors not requiring their deposits will withdraw these, which we refer to as early withdrawals or a bank run.

Deposits without early withdrawals Depositors do not know when they want to withdraw their deposits in advance, so their utility will be determined by the utility of withdrawing deposits, obtaining a repayment of $\left(1+r_D^1\right)D$, and the utility of remaining with the bank, obtaining a repayment of $\left(1+r_D^2\right)D$. This gives us an expected utility of

$$\Pi_D = \gamma u\left(\left(1+r_D^1\right)D\right) + (1-\gamma)\,u\left(\left(1+r_D^2\right)D\right). \tag{15.65}$$

Banks will ensure they hold sufficient cash reserves to pay those depositors withdrawing after the first time period, hence we will require that the expected repayments, $\gamma\left(1+r_D^1\right)$ do not exceed the reserves R, thus

$$\gamma\left(1+r_D^1\right)D \leq R. \tag{15.66}$$

In time period 2, banks need resources $(1-\gamma)\left(1+r_D^2\right)D$ to repay the remaining depositors. These resources are drawn from the cash reserves that have not been used to repay depositors in time period 1, $R-\gamma\left(1+r_D^1\right)D$, on which interest has been paid, and the repaid loans they have provided, $\pi\left(1+r_L\right)L$. Hence we require furthermore that

$$(1-\gamma)\left(1+r_D^2\right)D \leq R-\gamma\left(1+r_D^1\right)D+\pi\left(1+r_L\right)L. \tag{15.67}$$

Denoting by ξ_1, and ξ_2 the Lagrange coefficients for the constraints in equations (15.66) and (15.67), respectively, the first order conditions for maximizing the expected utility of depositors, equation (15.65), are given by

$$\frac{\partial \mathcal{L}}{\partial\left(1+r_D^1\right) D}=\gamma \frac{\partial u\left(\left(1+r_D^1\right) D\right)}{\partial\left(1+r_D^1\right) D}-\xi_1 \gamma-\xi_2 \gamma=0, \tag{15.68}$$

$$\frac{\partial \mathcal{L}}{\partial\left(1+r_D^2\right) D}=(1-\gamma) \frac{\partial u\left(\left(1+r_D^2\right) D\right)}{\partial\left(1+r_D^2\right) D}-\xi_2(1-\gamma)=0,$$

from which the second condition gives us $\xi_2=\frac{\partial u\left(\left(1+r_D^2\right) D\right)}{\partial\left(1+r_D^2\right) D}>0$ and inserting this expression into the first condition we obtain

$$\frac{\partial u\left(\left(1+r_D^1\right) D\right)}{\partial\left(1+r_D^1\right) D}-\xi_1=\frac{\partial\left(\left(1+r_D^2\right) D\right)}{\partial\left(1+r_D^2\right) D}. \tag{15.69}$$

As $\xi_2>0$, the constraint in equation (15.67) is binding. If $\xi_1>0$, then the condition in equation (15.66) is also binding, implying from solving these conditions as equalities that

$$\left(1+r_D^1\right) D=\frac{R}{\gamma}, \tag{15.70}$$

$$\left(1+r_D^2\right) D=\frac{\pi\left(1+r_L\right) L}{1-\gamma},$$

after inserting the solution for $\left(1+r_D^1\right) D$ into equation (15.67) to obtain the second expression.

If $\xi_1=0$, then equation (15.66) is not binding and we have $\left(1+r_D^1\right) D<\frac{R}{\gamma}$. In this case all deposit withdrawals can be met by the bank and it has additional cash reserves left to use when repaying those depositors that remain with the bank. Let us propose that in this case banks provide depositors remaining with them only with the minimal deposit rate at which they will not seek to withdraw their deposits and set $\left(1+r_D^2\right) D=\left(1+r_D^1\right) D$.

Inserting this relationship into the binding constraint from equation (15.67), we get the expression for $\left(1+r_D^1\right) D$ as

$$\left(1+r_D^1\right) D=\left(1+r_D^2\right) D=R+\pi\left(1+r_L\right) L. \tag{15.71}$$

The constraint on the cash reserves held by the bank, $\gamma\left(1+r_D^1\right) D \leq R$, becomes binding if

$$\pi>\pi^*=\frac{1-\gamma}{\gamma} \frac{R}{\left(1+r_L\right) L}. \tag{15.72}$$

If the constraint is binding, we have both constraints, equations (15.66) and (15.67), being fulfilled with equality and the repayments to depositors are given by equation (15.70), and if it is not binding the deposit rates are given by equation (15.71).

As we can see from the depiction of our result in figure 15.3, we have recovered the standard deposit contract for depositors withdrawing after one time period. If the repayments from loans are sufficiently high such that banks face no constraints on their cash reserves that would limit the ability to repay deposits, $\pi > \pi^*$, depositors obtain a fixed repayment and for lower loan repayments, depositors obtain their share of these repayments and the loans that have been repaid. For those depositors who have not withdrawn, they will obtain the same deposit rate if the bank faces constraints on their ability to repay depositors, treating them equally. If there are no constraints on repaying withdrawn deposits, those remaining with the bank can extract any surplus from the bank due to banks competing for deposits and obtain the loan repayments the bank receives.

Having established the deposit rates without allowing depositors to withdraw early, we can now allow for such early withdrawals by depositors and consider deposit rates in the presence of a bank run.

Deposits allowing early withdrawals Let us now assume that in addition to the fraction γ of depositors requiring their deposits returned, from those depositors that do not need to withdraw, a fraction $\hat{\gamma}$ also withdraws. Such early withdrawals can be interpreted as a partial bank run as the deposits of the bank get depleted more than would be necessary. In time period 1 we then observe that a fraction $\gamma + (1-\gamma)\hat{\gamma}$ of deposits are withdrawn. In this case, the constraints on the cash reserves from equation (15.66) and the total resources available to repay depositors in time period 2, equation(15.67), become

$$\begin{aligned}(\gamma + (1-\gamma)\hat{\gamma})\left(1+\hat{r}_D^1\right)D &\leq R \\ (1-\gamma)(1-\hat{\gamma})\left(1+\hat{r}_D^2\right)D &\leq \left(R - \left(1+\hat{r}_D^1\right)D\right) \\ &\quad +\pi\left(1+r_L\right)L,\end{aligned} \tag{15.73}$$

where $\hat{r}_D^1$ and $\hat{r}_D^2$ denote the deposit rates for those withdrawing in time period 1 and time period 2, respectively.

With depositors maximizing their utility, we know from equation (15.68) that the second constraint is binding. Let us furthermore assume that banks offer a deposit contract that promises to pay depositors $\left(1+\hat{r}_D^*\right)D$ if they withdraw early, provided the bank has sufficient cash reserves available, and they use all their reserves to repay deposits otherwise, in which case the first constraint becomes binding, too. We thus have $\left(1+\hat{r}_D^1\right)D = \min\left\{\left(1+\hat{r}_D^*\right)D; \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}\right\}$, where the second expression has been obtained from the first constraint in equation (15.73) being fulfilled with equality due to it being binding. Solving the binding second constraint in equation ((15.73)), the repayment for depositors remaining with the bank are given by

$$\left(1+\hat{r}_D^2\right)D = \frac{\left(R - (\gamma+(1-\gamma)\hat{\gamma})\left(1+\hat{r}_D^*\right)D\right) + \pi\left(1+r_L\right)L}{(1-\gamma)(1-\hat{\gamma})} \tag{15.74}$$

if $\left(1+\hat{r}_D^*\right)D \leq \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}$ and the bank has sufficient cash reserves to meet any withdrawals at the promised deposit rate of $\hat{r}_D^*$. Solving the condition for cash reserves being sufficient to meet the demand of withdrawing depositors for the early

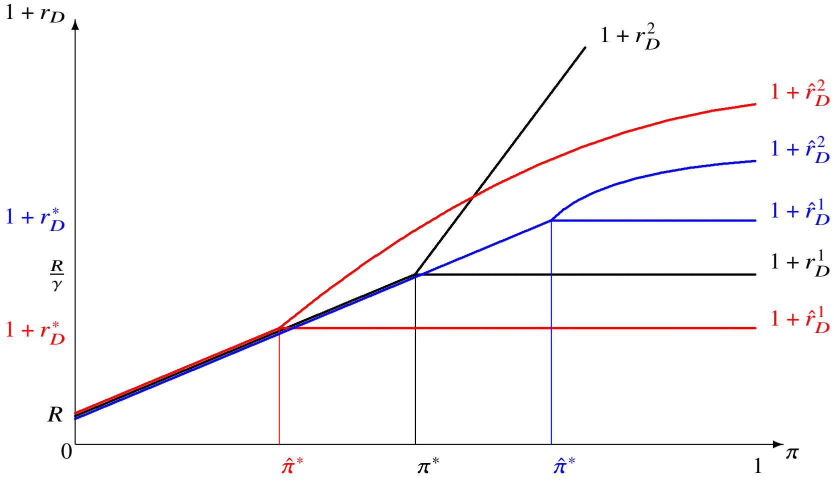


Fig. 15.3: Deposit rates with early withdrawals

withdrawal rate $\hat{\gamma}$, we obtain that this requires $\hat{\gamma} < \hat{\gamma}^* = \frac{R}{(1-\gamma)\left(1+\hat{r}_D^*\right)D} - \frac{\gamma}{1-\gamma}$. We thus observe that withdrawal rates of $0 < \hat{\gamma} < \hat{\gamma}^*$ are not feasible as the repayment, when remaining with the bank, in this range is higher than when withdrawing deposits.

If the cash reserves are not sufficient to repay all depositors withdrawing at this rate, $\left(1+\hat{r}_D^*\right) D > \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}$, the deposits rate for those remaining with the bank is given by

$$\left(1+\hat{r}_D^2\right) D = \frac{\pi\left(1+r_L\right) L}{(1-\gamma)\left(1-\hat{\gamma}\right)} \tag{15.75}$$

If we have a shortage of cash reserves, $\left(1+\hat{r}_D^*\right) D \geq \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}$, then we require that $\left(1+r_D^2\right) D = \left(1+r_D^1\right) D$. If the repayments in time period 1 were higher than in time period 2, all deposits would be withdrawn early and if the repayment in time period 2 was exceeding that of time period 1, no deposits would be withdrawn early. Hence the only equilibrium, is to offer the same repayments in both time periods. Thus solving from the first constraint in equation (15.73) and using the expression in equation (15.74), we obtain the equilibrium early withdrawal rate as

$$\hat{\gamma} = \frac{(1-\gamma) R - \gamma\pi\left(1+r_L\right) L}{(1-\gamma)\left(R+\pi\left(1+r_L\right) L\right)}. \tag{15.76}$$

Of course, in order to observe a shortage of cash reserves we need $\hat{\gamma} > \hat{\gamma}^*$, which was determined above. Solving this expression, we obtain the requirement that

$$\pi < \hat{\pi}^* = \frac{(1-\gamma)(1-\hat{\gamma}^*)}{1-(1-\gamma)(1-\hat{\gamma}^*)}\frac{R}{(1+r_L)L} \tag{15.77}$$
$$= \frac{\left(1+\hat{r}_D^*\right)D - R}{(1+r_L)L}.$$

Hence if $\pi < \hat{\pi}^*$, we observe that the cash reserves are not sufficient to repay all depositors withdrawing in time period 1. If we insert for $\hat{\gamma}^*$ into equations (15.74) and (15.74), we find the deposit rates of those remaining with the bank is given by

$$\left(1+\hat{r}_D^2\right)D = \begin{cases} R + \pi(1+r_L)L + \frac{\pi(1+r_L)L - \left(\left(1+\hat{r}_D^*\right)D - R\right)}{\pi(1+r_L)L}R & \text{if } \pi > \hat{\pi}^* \\ R + \pi(1+r_L)L & \text{if } \pi < \hat{\pi}^* \end{cases}. \tag{15.78}$$

If we set the deposit rate given to depositors with drawing in time period 1 such that $\left(1+\hat{r}_D^*\right)D < \frac{R}{\gamma}$, and hence the repayments when allowing for early withdrawals are below those in the case that no withdrawals are possible, we see that the threshold at which cash reserves are insufficient to meet the demand of deposit withdrawals is lower, $\hat{\pi}^* < \pi^*$. The lower deposit rate provides less incentives to retain deposits with the bank and hence we will observe early withdrawals.

We can now further compare the deposit rates if early withdrawals are allowed and figure 15.3 illustrates our result. We see that for low repayment rates of loans, π, and hence banks facing a cash reserve shortage the loan rates are identical in both times periods due to our assumption that banks do not pay more to depositors remaining with the bank than is necessary to prevent them withdrawing. The higher the loan repayments to banks are, the more resources they have available to repay depositors, increasing the deposit rate and once the promised deposit rate in the case of possible early withdrawals are reached, the shortage of cash reserves ceases. The same happens without the possibility of early withdrawal once the repayments to depositors withdrawn are fully made. Once cash reserves are sufficient to repay all deposits that are withdrawn, the deposit rates for the two time periods diverge. Competition between banks, allows depositors to extract any surplus from banks and the deposit rates will increase as the repayments from loans increase.

If banks promise a higher deposit rate to those withdrawing, $\left(1+\hat{r}_D^*\right)D > \frac{R}{\gamma}$, then the deposit rate for those remaining with the bank will be lower. The higher deposit rate when withdrawing increases the incentives to withdraw and hence larger cash reserves need to be held, limiting the amount that can be lent and hence the revenue the bank receives from loans and distributes to depositors remaining with them is reduced, resulting in a lower deposit rate for these depositors.

On the other hand, if banks promise a lower deposit rate to those withdrawing, $\left(1+\hat{r}_D^*\right)D < \frac{R}{\gamma}$, The incentives to withdraw early are small and the bank can hold lower cash reserves, allowing for larger investments into loans, which generates higher revenue that can be distributed amongst the remaining depositors. As this revenue from loans increases due to higher repayments, less and less depositors withdraw early and the revenue needs to be split with more remaining depositors, reducing the return to each of them and the deposit rate will be lower than without early withdrawals.

Summary If deposits can be withdrawn early, and thus a bank run can occur, this will affect deposit rates as depositors will take into account the possibility of losses arising from the depletion of cash reserves from such early withdrawals. Deposit rates will also to be set such that they do not provide an incentive to withdraw early, limiting the level of deposit rates that can be offered. The structure of the deposit contract itself is not significantly affected by the possibility of bank runs, we retain the fixed repayment of deposits, for as long as the resources of the bank permit this, and those remaining with the bank can then extract any surplus in future time periods. If high deposit rates are promised to depositors withdrawing, this will affect those remaining as they will obtain lower future deposit rates, given that bank need to maintain higher cash reserves to accommodate the higher withdrawal rate.

Thus banks will structure their deposit rates such that no incentives for a large withdrawals of deposits exist. They will do this by making an early withdrawal of deposit less attractive and offer sufficiently high deposit rates for those remaining with the bank. While this might be difficult to achieve with deposits constantly flowing in and out through the deposit rate directly, banks may use higher deposit rates for larger deposits, ensuring they are not withdrawn below a certain threshold, or providing bonuses for depositors who stay with them for long periods of time.

Reading Allen & Gale (1998)

15.4 The impact of competition

Banks may face the withdrawal of deposits and will have to use their cash reserves as well as the proceeds from liquidating assets to meet this demand. The degree of competition between banks might affect the extent to which they are able to meet such demands. Firstly, competition will affect the deposit rates that bank will offer, with a lower deposit rate reducing the amount that is withdrawn, allowing for larger withdrawal rates, which suggests that less competition will increase the resilience of banks. However, competition will also affect the lending behaviour of banks, If competition is less pronounced, banks may seek to take advantage of higher loan rates by providing more loans, reducing the amount of cash reserves they hold. The effect would be that less competition might adversely affect the vulnerability of banks to bank runs.

Let us assume that deposits D are provided to banks by consumers who do not know when they are requiring access to their funds. They only know that they will want to withdraw deposits with probability γ in time period 1, being paid $\left(1 + r_D^1\right) D$ to include interest r_D^1 for this one time period, and with probability $1 - \gamma$ they will retain the deposits until time period 2, being paid $\left(1 + r_D^2\right) D$ to include interest r_D^2 for retaining deposits with the bank. Banks use these deposits to finance a loan that is repaid with probability π at the end of two time periods, including interest r_L. As loans are given for two time periods, they cannot easily be recalled, but banks are able to liquidate any such loans and obtain a fraction λ of the initial loan amount L.

Banks will retain a fraction of the deposits as cash reserves such that they can repay the deposits of those depositors withdrawing early. Their cash reserves, $D-L$, are then given by

$$\gamma\left(1+r_D^1\right)D=D-L. \tag{15.79}$$

If we assume that banks are competitive, the remaining depositors will extract any surplus from banks. With the bank receiving the repayment of their loans, the payment to the remaining depositors is thus given by

$$(1-\gamma)\left(1+r_D^2\right)D=\pi\left(1+r_L\right)L. \tag{15.80}$$

Depositors seek to maximize their utility by choosing optimal level of interest for both time periods $\left(1+r_D^i\right)D$, subject to the constraints on the resources the bank has available from equations (15.79) and (15.80). We obtain the utility of the depositors, who do not know whether they want to withdraw early or late, as

$$\Pi_D=\gamma u\left(\left(1+r_D^1\right)D\right)+(1-\gamma)\,u\left(\left(1+r_D^2\right)D\right), \tag{15.81}$$

which after inserting for $\left(1+r_D^i\right)D$ from equations (15.79) and (15.80) gives us the optimal level of investment into loans L for the bank as

$$\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D}=\pi\left(1+r_L\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}. \tag{15.82}$$

In this scenario, the highest early withdrawal of deposits that can be supported is if all loans are liquidated, for which the bank obtains λL. Hence we need for the optimal amount of lending L^* and the optimal deposit rate $r_D^{1,*}$ that $\gamma\left(1+r_D^{1,*}\right)D\leq D-L^*+\lambda L^*$, or

$$\gamma\leq\gamma^*=\frac{D-(1-\lambda)\,L^*}{\left(1+r_D^{1,*}\right)D}. \tag{15.83}$$

If, in contrast, banks are not competitive but enjoy a monopoly, they would still be required to honour the early withdrawal of deposits and constraint (15.79) applies. Banks will have to repay depositors who have not withdrawn early their repayment of $(1-\gamma)\left(1+\hat{r}_D^2\right)D$ from the revenue that the loan repayments gives them. Hence their profits are given by

$$\Pi_B=\pi\left(1+r_L\right)\hat{L}-(1-\gamma)\left(1+\hat{r}_D^2\right)D. \tag{15.84}$$

Of course, depositors need to provide deposits and be better off than not making a deposit and keeping the deposits as cash. As a monopolist, the bank would extract any surplus from depositors, such that the utility from depositing and not depositing funds would be equal. Hence

$$\gamma u\left(\left(1+\hat{r}_D^1\right)D\right)+(1-\gamma)\,u\left(\left(1+\hat{r}_D^2\right)D\right)=u\,(D)\,. \tag{15.85}$$

Inserting the constraint arising from early withdrawal in equation (15.79) into equation (15.85), we can maximize the bank profits in equation (15.84) over the optimal investment into loans L and the repayment to the remaining depositors, $\left(1+\hat{r}_D^2\right)D$, subject to the constraint on deposits being made, equation (15.85). With ξ denoting the Lagrange multiplier, we easily get the first order conditions as

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\hat{L}} &= \pi\,(1+r_L)-(1-\gamma)\,\frac{\partial\left(1+\hat{r}_D^2\right)D}{\partial\hat{L}} \\ &\quad -\xi\left(-\frac{\partial u\left(\left(1+\hat{r}_D^1\right)D\right)}{\partial\left(1+\hat{r}_D^1\right)D}\right. \\ &\quad \left.+(1-\gamma)\,\frac{\partial\left(1+\hat{r}_D^2\right)D}{\partial\hat{L}}\frac{\partial u\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}\right)=0 \\ \frac{\partial\mathcal{L}}{\partial\left(1+\hat{r}_D^2\right)D} &= -\,(1-\gamma)-\xi\,(1-\gamma)\,\frac{\partial u\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}=0.\end{aligned} \tag{15.86}$$

From the second equation we directly obtain that the Lagrange multiplier is given as $\xi=-\frac{1}{\frac{\partial u\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}}$ and inserting this expression into the first condition, we recover the optimality condition (15.82) for depositors in the case that banks were competitive. Hence the decision by the bank is optimal for depositors.

Denote the optimal amount of loans given by the monopolist by $\hat{L}^*$ and the optimal deposit rate for those withdrawing deposits as $\hat{r}_D^{1,*}$. Let us now assume that $\hat{L}^*\leq L^*$, hence from the constraint on early withdrawals in equation (15.79) we have for those depositors withdrawing early that $\left(1+\hat{r}_D^{1,*}\right)D=\frac{D-\hat{L}^*}{\gamma}\geq\frac{D-L^*}{\gamma}=\left(1+r_D^{1,*}\right)D$. Given the usual assumption that the marginal utility is reducing, this implies that the marginal utility of the depositor with the monopolistic bank is lower than with the competitive bank, $\frac{\partial u\left(\left(1+\hat{r}_D^{1,*}\right)D\right)}{\partial\left(1+\hat{r}_D^1\right)D}\leq\frac{\partial u\left(\left(1+r_D^{1,*}\right)D\right)}{\partial\left(1+r_D^1\right)D}$ and the optimality condition for depositors in equation (15.82) similarly implies for the remaining depositors that $\frac{\partial u\left(\left(1+\hat{r}_D^{2,*}\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}\leq\frac{\partial u\left(\left(1+r_D^{2,*}\right)D\right)}{\partial\left(1+r_D^2\right)D}$, thus $\left(1+\hat{r}_D^{2,*}\right)D>\left(1+r_D^{2,*}\right)D$.

With the marginal utilities in the case of a monopolistic bank being smaller for those withdrawing early and those not withdrawing, the repayments would be larger for either type of depositor with a monopolistic bank. It is clear that if the repayments of deposits is larger if all surplus is extracted from depositors and $\Pi_D=u\,(D)$, the repayments in the case of competitive banks cannot be optimal. It therefore follows that our assumption that monopolistic banks lend less than competitive banks, $\hat{L}^*\leq L^*$, cannot be sustained and it must be that monopolistic banks lend more, $\hat{L}^*>L^*$.

In the case of a monopolistic bank we have the limits of early withdrawals given similarly as in equation (15.83); we obtain

$$\gamma \leq \hat{\gamma}^* = \frac{D - (1-\lambda)\hat{L}^*}{\left(1+\hat{r}_D^{1,*}\right)D}. \tag{15.87}$$

As $\hat{L}^* > L^*$ and from equation (15.79) we obtain that $\left(1+\hat{r}_D^{1,*}\right)D < \left(1+r_D^{1,*}\right)D$, we can now show that

$$\gamma^* < \hat{\gamma}^*. \tag{15.88}$$

To see this, note that $\gamma^* < \hat{\gamma}^*$ implies $\frac{D}{1-\lambda} > \frac{L^*\left(1+\hat{r}_D^{1,*}\right)D - \hat{L}^*\left(1+r_D^{1,*}\right)D}{\left(1+\hat{r}_D^{1,*}\right)D - \left(1+r_D^{1,*}\right)D}$ and given that $\hat{L}^* > L^*$, we have $\frac{D}{1-\lambda} > \frac{\hat{L}^*\left(1+\hat{r}_D^{1,*}\right)D - \hat{L}^*\left(1+r_D^{1,*}\right)D}{\left(1+\hat{r}_D^{1,*}\right)D - \left(1+r_D^{1,*}\right)D} > \frac{L^*\left(1+\hat{r}_D^{1,*}\right)D - \hat{L}^*\left(1+r_D^{1,*}\right)D}{\left(1+\hat{r}_D^{1,*}\right)D - \left(1+r_D^{1,*}\right)D}$, where the first inequality is true as $D > \hat{L}^* > (1-\lambda)\hat{L}^*$.

We thus see that monopolistic banks can accommodate a larger deposit withdrawals than competitive banks. The reason is that while monopolistic banks provide more loans to increase their profits and thus hold less cash reserves, they also pay lower deposit rates, reducing the resources required to repay any deposits that are being withdrawn. We have seen that the effect of lower loan rates dominates as monopolistic banks are able to accommodate larger withdrawals. We thus can expect them to be more resilient to any emerging bank run. Thus any policies aimed at increasing competition between banks might have to be accompanied by re-assurances to depositors about the stability of the banking system to avoid a higher likelihood of a bank run.

Reading Matsuoka (2013)

Conclusions

Bank runs can be self-fulfilling. Expecting other depositors to withdraw can make it rational to withdraw yourself, justifying the expectation of a bank run. No information about the bank itself is needed for such an outcomes, a shift in expectations is sufficient. It may even be that negative information about a bank has been obtained, but is only the expectation about how others will react to this information that may cause a bank run. Of course information itself can fully justify a bank run if the future return of deposits is affected and depositors seek to obtain repayments before losses increase further. If information is not available to all depositors at the same time, but only over time, it may even be optimal to not withdraw instantly, but retain deposits for the time being and obtain additional interest until finally withdrawing before the bank runs out of cash reserves. If any information is imperfect, banks runs might not always be efficient, they might overestimate the risks if receiving very negative information or doubt overly positive information, causing bank runs when none are justified; similarly, neutral information might be seen as overly reliable and bank runs that should occur will not occur.

It is common for deposit rates to include the risks associated with the risks the bank cafes from the loans they provide, but the possibility of bank runs should also be taking into consideration. If bank runs are a possibility, the higher cash reserves that are needed to accommodate any additional deposit withdrawals will reduce the profitability of the bank and hence the ability to give the remaining depositors high returns. Thus the possibility of bank runs will affect deposit rates.

Similarly, the competition between banks will affect their ability to withstand a bank run. Competition between banks sees them making very little profits from each loan, having to recover the lost profits by providing more loans. Having to give out more loans will reduce the cash reserves held by banks and hance make them more susceptible to the withdrawal of deposits. We can therefore expect more competitive banking systems to see more bank runs than banking systems where competition is less fierce.

Chapter 16

Interbank lending

A common observation is that banks do not only provide loans to companies, individuals, or governments, but also to other banks. This of course then implies that banks also take loans from other bank, thus deposits from the public get supplemented by these loans from other banks. To distinguish loans between banks from other loans and deposits, they are commonly referred to as interbank loans. Such loans are to a large degree short-term and of a fixed maturity, although it is not uncommon to extend interbank loans many times, but some of these interbank loans can have longer times to maturity comparable to that of loans to companies and individuals.

One view is that interbank loans are merely another type of loan a bank can provide, where the borrower happens to be a bank. Similarly, obtaining an interbank loan is comparable to obtaining a deposit from the general public. In chapter 16.1 we look at interbank loans seen as an investment and funding source by banks. However, the more common view of interbank loans is that they help to prevent bank runs by allowing banks to alleviate temporary liquidity shortages as will be shown in chapter 16.2. Similarly, chapter 16.3 shows that interbank loans can be seen as way of banks pooling their cash reserves to help a bank overcome a liquidity shock. While interbank loans are often seen as vehicles to overcome short-term liquidity shortages of otherwise healthy banks, there remains the risk that banks are facing solvency issues if loans they have provided to the general public are not repaid. Therefore, chapter 16.4 will take into account such credit risk and how this affects interbank lending.

Interbank lending is often conducted without the need to provide explicit collateral, but we show in chapter 16.5 how the availability of collateral obtained from loans to the general public can increase the ability of banks to secure additional interbank loans. They do so by being able to use the collateral they have obtained as a collateral to obtain loans themselves.

16.1 Interbank lending as investment

Providing an interbank loan can be seen as the provision of any other loan and will be done so if it is profitable to the bank. Similarly a bank will seek an interbank loan only if it is profitable for them. Hence, we can interpret interbank loans as a tool to maximize the profits of banks.

Interbank loans affect the amount of cash reserves R a bank holds. If an interbank loan of size M is obtained from another bank, cash reserves increase by this amount; treating interbank loans as deposits by another bank, the total deposits D will also increase by the amount of this interbank loan. If the bank provides an interbank loans to another bank, it reduces its cash reserves by the amount of the loan, M, but it does not affect their deposits.

Let us now assume that banks have an optimal level of cash reserves, R^*, that might be determined from the amount of deposits they expect to be withdrawn. Holding larger cash reserves, $R > R^*$, will make the bank less vulnerable to the withdrawal of deposits and we therefore assume that this will increase the utility of the bank. Given that cash reserves are held to cover any withdrawals of deposits, we propose that the ration of cash reserves and deposits if a relevant measure of this liquidity of the bank.

Banks will not only be concerned about their liquidity, but also the profits they make. If the loan rate for obtaining an interbank loan is $\hat{r}_M$, the profits of the bank are reduced from Π_B^0 to $\Pi_B^1 = \Pi_B^) - \hat{r}_M M$. Similarly if an interbank loans is granted to another bank at loan rate r_M, the profits increase to $\Pi_B^1 = \Pi_B^0 + r_M M$. We now propose that for the utility of the bank, they weigh the benefits from holding cash reserves against the profits they are making, where we assign a weight of θ to the cash reserves and a weight of $1 - \theta$ to the profits of the bank.

Hence, a bank prior to providing or obtaining interbank loans will have a utility of

$$U = \theta \frac{R - R^*}{D} + (1 - \theta) \Pi_B^0. \tag{16.1}$$

If the bank borrows the amount of M from another bank at rate $\hat{r}_M$, this utility changes to

$$U_B = \theta \frac{(R - R^*) + M}{D + M} + (1 - \theta) \left(\Pi_B^0 - \hat{r}_M M \right) \tag{16.2}$$

and if the bank provides a loan M to another bank, its utility will be

$$U_L = \theta \frac{(R - R^*) - M}{D} + (1 - \theta) \left(\Pi_B^0 + r_M M \right). \tag{16.3}$$

A bank will obtain or provide a loan only if it is beneficial for them to do so, thus we require that $U_B \geq U$ for a bank to borrow from other banks and $U_L \geq U$ to provide a loan to another bank. We thus require

$$\hat{r}_M \leq \hat{r}_M^* = \frac{\theta}{1-\theta} \frac{D-(R-R^*)}{D(D+M)}, \tag{16.4}$$
$$r_M \geq r_M^* = \frac{\theta}{1-\theta} \frac{1}{D}.$$

Banks are only borrowing from other banks if the interbank loan rate is sufficiently low such that the benefits of the increased cash reserves are outweighing the costs of this loan. Similarly, the interbank loan rate must be sufficiently high to ensure the benefits from the higher profits the bank will make are outweighing the costs of having lower cash reserves. We easily note that $0 < \hat{r}_M^* \leq r_M^*$, provided that $M > R^* - R$. Thus, as long as the interbank loan is at least the size of any shortfall in cash reserves, the interbank rate at which a bank is willing to borrow is lower than the interbank at which it is willing to lend.

Banks van only lend if another bank is seeking such an interbank loan. Generally banks will not have the same cash reserves, and if the denote the cash reserves of bank i by R_i, and the associated interbank loan rates by r_M^i and $\hat{r}_M^i$, respectively, then an interbank loan from bank i to bank j is feasible if the loan rate bank i charges, r_M^i is below the maximum rate at which bank j is willing to borrow, $\hat{r}_M^{j,*}$. Hence, for a possible agreement of an interbank loan we need $r_M^{i,*} \leq \hat{r}_M^{j,*}$, which easily becomes $M < R^* - R_j$. Hence if the cash reserves of the borrowing bank are sufficiently low, an interbank loan can be agreed. Especially, we require that $R_j < R^*$ for a feasible solution, implying that the cash reserves of the borrowing bank have to be below their optimal cash reserves.

Interbank loans are only given if it is profitable to do so, taking into account the impact these loans have on the cash reserves of a bank. Banks borrow from other banks if they face a shortfall in their desired cash reserves and while this imposes costs on them in the form of the interbank loan rate, the increased cash reserves compensate them for these costs. On the other hand, interbank loans are a means of banks with excess cash reserves to increase their profitability without reducing their cash reserves too much.

Reading Xiao & Krause (2022)

16.2 Insurance against bank runs

Banks hold cash reserves in order to be able to repay deposits that are withdrawn prior to the maturity of the loans they provide, see for example chapter 4.1 for a justification for such cash reserves. However, often it is not known how many of the deposits are going to be withdrawn and banks might hold excess cash reserves if they overestimate the amount that is withdrawn, or they might hold not enough cash reserves if the withdrawals are higher than anticipated. As long as banks are not all affected in the same way, but some banks hold excess cash reserves while others face a shortage of cash reserves, banks could provide each other with liquidity to allow them to repay withdrawn deposits. Banks that hold excess cash reserves could lend them to banks with a shortage of cash reserves. As banks are facing cash shortages only due to the unexpected high withdrawal of deposits, but not because of the loans they have given being of lower value, there is no risk associated with banks lending each other; after the loans are repaid, the banks will have sufficient resources to repay all remaining depositors and the interbank loans.

Banks provide loans L for two time periods, fully financed by deposits D; these deposits can be withdrawn at any time. Let us now assume that banks do not know the withdrawal rate of deposits after time period 1 which they face, but they know that it either a fraction γ_H will be withdrawn, or a fraction $\gamma_L < \gamma_H$. They also know that the high withdrawal rate γ_H occurs with probability p, and therefore the low withdrawal rate γ_L with probability $1 - p$. Depositors withdrawing in time period 1 obtain a repayment of $\left(1 + r_D^1\right) D$, where r_D^1 denotes the deposit rate applied to them and those remaining with the bank obtain $\left(1 + r_D^2\right) D$ in time period 2.

If we consider the first best solution where depositors on aggregate can be repaid if they withdraw, then cash reserves will be such that they pay for the average fraction γ of deposits being withdrawn, this

$$\gamma \left(1 + r_D^1\right) D = R, \tag{16.5}$$

with $\gamma = p\gamma_H + (1 - p)\gamma_L$ representing the average withdrawal rate

In time period 2, we assume that competition between banks ensures that depositors can extract all surplus from the bank, and hence the remaining depositors will be able to secure repayments that equal the revenue the bank obtains from the loans they have provided. These loans have been granted with interest r_L and are repaid with probability π. Hence we have

$$(1 - \gamma) \left(1 + r_D^2\right) D = \pi (1 + r_L) L. \tag{16.6}$$

This result, though, cannot be implemented easily as the actual withdrawal rate a bank faces will never be the average withdrawal rate γ, but either the high withdrawal rate γ_H or the low withdrawal rate γ_L. Hence, by holding cash reserves $\gamma \left(1 + r_D^1\right) D$, the bank would either hold too much cash reserves if γ_L is realized, or too little cash reserves if γ_H is realized. In the latter case the bank would fail, opening itself to bank runs.

In order to avoid this inefficiency when holding cash reserves, let us assume that banks can provide each other with interbank loans to cover any cash reserve shortage. Banks hold cash reserves R and the cash demands from withdrawn deposits are $\gamma_i \left(1 + r_D^1\right) D$, hence the cash shortfall of banks facing the high withdrawal rate γ_H is given by

$$M_H = \gamma_H \left(1 + r_D^1\right) D - R. \tag{16.7}$$

Similarly, banks facing the low withdrawal rate γ_L have excess cash reserves to the amount of

$$M_L = R - \gamma_L \left(1 + r_D^1\right) D. \tag{16.8}$$

Banks with excess cash M_L can now lend this amount as an interbank loan to banks with a shortage of cash reserves. With a fraction $1-p$ of banks having excess cash reserves due to low deposit withdrawals and a fraction p of banks with a cash shortages due to high deposit withdrawals, the demand for interbank loans and the supply of such loans matches if $pM_H = (1-p)\,M_L$. Inserting from equations (16.7) and (16.8), we recover condition (16.5) for the holding of cash reserves in the social optimum. Thus if we obtain the social optimum, the market for interbank loans will always be in equilibrium.

Those banks facing a cash shortages, thus having the high withdrawal rate γ_H, will have to pay an interest r_M on the interbank loan they obtain. The resources the bank has available to repay the loan emerge from the repaid loans, $\pi\,(1+r_L)\,L$, less the amount paid out to the those depositors remaining in time period 2, $(1-\gamma_H)\left(1+r_D^2\right)D$. Thus, assuming any surplus to the bank is extracted through the interbank loan, we have

$$\begin{aligned}(1+r_M)\,M_H &= \pi\,(1+r_L)\,L-(1-\gamma_H)\left(1+r_D^2\right)D \\ &= \frac{\gamma_H-\gamma}{1-\gamma}\pi\,(1+r_L)\,L,\end{aligned} \tag{16.9}$$

where the final equality has been obtained from solving equation (16.6) for $\left(1+r_D^2\right)D$ and inserting here. Solving equation (16.5) for $\left(1+r_D^1\right)D$and inserting this into equation (16.7) we obtain

$$M_H = \frac{\gamma_H-\gamma}{\gamma}R. \tag{16.10}$$

Once we insert this expression into equation (16.9), we get

$$1+r_M = \frac{\gamma}{1-\gamma}\pi\,(1+r_L)\,\frac{L}{R}. \tag{16.11}$$

We now see that the interbank loan rate is increasing in the average early withdrawals (γ) as the increased demand for cash in general raises interest rates. Higher returns from loans, $\pi\,(1+r_L)$, allow to extract more surplus from banks, thus increasing interbank loan rates. Finally, lower cash reserves, relative to the amount of loans provided and thus relative to deposits, also increases interbank rates due to higher demand for additional cash.

The profits of banks providing the interbank loan are given by

$$\Pi_L = \pi\,(1+r_L)\,L-(1-\gamma_L)\left(1+r_D^2\right)D+(1+r_M)\,M_L = 0, \tag{16.12}$$

where the final equation emerges if we insert for $\left(1+r_D^2\right)D$ from equation (16.6), and we note that $M_L = \frac{p}{1-p}M_H$ due to market the interbank loan market clearing, and M_H being given by equation (16.10). Hence interbank lenders do not make any profits that might incentivize them to hold excess cash reserves with the aim to provide interbank loans later.

Banks could, however, have incentives to hold too low cash reserves if the costs of borrowing in the interbank market is less than what they can earn from providing loans to the general public. Hence we require that $1 + r_M \geq \pi(1 + r_L)$, or with $L = D - R$

$$\gamma \geq \frac{R}{D}. \tag{16.13}$$

Hence if the rate of early withdrawals is too low, banks would have an incentive to reduce their cash holdings. Using equation (16.5), we can rewrite equation (16.13) as

$$r_D^1 \leq 0, \tag{16.14}$$

thus depositors withdrawing should not obtain interest. While this requirement seems rather strict, it is not completely unrealistic as deposits are risk-free in our model, thus they are always repaid, and given the early nature of their withdrawal, banks might not be willing to provide them with interest.

If banks face different levels of deposit withdrawals, those banks that have low withdrawal rates and thus have excess cash reserves would be willing to provide an interbank loan to other banks who face higher withdrawal rates and thus a cash shortage to meet this demand by depositors. It is therefore that interbank loans can be an efficient way to re-distribute cash reserves across banks and prevent some banks from having to fail, while other banks have excess cash reserves. Alternatively, all banks would have to hold high cash reserves, leading to lower lending. Holding such high cash reserves will be unnecessary for all those banks facing low withdrawal, leading to lower revenue from loans and subsequently lower repayments to those depositors not withdrawing.

Interbank loans can also be used to provide liquidity support to banks facing a bank run. Provided banks are re-assured that the bank faces no solvency problem in that loan repayments are lower than anticipated, they could provide the bank with interbank loans, allowing the bank to repay all depositors withdrawing and averting the failure of this bank. As long as depositors have confidence in the banking system as whole and move their deposits to other banks, these banks will have the requisite excess cash reserves arising from these transferred deposits to provide interbank loans.

Reading Bhattacharya & Gale (1987)

16.3 Insurance against liquidity shocks

Banks are faced with the potential withdrawal of deposits, either because depositors require their deposits returned, for example for consumption, or because they withdraw deposits to safeguard their deposits from future losses, a bank run. It might, however, also be the case that in addition to the withdrawals of deposits for consumption, unexpected demands arise on depositors that necessitate the withdrawal of deposits. Such an event, often referred to as a liquidity shock, can arise from the specific circumstances of depositors, such as the loss of employment, and would thus be specific to a single bank, in which case we speak of an idiosyncratic liquidity shock. However, it might also be that this liquidity shock affects all banks, for example in connection with a recession, referred to as a common liquidity shock. Banks can seek to support banks facing idiosyncratic liquidity shocks through interbank loans and prevent their failure.

Let us consider an economy with three banks, each holding cash reserves of R. We now consider the case that exactly one of these banks faces a significant idiosyncratic liquidity shock $S > 2R$, thus the liquidity shock this bank faces would exceed not only its own cash reserves, but another bank lending it its entire cash reserves would not be sufficient to prevent the bank to fail. Only if both of the other banks provide the bank with liquidity through interbank loans from their cash reserves, can the bank survive their idiosyncratic liquidity shock. The bank can only avoid failure if $S < 3R$ as otherwise the cash reserves of all banks combined would not be sufficient. We this assume that the idiosyncratic liquidity shock is such that $2R < M \leq 3R$.

If a bank faces this idiosyncratic liquidity shock, the other bank may be provide it with interbank loans. Once these interbank loans have been given, the banking system might face a common liquidity shock $\hat{S}$. This common liquidity shock occurs with probability p and its size is such that $0 < \hat{S} < R$. Hence, a bank facing the common liquidity shock has sufficient cash reserves and would not fail. However, if banks have provided interbank loans to the bank facing the idiosyncratic liquidity shock, they might have depleted their cash reserves sufficiently to fail due to the common liquidity shock.

Let us assume that bank 1 faces the idiosyncratic liquidity shock and banks 2 and 3 provide interbank loans of size M_i to this bank. Bank 1, facing the idiosyncratic liquidity shock, will have cash reserves of $R - S + M_2 + M_3$. If these cash reserves are less than the common liquidity shock $\hat{S}$, the bank will fail. The other two banks, those not facing the idiosyncratic liquidity shock, will have cash reserves of $R - M_i$ and they will fail if this is less than the common liquidity shock $\hat{S}$.

If we assume for simplicity that the common liquidity shock has a uniform distribution on the interval $[0; R]$, then we have the probability that a common liquidity shock will cause the bank to fail given as

$$
\begin{aligned}
Prob\left(\hat{S} > R - S + M_2 + M_3\right) &= 1 - \frac{R - S + M_2 + M_3}{R}, \\
Prob\left(\hat{S} > R - M_i\right) &= 1 - \frac{R - M_i}{R}
\end{aligned}
\tag{16.15}
$$

for the bank facing the idiosyncratic liquidity shock, and the two other banks, respectively. The expected number of banks failing if a common liquidity shock occurs is then given by

$$
\begin{aligned}
&Prob\left(\hat{S} > R - S + M_2 + M_3\right) + Prob\left(\hat{S} > R - M_2\right) \\
&\quad + Prob\left(\hat{S} > R - M_3\right) = \frac{S}{R}.
\end{aligned}
\tag{16.16}
$$

As the common liquidity shock only occurs with probability p, the total expected number of failing banks is given as $p\frac{S}{R}$.

If banks were not providing interbank loans to bank 1, facing an idiosyncratic liquidity shock, then this bank would fail and the remaining two banks would survive; thus we have one bank failing. The provision of interbank loans is therefore desirable if $p\frac{S}{R} \leq 1$, or

$$S \leq \frac{R}{p}. \tag{16.17}$$

As long as the idiosyncratic liquidity shock is not too large, it should be insured through the provision of interbank loans by other banks. Larger idiosyncratic liquidity shocks should not be insured as the required interbank loans would bo so large that it exposes the banks providing these loans to the risk of failing if facing the common liquidity shock. We note that the size of the interbank loan, M_i, is irrelevant as long as it allows bank 1 to survive. This is because an interbank loan in excess of this minimum amount would increase its likelihood of surviving the common liquidity shock, but diminish that of the banks providing the interbank loans by the same amount. Hence we might want to set the interbank loans such that it allows bank 1 to survive, $M_i = \frac{1}{2}(S - R)$. As we assumed that $S \leq 3R$, we see that $M_i \leq R$ and hence interbank loans of this size can be provided.

While the provision of such interbank loans might be desirable, banks giving them would seek to minimize any losses they have when providing such interbank loans. If the common liquidity shock materialises, which happens with probability p, and the bank they lend to fails, probability $1 - \frac{R-S+M_2+M_3}{R}$, they lose their interbank loan, M_i. In addition, they will face losses if the common liquidity shock occurs, p, and their cash reserves are not sufficient, $1 - \frac{R-M_i}{R}$; in this case banks would lose their equity E. We thus have the total losses when providing interbank loans given as

$$\Pi_B^i = p\left(1 - \frac{R - S + M_2 + M_3}{R}\right)M_i + p\left(1 - \frac{R - M_i}{R}\right)E. \tag{16.18}$$

Minimizing over the size of the interbank loan gives us for bank i the first order condition

$$\frac{\partial \Pi_B^i}{\partial M_i} = \frac{p}{R}\left(M - M_j + E - 2M_i\right) = 0, \tag{16.19}$$

where j indicates the other bank providing the interbank loan. These two conditions for banks 2 and 3 solve for

$$M_2 = M_3 = \frac{S + E}{3}. \tag{16.20}$$

As the optimal interbank loans cannot exceed the cash reserves of a bank, $M_i \leq R$, we need $S \leq 3R - E$ for this optimal solution to be implemented, which we assume to be the case. Inserting the optimal interbank loan size into equation (16.18) we obtain the total losses to the banks providing interbank loans as

$$\Pi_B^2 = \Pi_B^3 = \frac{1}{9}(S + E)^2. \tag{16.21}$$

These banks would provide interbank loans if these losses were less than those when not providing interbank loans. If a bank faces an idiosyncratic loss and no interbank loan is provided, it will fail and lose its equity E. With all banks being equal, there is a $\frac{1}{3}$ chance of facing such an idiosyncratic liquidity shock and the losses when not providing interbank loans would be $\frac{E}{3}$. In order for the provision of interbank loans to be profitable, we need $\frac{1}{9}(M+E)^2 \leq \frac{E}{3}$. This can easily be shown to require

$$S \leq \sqrt{3E} - E. \tag{16.22}$$

If we compare this requirement with the condition that interbank loans are socially desirable, $p\frac{S}{R} \leq 1$, we see that interbank loans are provided for larger idiosyncratic liquidity shocks than socially optimal if $p > \frac{R}{\sqrt{3E}-E}$ and if $p < \frac{R}{\sqrt{3E}-E}$ the provision of interbank loans is too restrictive in that only smaller idiosyncratic interbank loans are insured against through interbank loans. Thus if the likelihood of common liquidity shocks is high, interbank loans are given for too large idiosyncratic liquidity shocks, while in situations where common liquidity shocks are rarer, interbank loans are not forthcoming enough for larger idiosyncratic liquidity shocks.

The minimum size of an interbank loan to prevent the failure of a bank from an idiosyncratic liquidity shock is, as detailed above, $M_i = \frac{1}{2}(S-R)$ and comparing this with the interbank loan size in equation (16.20) we see that interbank loan actually given is larger than this minimum if

$$S < 3R + 2E. \tag{16.23}$$

As we had assumed that $M < 3R$, this condition is always fulfilled. Thus, banks will provide interbank loans that are larger than required by the idiosyncratic liquidity shock. These larger interbank loans are made to ensure a higher probability of these being repaid, because the receiving bank can withstand the common liquidity shock better. A bank being saved from the idiosyncratic liquidity shock could subsequently fail from the common liquidity shock if they do not have sufficient cash reserves. In this case, the banks providing the interbank loans would make losses from providing them; they will take this into account and provide a larger interbank loan, even if it increases the likelihood of them succumbing to the common liquidity shock.

Thus we see that if common liquidity shocks are rare, banks will not provide interbank loans to other banks facing larger idiosyncratic liquidity shocks, even though it would be desirable to do so. On the other hand if common liquidity shocks are more frequent, banks will happily provide interbank loans for smaller idiosyncratic liquidity shocks. If they provide interbank loans, though, these are more than merely covering the shortfall arising from the initial idiosyncratic liquidity shock to reduce the possibility of these loans being defaulted on due to the common liquidity shock.

We thus see that banks are willing to provide interbank loans to banks facing idiosyncratic liquidity shocks, although there are some deviations from the social optimum, and they generally provide interbank loans in excess of the amount needed. In order to obtain interbank loans themselves in order to survive an idiosyncratic liquidity shock, banks are willing to provide interbank loans to other banks in such a situation, even if this exposes them to the risk of a future liquidity shock. In most cases it will be sufficient to rely on banks insuring themselves against liquidity shocks without the need for the interference and support of central banks.

Reading Castiglionesi & Wagner (2013)

16.4 Counterparty risk

It is common to assume that interbank loans are risk-free. The justification for this assumption is that interbank loans are used to obtain additional cash reserves for banks facing a cash shortage due to the withdrawal of deposits, whether these are higher than expected or the result of a bank run. In such a scenario, the bank receiving the interbank loan only faces a temporary cash shortage and deposits, as well as interbank loans, can be repaid from the revenue that loans repaid by the general public generate. However, when providing interbank loans, it cannot necessarily be assumed that they are going to be repaid as banks might face losses from lower than expected repayment rates on loans they have provided to the general public. This possibility imposes counterparty risk on banks providing interbank loans, which will be taken into account when providing interbank loans.

Let us assume that banks provide loans L over multiple time periods to the general public on which interest r_L is payable; these loans get repaid with either probability π_H or $\pi_L < \pi_H$. Banks will not be aware of the probability of repayment on their loans when making lending decisions and only learn this information after one time period; they only know that the high repayment rate of π_H is achieved with probability p and the low repayment rate with probability $1 - p$.

While loans are given for long time periods, they can be liquidated at any time and yield a fraction λ_i of the face value of the loan. Loans with low repayment rates will yield a lower liquidation value than loans with higher repayment rates, $\lambda_L < \lambda_H < 1$. This can be justified by the observation that more risky loans, those with low repayment rates, are firstly having a lower value at maturity, $\pi_L (1 + r_L) L$, and will secondly be in general less in demand with possible buyers. In addition, we assume that $\pi_i (1 + r_L) > \lambda_i$ and hence liquidating loans is always inferior to holding them to maturity.

After one time period banks experience a liquidity shock due depositors withdrawing a fraction γ_j of their deposits. This withdrawal can be high at γ_H or low at $\gamma_L < \gamma_H$. Knowing that deposits might be withdrawn, banks will hold some cash reserves R and we assume that with deposits of D, the small liquidity shock would not exhaust the cash reserves of the bank, but the larger liquidity shock would not allow banks to repay the withdrawn deposits with cash reserves alone. Hence we assume that

$$\gamma_L D \leq R \leq \gamma_H D \tag{16.24}$$

Banks will not hold cash reserves of less than $\gamma_L D$ as in this case all banks would have a cash shortage to repay depositors, thus no interbank loans could be given to alleviate the cash shortage, and not more cash reserves than $\gamma_H D$ as in that case every bank would hold excess cash that does not generate any return with the prospect of being able to use these cash reserves to provide interbank loans.

In reaction to the liquidity shock, banks can liquidate a fraction α_i^j of loans to increase the cash reserves, where i denotes the repayments of loans and j the level of the liquidity shock. In addition to or instead of liquidating loans, banks might take an interbank loan M_i^j. Any excess cash reserves they hold can be retained as cash reserves, and we assume a fraction $\hat{\alpha}_i^j$ of the total cash available is kept as such, but banks may also give an interbank loan $M_i^{j,k}$, if they have loan repayments of type i, a liquidity shock of type j and provide a loan to a bank with loan repayments of type k.

We can now investigate the provision of interbank loans by analysing the behaviour of banks facing a liquidity shortage as they are subjected to high deposit withdrawals of $\gamma_H D$ and then turn to banks having excess cash reserves due to facing low deposit withdrawals of $\gamma_L D$.

Banks facing a cash shortage Let us at first consider a bank facing a high deposit withdrawal $\gamma_H D \geq R$ and thus a cash shortage. Such a bank would not provide interbank loans to other banks as it already has insufficient cash reserves to meet the demand for deposit withdrawals. This bank will obtain the revenue from the fraction $1-\alpha_i^H$ of loans L that have not been liquidated, yielding $\left(1-\alpha_i^H\right)(1+r_L)$ and will have cash reserves consisting of their original cash reserves and the cash raised from liquidating loans, $R+\alpha_i^H \lambda_i L$, of which they retain a fraction $\hat{\alpha}_i^H$. They then repay their interbank loan M_i^H, including interest r_M^i, as well as the remaining deposits $(1-\gamma_H)D$, on which interest r_D is paid. The bank can only obtain these profits if the loans to the general public are repaid, which happens with probability π_i, as only then is any revenue being generated. Thus the bank profits are given by

$$\begin{aligned}\Pi_B^{H,i} = \pi_i \Big(& (1+r_L)\left(1-\alpha_i^H\right)L + \hat{\alpha}_i^H\left(R+\alpha_i^H\lambda_i L\right) \\ & - \left(1+r_M^i\right)M_i^H - (1-\gamma_H)(1+r_D)D\Big). \end{aligned} \tag{16.25}$$

We assume here that deposits withdrawn in tim period 1 do not attract interest, nor do cash reserves attract any interest.

The maximization of the bank profits in equation (16.25) will be subject to constraints, that will attract Lagrange multipliers ξ_k. The amount of cash reserves required consists of the deposit withdrawals $\gamma_H D$ and the amount retained in cash, $\hat{\alpha}_i^H\left(R+\alpha_i^H\lambda_i L\right)$. The cash reserves available is the existing cash reserves R, the amount raised from liquidating loans, $\alpha_i^H\lambda_i L$, and the interbank loan, M_i^H. The cash reserves required cannot exceed the cash reserves available, hence we require

$$\gamma_H D + \hat{\alpha}_i^H\left(R+\alpha_i^H\lambda_i L\right) \leq R + \alpha_i^H\lambda_i L + M_i^H, \tag{16.26}$$

and associate Lagrange multiplier ξ_1 with his constraint. In addition, the interbank loan M_i^H cannot be negative, the fraction of loans liquidated must fulfill $0 \leq \alpha_i^H \leq 1$ and the fraction of cash reserves retained will also fulfill $0 \leq \hat{\alpha}_i^H \leq 1$. Hence the following restrictions are associated with Lagrange multipliers ξ_2, ξ_3, ξ_4, ξ_5, and ξ_6:

$$M_i^H \geq 0, \tag{16.27}$$

$$\alpha_i^H \geq 0, \tag{16.28}$$

$$\alpha_i^H \leq 1, \tag{16.29}$$

$$\hat{\alpha}_i^H \geq 0, \tag{16.30}$$

$$\hat{\alpha}_i^H \leq 0. \tag{16.31}$$

The first order conditions for the bank selecting the optimal amount of interbank loans, the optimal fraction of loans to liquidate, and the optimal fraction of cash reserves to retain are given by

$$\frac{\partial \mathcal{L}}{\partial M_i^H} = -\pi_i \left(1 + r_M^i\right) + \xi_1 + \xi_2 = 0, \tag{16.32}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \alpha_i^H} &= -\pi_i \left(1 + r_L\right) L + \pi_i \hat{\alpha}_i^H \lambda_i L \\ &\quad + \xi_1 \left(1 - \hat{\alpha}_i^H\right) \lambda_i L + \xi_3 - \xi_4 = 0,\end{aligned} \tag{16.33}$$

$$\frac{\partial \mathcal{L}}{\partial \hat{\alpha}_i^H} = \pi_i \left(R + \alpha_i^H \lambda_i L\right) - \xi_1 \left(R + \alpha_i^H \lambda_i L\right) + \xi_5 - \xi_6 = 0. \tag{16.34}$$

Let us now assume that the bank seeks an interbank loan, $M_i^H > 0$ and hence $\xi_2 = 0$ as constraint (16.27) is not binding. Thus the first order condition (16.32) solves for

$$\xi_1 = \pi_i \left(1 + r_M^i\right) > 0 \tag{16.35}$$

and after inserting this expression, the first order condition (16.34) simplifies to

$$-\pi_i r_M^i \left(R + \alpha_i^H \lambda_i L\right) + \xi_5 - \xi_6 = 0. \tag{16.36}$$

As the first term is negative, we need $\xi_5 > 0$ as all Lagrange multipliers are non-negative and hence $\xi_6 \geq 0$, implying that constraint (16.30) is binding and hence no cash reserves are retained by banks obtaining interbank loans, $\hat{\alpha}_i^H = 0$.

Knowing that $\hat{\alpha}_i^H = 0$, we can rewrite the first order condition (16.33) as

$$-\pi_i \left(\left(1 + r_L\right) - \left(1 + r_M^i\right) \lambda_i\right) L + \xi_3 - \xi_4 = 0. \tag{16.37}$$

This implies that for $1 + r_L - \left(1 + r_M^i\right) \lambda_i > 0$, we need $\xi_3 > 0$, therefore from constraint (16.28) that $\alpha_i^H = 0$ and no loans are liquidated if this condition is fulfilled, which can be rewritten as

$$1 + r_M^i \leq \frac{1 + r_L}{\lambda_i}. \tag{16.38}$$

Thus, if the interbank loan rate is not too high, banks facing a cash shortage will seek interbank loans rather than liquidate loans.

We have established that banks seeking interbank loans if the condition in equation (16.38) is fulfilled, will not seek to sell loans, thus $\alpha_i^H = 0$, and these banks will never retain any excess cash reserves, $\hat{\alpha}_i^H = 0$; hence the constraint on cash reserves from equation (16.26) is binding as well.

Having established the optimal excess cash reserves and liquidation of loans by banks facing a cash shortage, we can now assess the provision of interbank loans by banks with excess cash.

Banks with excess cash reserves Turning to the bank facing low deposit withdrawals $\gamma_L D \leq R$, who will therefore have excess cash reserves, they would not seek additional cash reserves through interbank loans as they already hold excess cash reserves. Their revenue is obtained from the loans they have not liquidated, $(1 + r_L)\left(1 - \alpha_i^L\right) L$ and the fraction of cash reserves retained, consisting of the existing cash reserves and the proceeds from the liquidation of loans, $\hat{\alpha}_i^L \left(R + \alpha_i^L \lambda_i L\right)$; in addition they obtain revenue from the interbank loans they have given to banks with high (low) repayments of loans, $\left(1 + r_M^H\right) M_i^{L,H}$ $(\left(1 + r_M^L\right) M_i^{L,L})$ and which are only repaid if these banks obtain the repayments from their loans to the general public. The bank finally repays the deposits that have not been withdrawn, $(1 - \lambda_L)(1 + r_D) D$. The bank can only obtain these profits if the loans to the general public are repaid, which happens with probability π_i, as only then is any revenue being generated. Thus the bank profits are given by

$$\begin{aligned} \Pi_B^{L,i} = \pi_i \Big(& (1 + r_L)\left(1 - \alpha_i^L\right) L + \hat{\alpha}_i^L \left(R + \alpha_i^L \lambda_i L\right) \\ & + \pi_H \left(1 + r_M^H\right) M_i^{L,H} + \pi_L \left(1 + r_M^L\right) M_i^{L,L} \\ & - (1 - \gamma_L)(1 + r_D) D \Big). \end{aligned} \tag{16.39}$$

The maximization of the bank profits in equation (16.25) will be subject to constraints, that will attract Lagrange multipliers $\hat{\xi}_k$. The cash reserves required consist of the withdrawn deposits $\gamma_L D$, the retained cash reserves $\hat{\alpha}_i^L \left(R + \alpha_i^L \lambda_i L\right)$, as well as the interbank loans given to banks with high and low repayment rates, $M_i^{L,H}$ and $M_i^{L,L}$. The cash reserves available consists of the original cash reserves R and the proceeds from the liquidated loans, $\alpha_i^L \lambda_i L$. The cash reserves required cannot exceed the cash reserves available, hence we require

$$\gamma_L D + \hat{\alpha}_i^L \left(R + \alpha_i^L \lambda_i L\right) + M_i^{L,H} + M_i^{L,L} \leq R + \alpha_i^L \lambda_i L, \tag{16.40}$$

and associate Lagrange multiplier $\hat{\xi}_1$ with his constraint. In addition, the interbank loans cannot be negative, $M_i^{L,H} \geq 0$ and $M_i^{L,L} \geq 0$, the fraction of loans liquidated has to fulfill $0 \leq \alpha_i^L \leq 1$ and the fraction of cash reserves retained also have to fulfill $0 \leq \hat{\alpha}_i^L \leq 1$. Hence the following restrictions are associated with Lagrange multipliers $\hat{\xi}_2$, $\hat{\xi}_3$, $\hat{\xi}_4$, $\hat{\xi}_5$, $\hat{\xi}_6$, and $\hat{\xi}_7$:

$$M_i^{L,H} \geq 0, \tag{16.41}$$
$$M_i^{L,L} \geq 0, \tag{16.42}$$
$$\alpha_i^L \geq 0, \tag{16.43}$$
$$\alpha_i^L \leq 1, \tag{16.44}$$
$$\hat{\alpha}_i^L \geq 0, \tag{16.45}$$
$$\hat{\alpha}_i^L \leq 1. \tag{16.46}$$

The first order conditions for the bank selecting the optimal amount of interbank loans, the optimal fraction of loans to liquidate, and the optimal fraction of cash reserves to retain are given by

$$\frac{\partial \mathcal{L}}{\partial M_i^{L,H}} = \pi_i \pi_H \left(1 + r_M^H\right) - \hat{\xi}_1 + \hat{\xi}_2 = 0, \tag{16.47}$$
$$\frac{\partial \mathcal{L}}{\partial M_i^{L,L}} = \pi_i \pi_L \left(1 + r_M^L\right) - \hat{\xi}_1 + \hat{\xi}_3 = 0, \tag{16.48}$$
$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \alpha_i^L} = {} & -\pi_i \left(1 + r_L\right) L + \hat{\alpha}_i^L \pi_i \lambda_i L + \\ & \left(1 - \hat{\alpha}_i^L\right) \lambda_i L \hat{\xi}_1 + \hat{\xi}_4 - \hat{\xi}_5 = 0,\end{aligned} \tag{16.49}$$
$$\frac{\partial \mathcal{L}}{\partial \hat{\alpha}_i^L} = \pi_i \left(R + \alpha_i^L \lambda_i L\right) - \hat{\xi}_1 \left(R + \alpha_i^L \lambda_i L\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. \tag{16.50}$$

Let us at first assume that the bank gives an interbank loan to a bank whose repayment rate is high, $M_i^{L,H} > 0$, but not to a bank whose repayment rate is low, $M_i^{L,L} = 0$. This implies from constraint (16.41) that this constraint is not binding and hence $\hat{\xi}_2 = 0$. Then the first order condition in equation (16.47) solves for

$$\hat{\xi}_1 = \pi_i \pi_H \left(1 + r_M^H\right) > 0, \tag{16.51}$$

thus constraint (16.40) is binding.

Inserting for $\hat{\xi}_1$ from equation (16.51) into equation (16.50) we obtain

$$\pi_i \left(R + \alpha_i^L \lambda_i L\right) \left(1 - \pi_H \left(1 + r_M^H\right)\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. \tag{16.52}$$

If $\pi_H \left(1 + r_L^H\right) < 1$, then the first term is positive and as $\hat{\xi}_6 \geq 0$, we need $\hat{\xi}_7 > 0$, thus making constraint (16.46) binding such that $\hat{\alpha}_i^L = 1$. But in this case all proceed would be held in cash, allowing no interbank loan, contradicting our assumption that $M_i^{L,H} > 0$. Thus we need $\pi_H \left(1 + r_L^H\right) \geq 1$.

With $M_i^{L,L} = 0$, constraint (16.42) is binding and hence $\hat{\xi}_3 > 0$, giving us from first order condition (16.48) that

$$\hat{\xi}_1 = \pi_i \pi_L \left(1 + r_M^L\right) + \hat{\xi}_3 > 0. \tag{16.53}$$

Inserting this result into the first order condition (16.50) we get

$$\pi_i \left(R - \alpha_i^L \lambda_i L\right) \left(1 - \pi_L \left(1 + r_M^L\right)\right) - \hat{\xi}_3 \left(R - \alpha_i^L \lambda_i L\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. \tag{16.54}$$

If $\pi_L \left(1 + r_M^L\right) > 1$, then the first two terms are negative and thus we need $\hat{\xi}_6 > 0$, implying from constraint (16.45) that this is binding, thus $\hat{\alpha}_i^L = 0$ and no cash is retained. But from equation (16.53) we know that constraint (16.40) is binding and the excess cash needs to be invested. Hence the assumption that $\pi_L \left(1 + r_M^L\right) > 1$ cannot hold and we find that

$$\pi_L \left(1 + r_M^L\right) \leq 1 \leq \pi_H \left(1 + r_M^H\right). \tag{16.55}$$

Let us now consider a bank providing interbank loans to both types of banks, thus $M_i^{L,H} > 0$ and $M_i^{L,L} > 0$ such that $\hat{\xi}_2 = \hat{\xi}_3 = 0$ as constraints (16.41) and (16.42) are not binding. From the first order conditions (16.47) and (16.48) we then immediately get that the returns on interbank loans to the two types of banks are identical,

$$\pi_L \left(1 + r_M^L\right) = \pi_H \left(1 + r_M^H\right) \tag{16.56}$$

and

$$\hat{\xi}_1 = \pi_i \pi_H \left(1 + r_M^H\right) > 0. \tag{16.57}$$

The first order condition (16.50) then becomes

$$\pi_i \left(R + \alpha_i^L \lambda_i L\right) \left(1 - \pi_H \left(1 + r_M^H\right)\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0 \tag{16.58}$$

and if $\pi_H \left(1 + r_M^H\right) < 1$, the first term is positive, thus requiring $\hat{\xi}_7 > 0$, such that constraint (16.46) is binding and $\hat{\alpha}_i^L = 1$, meaning all funds are held as cash reserves and no interbank loan can be given; this implies that we need $\pi_H \left(1 + r_M^H\right) \geq 1$. In that case we then require $\hat{\xi}_6 > 0$, implying no holding of cash reserves and either $M_i^{L,H} > 0$ or $M_i^{L,L} > 0$. As the returns on interbank loans to the two types of banks are equal from equation (16.56), in general both banks will obtain interbank loans and we only consider this case.

If we assume that $M_i^{L,L} > 0$, then $\hat{\xi}_3 = 0$ and the first order condition (16.48) becomes

$$\hat{\xi}_1 = \pi_i \pi_L \left(1 + r_M^L\right) > 0 \tag{16.59}$$

and first order condition (16.50) will be

$$\pi_i \left(R + \alpha_i^L \lambda_i L\right) \left(1 - \pi_L \left(1 + r_M^L\right)\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. \tag{16.60}$$

Knowing that $\pi_L \left(1 + r_M^L\right) = \pi_H \left(1 + r_M^H\right) > 1$, we see that we need $\hat{\xi}_6 > 0$, implying that no cash reserves are retained, $\hat{\alpha}_i^L = 0$. Similarly for $M_i^{L,H} > 0$ we have also find that no cash reserves are retained, $\hat{\alpha}_i^H = 0$. Thus banks providing interbank loans do not retain cash reserves.

Looking at the provision of interbank loans to a bank with high repayment rates, $M_i^{L,H} > 0$, we see that $\hat{\xi}_2 = 0$ as the constraint is not binding and hence from the first order condition (16.47) we obtain that $\hat{\xi}_1 = \pi_i \pi_H \left(1 + r_M^H\right) > 0$. Using this result and $\hat{\alpha}_i^L = 0$, we can rewrite first order condition (16.49) as

$$-\pi_i \left(1 + r_L - \lambda_i \pi_H \left(1 + r_M^H\right)\right) L + \hat{\xi}_4 - \hat{\xi}_5 = 0. \tag{16.61}$$

If $1 + r_L - \lambda_i \pi_H \left(1 + r_M^H\right) > 0$, then the first term is negative and we need $\hat{\xi}_4 > 0$, implying from constraint (16.43) that $\alpha_i^L = 0$ and the lender does not liquidate any loans. The same result we obtain for interbank loans provided to banks with low repayment rates. Hence we require

$$\pi_H \left(1 + r_M^H\right) = \pi_L \left(1 + r_M^L\right) \leq \frac{1 + r_L}{\lambda_i} \tag{16.62}$$

such that the bank does not liquidate any loans they have given to the general public. Thus, if the interbank loan rate is not too high, the bank will not liquidate loans to generate additional cash reserves in order to be able to provide more interbank loans. This constraint is most binding for $\lambda_i = \lambda_H$ and hence if equation (16.62)is fulfilled for banks with high repayment rates, it will also be fulfilled for banks with low repayment rates. And similarly, if equation (16.69) is not fulfilled for $\lambda_i = \lambda_L$, then it will also not be fulfilled for $\lambda_i = \lambda_L$. Thus, if banks with high repayment rates do not liquidate loans, those banks with low repayment rates will not do so; and if banks with low repayment rates liquidate loans, banks with high repayment rates will also do so. Banks with low repayment rates are less inclined to liquidate their loans as they obtain a lower fraction of the loan than those banks with higher repayment rates, making the liquidation of loans more costly for banks with lower repayment rates than for banks with higher repayment rates.

We have thus established that banks providing interbank loans do not hold excess cash reserves, $\alpha_i^L = 0$, and if the condition in equation (16.62) is fulfilled they do not liquidate any loans, $\hat{\alpha}_i^H = 0$. The constraint on cash reserves from equation (16.40) is binding as well.

Having established the provision of interbank loans by banks facing excess cash reserves, we can now establish the conditions under which interbank loans can actually be agreed between banks.

Interbank lending without liquidation Firstly, we compare the condition in equation (16.62) for banks to not liquidating loans to provide more interbank loans and equation (16.38) for banks to seek interbank loans rather than liquidate loans, and see that the latter is more strict and hence we require that

$$1 + r_M^i \leq \frac{1 + r_L}{\lambda_i} \tag{16.63}$$

for banks to not liquidate any loans.

We can now insert our results so far and rewrite the profits of the banks seeking and providing interbank loans, respectively, by inserting these into equations (16.25) and (16.39), obtaining the profits as

$$\begin{aligned} \Pi_B^{H,i} &= \pi_i \left((1 + r_L) L - \left(1 + r_M^i\right) M_i^H - (1 - \lambda_H)(1 + r_D) D \right), \qquad (16.64) \\ \Pi_B^{L,i} &= \pi_i \left((1 + r_L) L + \pi_H \left(1 + r_M^H\right) M_i^{L,H} \right. \\ &\quad \left. + \pi_L \left(1 + r_M^L\right) M_i^{L,L} - (1 - \lambda_L)(1 + r_D) D \right). \end{aligned}$$

Similarly, the constraints on interbank loans from constraints (16.26) and (16.40), who are binding, become

$$\begin{aligned} M_i^{L,H} + M_i^{L,L} &= R - \lambda_L D, \\ M_i^H &= \lambda_H D - R. \end{aligned} \tag{16.65}$$

Thus the amount of interbank loans given is equal to the remaining cash reserves after repaying deposits and the interbank loan demanded is the cash shortfall of the bank.

While banks know the repayment rate of their loans and the size of the deposit withdrawals in time period 1, we assume that they are unaware of this when making the lending decision. In this case their expected profits are given by

$$\Pi_B = \gamma_L \left(p \Pi_B^{L,H} + (1 - p) \Pi_B^{L,L} \right) + \gamma_H \left(p \Pi_B^{H,H} + (1 - p) \Pi_B^{H,L} \right). \tag{16.66}$$

Inserting equations (16.65) into equation (16.64) and this result into equation (16.66), we can get the first other condition for the optimal amount of loans to the public, after noting that $R = D - L$ and $\pi_H \left(1 + r_M^H\right) = \pi_L \left(1 + r_M^L\right)$, as

$$\frac{\partial \Pi_B}{\partial L} = (\gamma_H + \gamma_L) \pi (1 + r_L) - \pi_H (\gamma_H + \gamma_L \pi) \left(1 + r_M^H\right) = 0, \tag{16.67}$$

where $\pi = p\pi_H + (1 - p)\pi_L$ denotes the average repayment rate of loans. This first order condition solves for

$$1 + r_M^H = \frac{\pi}{\pi_H} \frac{\gamma_H + \gamma_L}{\gamma_H + \pi \gamma_L} (1 + r_L). \tag{16.68}$$

Inserting this result into the condition that no interbank loans are liquidated, equation (16.63), we get the requirement that

$$\frac{\pi}{\pi_H} \frac{\gamma_H + \gamma_L}{\gamma_H + \pi \gamma_L} \leq \frac{1}{\lambda_i}. \tag{16.69}$$

Hence if this condition is fulfilled, the borrowing bank will not liquidate loans, but rely on interbank borrowing and the lending bank will not liquidate loans to increase their ability to provide additional interbank loans. This condition can be rewritten as

$$\pi_L \leq \pi_L^* = \frac{\gamma_H \pi_H}{(1-p)\left(\lambda_i\left(\gamma_H+\gamma_L\right)-\gamma_L \pi_H\right)}-\frac{p}{1-p} \pi_H . \tag{16.70}$$

Of course, banks need to be willing to provide interbank loans, hence this cannot be imposing a loss on them and the expected return must cover at least the funding costs by deposits, hence we require $\pi_H\left(1+r_M^H\right) \geq 1+r_D$. Inserting from equation (16.68), this easily solves for

$$\pi_L \geq \pi_L^{**}=\frac{\gamma_H \pi_H}{(1-p)\left(\left(\gamma_H+\gamma_L\right)\left(1+r_L\right)-\gamma_L \pi_H\left(1+r_D\right)\right)}-\frac{p}{1-p} \pi_H . \tag{16.71}$$

Hence, if $\pi_L^{**} \leq \pi_L \leq \pi_L^*$, we see that interbank loans are provided and no loans to the general public are liquidated. If $\pi_L<\pi_L^{**}$, then no interbank loans are offered as the interbank loans rate does not cover the funding costs and if $\pi_L>\pi_L^*$, the interbank loan rate is too high and banks prefer liquidating loans if needing additional cash reserves.

Interbank lending with liquidation If the condition in equation (16.69) is not fulfilled, then no interbank lending occurs. In the case that $\frac{1}{\lambda_H}<\frac{\pi}{\pi_H} \frac{\gamma_H+\gamma_L}{\gamma_H+\pi \gamma_L} \leq \frac{1}{\lambda_L}$, the bank with low repayment rates will seek to obtain interbank loans, but the bank with high repayment rates will rather liquidate their loans; we thus find that $M_H^H=0$.

Let us now assume that in this case constraint (16.26) is not binding, hence $\xi_1=0$, and from the first order condition (16.34) we easily get that

$$\pi_H\left(R+\alpha_H^H \lambda_H L\right)+\xi_5-\xi_6=0, \tag{16.72}$$

and as the first term is positive, we require that $\xi_6>0$. Hence from constraint (16.31) we obtain that $\hat{\alpha}_H^H=1$, thus constraint (16.26) becomes $\gamma_H D \leq 0$, which is a contradiction and we will have $\xi_1>0$ such that constraint (16.26) is binding and as $R \leq \gamma_H D$, we need $\alpha_H^H>0$, implying from constraint (16.28) that $\xi_3=0$. Inserting this into the first order condition (16.33) we get

$$\left(\pi_H-\xi_1\right) \hat{\alpha}_H^H \lambda_H L+\left(\xi_1 \lambda_H-\pi_H\left(1+r_L\right)\right) L-\xi_4=0 . \tag{16.73}$$

The condition that $\alpha_H^H>0$ shows that the bank will liquidate some of their loans to increase their cash position and be able to repay all deposits that have been withdrawn.

If $\hat{\alpha}_H^H>0$, then from constraint (16.30) we find $\xi_5=0$ and the first order condition (16.34) becomes

$$\left(\pi_H-\xi_1\right)\left(R+\alpha_H^H \lambda_H L\right)-\xi_6=0 . \tag{16.74}$$

From equation (16.35) it is obvious that $\xi_1>\pi_H$ and hence that the first term is negative, requiring $\xi_6<0$, which is impossible and hence $\hat{\alpha}_H^H=0$. Thus no cash reserves are retained. Using these results we obtain the constraints on the maximization of the bank profits as

$$\begin{aligned} M_L^H &= \lambda_H D - R, \\ M_H^H &= 0, \\ M_i^{L,H} + M_i^{L,L} &= R - \lambda_L D, \\ \alpha_H^H \lambda_H L &= \lambda_H D - R. \end{aligned} \tag{16.75}$$

Here the first and third terms show that the interbank loan is such that it covers the excess cash and the cash shortfall, respectively, and the final term that the proceeds from loans sold have to cover the cash shortfall.

Inserting conditions (16.75) into the profits of the bank in equation (16.64) and solving the first order condition for the optimal loan amount from $\frac{\partial \Pi_B}{\partial L} = 0$, we get the interbank loan rate as

$$1 + r_M^H = \frac{\pi(\gamma_H + \gamma_L) - \gamma_H p \pi_H \frac{1}{\lambda_H}}{(1-p)\gamma_H + \pi \gamma_L} \frac{1 + r_L}{\pi_H}. \tag{16.76}$$

Inserting this interbank loan rate into equation (16.63), we obtain that

$$\pi_L \leq \hat{\pi}_L^* = \frac{\pi_H (1-p)\gamma_H + \frac{\lambda_L \gamma_H p \pi_H}{\lambda_H}}{(1-p)(\lambda_L(\gamma_H + \gamma_L) - \pi_H \gamma_L)} - \frac{p}{1-p}\pi_H. \tag{16.77}$$

Thus if $\pi_L \hat{\pi}_L^*$ both types of banks would use interbank loans. If $\pi_L \leq \hat{\pi}_L^*$ only the banks with low repayment rates.

Banks need to be willing to make interbank loans, hence this cannot be imposing a loss, hence we need $\pi_H \left(1 + r_M^H\right) \geq 1 + r_D$

$$\pi_L \geq \hat{\pi}_L^{**} = \frac{\frac{\gamma_H p \pi_H}{\lambda_H} + (1-p)\pi_H}{(\gamma_H + \gamma_L)\frac{1+r_L}{1+r_D} - \gamma_L} - \frac{p}{1-p}\pi_H. \tag{16.78}$$

Therefore, if $\hat{\pi}_L^{**} \leq \pi_L \leq \hat{\pi}_L^*$, we see that interbank loans demanded by banks with low repayment rates, while those with high repayment rates liquidate loans to increase their cash reserves. If $\pi_L < \hat{\pi}_L^{**}$, then no interbank loans are offered as the interbank loans rate does not cover the funding costs and if $\pi_L > \hat{\pi}_L^*$, the interbank loan rate is too high for the bank with low repayment rates and it prefers liquidating loans if needing additional cash reserves.

Figure 16.1 visualizes our results. We see that if the low repayment rate is sufficiently high, the interbank loan rate will be too high for banks to demand interbank loans and they will prefer to liquidate loans instead. Thus, if there is not much difference in the risk between banks and banks with lower repayment rates, high-risk banks, are not too common, all banks will liquidate loans instead of turning to interbank loans. This suggests that in banking systems where there is very little adverse selection between banks and high-risk banks are rare, interbank loans are demanded much.

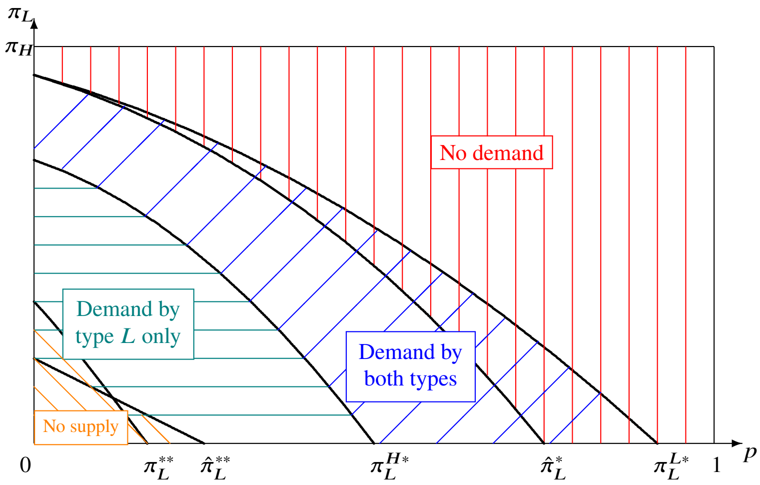


Fig. 16.1: Interbank loans and liquidation

If the differences between banks become larger, thus the risk of the high-risk banks increase or they become more common, it will only be the high-risk banks demanding interbank loans. The higher losses when liquidating loans let them turn to interbank loans as despite the high interbank rate, these costs are smaller. The costs to banks with higher success rates, low-risk banks, will still be lower from liquidating loans. We observe an area where both equilibria can exist. Increasing the risk of high-risk banks further, makes liquidating loans less and less attractive as the costs are increasing and the interbank loan rate becomes attractive to both bank types, thus they will both demand interbank loans. As risks increase further, the provision of interbank loans becomes unprofitable and eventually they will not longer be offered.

Summary If banks might not be able to repay their interbank loans due to the loans they have provided to the general public defaulting, interbank loans will still be supplied to cover any shortfall in cash reserves by those with excess cash reserves, unless the risks are too high. However, if risks are sufficiently low, the demand for interbank loans will be limited as banks can find it more attractive to liquidate their assets as the associated costs might be lower than the interest payable on interbank loans. These interbank loans rates will, of course, take into account the risks banks take when providing them, and can therefore be quite substantial.

We can therefore expect that banks turn more towards liquidating assets in times of liquidity shortages rather than relying on the interbank market if banks are quite homogenous. If the differences in the risks banks are taking are small, and the high-risk banks are not too common, demand for interbank loans might be low. However, it has to be taken into account that the liquidation of assets cannot be conducted within a very short time frame, while interbank loans can be agreed more easily and faster. Thus in the case of sudden liquidity shortages, there is no alternative to interbank loans, but longer-term liquidity shortage might well lead to the liquidation of assets.

Reading Heider, Hoerova, & Holthausen (2015)

16.5 Interbank lending with collateral pyramids

Banks provide interbank loans typically without obtaining collateral as they are provided to assist in the management of cash reserves and are deemed to be risk-free. However, banks may use interbank loans instead to expand their provision of loans, which is an inherently risky; in this case interbank loans may only be given if collateral is provided. bank can now use the returns generated from this additional investment as collateral to obtain further interbank loans. Hence using the initial collateral, banks can secure loans that are secured on future revenue based on investments made using such a loan. This is referred to as a collateral pyramid.

A bank has deposits D, on which it pays interest r_D, and obtains an interbank loan M_0 at a loan rate of r_M and invests these proceeds into a loan L to the general public with a loan rate r_L, which is repaid with probability π, and into a risk-free bond B, where the risk-free rate is r. The profits of this bank are given by

$$\begin{aligned}\Pi_B = {} & \pi\left(1+r_L\right)\left(D-B+M_0\right)+\left(1+r\right)B \\ & -\left(1+r_M\right)M_0-\left(1+r_D\right)D,\end{aligned} \tag{16.79}$$

If instead of providing the loan to the general public, a bank can also just provide interbank loans. In this case they invest their deposits into such interbank loans and the bond, such that its profits are

$$\hat{\Pi}_B = \left(1+r_M\right)\left(D-B\right)+\left(1+r\right)B-\left(1+r_D\right)D. \tag{16.80}$$

In order for banks to provide loans to the general public, we need $\Pi_B \geq \hat{\Pi}_B$, or

$$\pi \geq \pi^* = \frac{1+r_M}{1+r_L}, \tag{16.81}$$

using that $D-B+M>0$.

Let us now assume that the bond B acts as a collateral for the interbank loan M_0. If the bank decided to default strategically, they would obtain the proceeds from their loan, but not repay their interbank loans and lose its collateral, the bond B. If we assume they would nevertheless repay their depositors, their profits from strategic default is given by

$$\Pi_B^S = \pi\left(1+r_L\right)\left(D-B+M_0\right)-\left(1+r_D\right)D. \tag{16.82}$$

If repaying the interbank loan, the bank would make the profits as detailed in equation (16.79), but also be able to raise more interbank loans to a total size of M in the future which are invested into loans to the general public, such that its profits are then

$$\Pi_B^R = \Pi_B + (\pi (1 + r_L) - (1 + r_M)) (M - M_0), \tag{16.83}$$

To avoid strategic default, we need that it is more profitable to repay the interbank loan than it is not to do so, hence $\Pi_B^R \geq \Pi_B^S$. This requires $1 + r_M \geq \frac{1+r_M}{1+r_L}$, which is the same condition as in equation (16.81) for bank to lend to the general public.

Banks can now use the revenue generated from investing into loans to the general public the funds obtained an interbank loan M_k as collateral for another interbank loan, M_{k+1}. We assume that due to the risks associated with these loans, they can only obtain a fraction λ of this revenue as a new interbank loan, hence $M_{k+1} = \lambda\pi (1 + r_L) M_k$. The total revenue arising from these k interbank loans is $\pi (1 + r_L) (D - B - M_0) + \sum_{j=1}^{k} (\pi (1 + r_L))^j (D - B - M_0)$ and the costs of these loans are $(1 + r_M) \sum_{j=1}^{k} M_j$. Thus the profits generated to the bank is given by

$$\begin{aligned}\hat{\Pi}_B^R = {} & \pi (1 + r_L) (D - B - M_0) + (1 + r) B - (1 + r_D) D \\ & + \sum_{j=1}^{k} (\lambda\pi (1 + r_L))^j (D - B - M_0) - (1 + r_M) \sum_{j=1}^{k} M_j.\end{aligned} \tag{16.84}$$

If they strategically defaulted on the final interbank loan, M_k, they will lose the collateral they pledged, which is the revenue from interbank loans M_{k-1}, $\lambda\pi (1 + r_L) M_{k-1}$ and do not repay this interbank loan. Hence the profits with strategic default are given by

$$\begin{aligned}\hat{\Pi}_B^S = {} & \pi (1 + r_L) (D - B - M_0) + (1 + r) B - (1 + r_D) D \\ & + \sum_{j=1}^{k-1} (\lambda\pi (1 + r_L))^j (D - B - M_0) - (1 + r_M) \sum_{j=1}^{k-1} M_j.\end{aligned} \tag{16.85}$$

The collateral of other interbank loans, including the original bond, are not affected as these are repaid. We see that banks repay their interbank loan M_k if $\hat{\Pi}_B^R \geq \hat{\Pi}_B^S$, which requires

$$M_k \leq \frac{(\lambda\pi (1 + r_L))^k (D - B + M_0)}{1 + r_M}. \tag{16.86}$$

As banks seek to maximize their profits, they will obtain the highest possible interbank loan, such that this condition will be met with equality. We also can derive from this condition that $M_k = \lambda\pi (1 + r_L) M_{k-1}$, very much in line with the constraint on the provision of collateral to obtain additional interbank loans. Therefore, if this constraint is fulfilled for M_{k-1}, it will be fulfilled for M_k, which leaves us to show the condition that needs to be fulfilled for M_1. Setting $k = 1$ in equation (16.86) gives us

$$M_1 \leq \frac{\lambda\pi (1 + r_L) (D - B + M_0)}{1 + r_M}. \tag{16.87}$$

The total amount of interbank lending is now given by summing up all the possible interbank loans, such that

$$\begin{aligned} M &= M_0 + \sum_{k=1}^{+\infty} M_k \qquad (16.88) \\ &= M_0 + \sum_{k=1}^{+\infty} \left(\lambda\pi^* \left(1 + r_L\right)\right)^{k-1} M_1 \\ &= M_0 + \sum_{k=0}^{+\infty} \left(\lambda\pi^* \left(1 + r_L\right)\right)^{k} M_1 \\ &= M_0 + \frac{M_1}{1 - \lambda\pi^* \left(1 + r_L\right)}, \end{aligned}$$

where we assume that $\lambda\pi^* \left(1 + r_L\right) < 1$. The first term, M_0 denotes the amount of interbank lending, and hence provision of loans to the general public, that is based on outside collateral, the value of the bond; the represents the interbank loans that are based on inside collateral, the revenue generated from using as collateral the revenue from loans given by the use of other interbank loans. The value of the bond at maturity is $(1 + r)\, B$ and the repayment of the initial interbank loan is $(1 + r_M)\, M_0$. Given the bond is risk-free, we can assume that the interbank loan the bank can obtain for this bond is such that these values at repayment are equal and hence

$$M_0 = \frac{1 + r}{1 + r_M} B. \qquad (16.89)$$

Providing collateral to obtain an interbank loan, banks can use the revenue generated from investing its proceeds as collateral for an additional interbank loan that in turn can be invested, leading to an ever increasing amount of interbank loans that are collaterized against the revenue obtained from investments made with other interbank loans. While the original collateral is not re-used, its use to generate more revenue can increase the use of interbank loans significantly. Such pyramid of ever smaller interbank loans can expand lending by banks significantly and interbank loans can be interpreted as deposits, which are secured on specific revenue streams. It allows banks that are able to lend profitably to expand their lending and banks without access to such borrowers will provide the additional interbank loans to finance these loans. The use of collateral pyramids can lead to a more efficient allocation of capital as it redistributes the deposits that one bank has received to a more efficient use by another bank.

Reading Boissay & Cooper (2020)

Conclusions

Interbank loans can be used by banks to provide an insurance against liquidity imbalances in the banking system. Banks may face unexpected liquidity shortfalls due to the withdrawal of deposits or expected withdrawals of deposits are not replaced by an inflow of new deposits in the way anticipated. As long as other banks have an excess of cash reserves, they will be happy to provide those banks facing liquidity shortages with funds to temporarily provide this liquidity to banks. Such excess cash reserves might have been accumulated from lower than expected deposit withdrawals or larger than anticipated inflows of deposits.

As banks will provide their excess cash reserves to those banks that need additional liquidity, they are exposing themselves to a possible liquidity shortage if subsequently they are affected by a liquidity shock. While this increases the risk of the bank failing, it is nevertheless optimal to provide other banks with liquidity if they need it, as the bank itself would have to rely on such assistance if faced a liquidity shortage. Hence the provision of interbank loans provides a mutual insurance between banks against liquidity shortages.

Such interbank loans are still given if banks might not be able to repay their loans as the loans they have given are defaulting. However, as the interbank loan rate will take into account the risks the banks take when providing interbank loans, it might be more beneficial for banks to raise cash reserves through the liquidation of assets, if the time frame available to raise liquidity allows so. This will be particularly the case if banks are homogenous and the losses from liquating assets are not very high.

While interbank loans can be used to redistribute liquidity within the banking system, they can also be used to optimise the profits to the bank. The revenue they can generate will increase the profits of the bank and as long as the cash reserves are not depleted such that liquidity risks to the bank increase significantly, providing such loans will be profitable. Similarly, banks might be taking interbank loans in order to invest the proceeds and generate more revenue and profits. Using this anticipated revenue as collateral, banks can extend the availability of interbank loans and generate even more profits to the bank.

We have thus seen that interbank loans provide a mechanism for banks to overcome liquidity imbalances in the banking system, but they can also be used to increase the leverage of banks by accessing ever more loans and making investments. In this context they can be seen as a way to circumvent the limited deposits that a bank has available and allow more investments than otherwise would be possible. In this way, deposits might be redistributed between banks if banks have access to investments, mostly loans, of different qualities, ensuring the most efficient allocation of resources within the economy.

Chapter 17

Repurchase agreements

A repurchase agreement, often referred to as repos, consists of an agreement between two parties in which one party sells the other an asset and agrees to repurchase this asset at a fixed time in the future at a price already agreed. The asset on which the agreement is based is usually a marketable security, often a government bond. Repurchase agreements can be interpreted as a loan provided by the purchaser of the asset to its seller, where the asset serves as collateral. This interpretation stems from the fact that if the bank is not able to repurchase the asset, thus repay the loan, it will forfeit the asset, which the purchaser can sell.

With many repurchase agreements covering only short time periods, we can interpret them as short-term liquidity being provided to the seller of the asset by its purchaser. Banks can use the proceeds of repurchase agreements to cover their own liquidity shortfalls, but also provide further loans if new lending opportunities emerge. Similarly banks holding excess cash reserves can use repurchase agreements to invest some of their cash and obtain profits.

Chapter 17.1 will determine why repurchase agreements are preferable to the alternative to generate cash, banks selling the asset in the free market without an agreement for a later repurchase of the same asset. Once their cash demands have reduced they could repurchase the asset again, even without a prior agreement to do so. Given the short-term nature of many repurchase agreements, they will have to be rolled over if the demand on the cash reserves persist, allowing them to be used for long-term investments. This, however, exposed banks to the risk of such roll overs being denied, leading to a so-called repo run, in analogy to the withdrawal of deposits in a bank run. In chapter 17.2 we will explore under which condition such a repo run could lead to the failure of the bank.

17.1 Financing short-term investments

Banks may have purchased long-term marketable securities such as bonds using their cash reserves at a time when they did not have more profitable alternative investments available; for example, the demand for loans might have been low or companies applying for loans were not creditworthy. Such a situation might change, however, and the bank might want to release the funds invested into the securities to provide additional loans. One way the bank could achieve this, is by selling the security and then investing the proceeds into new loans. Alternatively, banks might use repurchase agreements to obtain additional funds if the requirement for these funds are for a fixed time period only.

Let us now assume that a bank A has the opportunity to provide a short-term loan L, which yields $\pi(1 + r_L) L$ after one time period, where r_L is the loan rate and π the probability with which this loan is repaid. In addition, this bank own a long-term risk-free bond B that generates $(1 + r)^2 B$ in two time periods, where r denotes the risk-free rate and we assume interest accumulates over time periods. Another bank B currently has a cash surplus B, but will require this cash in the coming time period. Failing to obtain this amount of cash will increase their funding costs from emergency loans or penalties imposed on them, causing the bank losses. These losses are equivalent to only obtaining a fraction $\lambda \leq 1$ of the accumulated value of the cash, $\lambda(1 + r)^2 B$. A third bank C would be able to purchase the bond in two time periods with its excess cash reserves.

In order to provide the loan, bank B has purchased the bond from bank A and in the next time period needs to sell the bond to either bank B or bank C to raise the required cash. Let us assume bank B approaches these banks sequentially and we initially consider the second approach after the first approach has not yielded a sale of the bond. The price obtained will be denoted $\hat{P}_1$ and bank B makes profits of $\hat{P}_1 - \lambda(1 + r)^2 B$, the difference of the cash obtained and the value it would obtain if the cash is not raised. The buyer of the bond, regardless of the type of bank, would obtain a bond that yields them $(1 + r)^2 B$ and for which they pay $\hat{P}_1$, giving them profits of $(1 + r)^2 B - \hat{P}_1$. Using Nash bargaining over the price $\hat{P}$, we seek to maximize the objective function

$$\hat{\mathcal{L}}_1 = \left(\hat{P}_1 - \lambda(1 + r)^2 B\right)\left((1 + r)^2 B - \hat{P}_1\right), \tag{17.1}$$

whose first order condition for a maximum, $\frac{\partial \hat{\mathcal{L}}_1}{\partial \hat{P}_1} = 0$, yields

$$\hat{P}_1^* = \frac{1}{2}(1 + \lambda)(1 + r)^2 B \leq (1 + r)^2 B, \tag{17.2}$$

where the last inequality arises from $\lambda \leq 1$ and hence the agreed price will be below the value of the bond.

Knowing the outcome of the second approach, bank B knows that it will receive $\hat{P}_1^*$ as determined in equation (17.2), if declining the offer P_1 from the first approach, the net surplus will be $P_1 - \hat{P}_1^*$. The net surplus of the other bank will remain at $(1 - r)^2 B - P_1$. Hence Nash bargaining seeks to maximize

$$\mathcal{L}_1 = \left(P_1 - \hat{P}_1\right)\left((1 + r_B)^2 B - P_1\right), \tag{17.3}$$

which gives us from the first order condition for a maximum, $\frac{\partial \mathcal{L}_1}{\partial \hat{P}_1} = 0$, after inserting from equation (17.2) for $\hat{P}_1^*$ that

$$P_1 = \frac{1}{2}\left(\hat{P}_1 + (1+r)^2 B\right) = \frac{\lambda+3}{4}(1+r)^2 B. \tag{17.4}$$

We can easily see that $\hat{P}_1 \leq P_1 \leq (1+r_B)^2 B$. Thus bank B would prefer to accept the price from the first approach, P_1, as it is higher. This higher price in the first approach arises from to the existence of the outside option, namely to make a second approach to the other bank.

As the first offer is accepted by bank B, it could thus approach bank A, from which it purchased the bond, to re-sell it the bond after one time period, that is after the short-term loan they have provided has been repaid. This constitutes a repurchase agreement and they could agree the price of this repurchase of the bond, P_1, in advance. The price at which bank B obtains the bond from bank A in the first instance in this repurchase agreement is denoted P_0^R. The net surplus for bank B will be $P_1 - P_0^R$, the price difference between what it paid for the bond, P_0^R, and what it sells it for, P_1. If bank A enters this repurchase agreement, they will be able to invest these proceeds into the loan and obtain $\pi(1+r_L)P_0^R$, having to repurchase the loan at P_1, giving them profits of $\pi(1+r_L)P_0^R - P_1$. Hence the Nash bargaining maximizes the expression

$$\mathcal{L}_0^R = \left(P_1 - P_0^R\right)\left(\pi(1+r_L)P_0^R - P_1\right). \tag{17.5}$$

The first order condition for maximizing the objective function, $\frac{\partial \mathcal{L}_0^R}{\partial P_0^R} = 0$ solves for

$$P_0^R = \frac{\lambda+3}{4}\frac{\pi(1+r_L)-1}{\pi(1+r_L)}(1+r)^2 B \leq P_1 \tag{17.6}$$

when inserting for P_1 from equation (17.4).

Alternatively, bank B may sell the bond to a third party, which here is banks C. In this case bank A will not repurchase the bond, but simply loose its payoff, $(1+r_B)^2 B$ in exchange for obtaining the purchase price P_0 and investing this into loans, giving $\pi(1+r_L)P_0$. Thus, the profits of banks A in this case are $\pi(1+r_L)P_0 - (1+r)^2 B$ and bank B will obtain the price difference between what it paid for the bond, P_0, and what it sells it for, P_1, giving profits $P_1 - P_0$. Using Nash bargaining to determine the price banks B pays bank A, the objective function becomes

$$\mathcal{L}_0^R = (P_1 - P_0)\left(\pi(1+r_L)P_0 - (1+r)^2 B\right), \tag{17.7}$$

from which we obtain from the first order condition of maximizing this expression, $\frac{\partial \mathcal{L}_0^R}{\partial P_0} = 0$, that

$$P_0 = \left(\frac{\lambda + 3}{8} - \frac{1}{2\pi\,(1 + r_L)}\right)(1 + r)^2\,B \leq P_1. \tag{17.8}$$

We easily see that with a repurchase agreement, the price agreed is higher, $P_0^R > P_0$, given that $\lambda \leq 1$. Hence a repurchase agreement allows the seller of the bond, banks A, to obtain a higher price for the bond than a direct sale of the bond to raise cash, making such an arrangement more attractive to the seller of the bond. While a repurchase agreement results in a higher price to be paid for the bond by bank B, it would still be profitable for bank B to offer a repurchase agreement. Banks obtain a higher price in repurchase agreements as they retain a stronger bargaining position; failure to agree the repurchase of the asset without a repurchase agreement, will lead to a larger loss due to the bond and the associated interest not being returned. The fact that the repurchase itself occurs at a lower price than the value of the bond, make the repurchase agreement more valuable to the seller.

The total surplus with a repurchase agreement is given by the surplus of the banks selling the bond, bank A, consisting of the return from investing the initial sale price P_0^R and then repurchasing the bond at P_1, while the purchaser of the bond, banks B, make profits from the difference in the prices at which it sells the bond back to the original seller and the price it purchased it from this bank, $P_1 - P_0^R$. Hence the combined profits are given by

$$\begin{aligned} \Pi_R &= \left(\pi\,(1 + r_L)\,P_0^R - P_1\right) + \left(P_1 - P_0^R\right) \\ &= P_0\,(\pi\,(1 + r_L) - 1)\,, \end{aligned} \tag{17.9}$$

where we obtain the second equality from inserting the expressions in equations (17.4) and (17.6).

Similarly, the surplus when not entering a repurchase agreement is for bank A selling the bond given as the difference between investing the purchase price P_0 into the loans and giving up the claim on the bond, while banks B purchasing the bond has the same profits as before. Using the expressions in equations (17.4) and (17.8), we get the total surplus when no repurchase agreement is in place as

$$\begin{aligned} \Pi &= \left(\pi\,(1 + r_L)\,P_0 - (1 + r)^2\,B\right) + (P_1 - P_0) \\ &= P_0\,(\pi\,(1 + r_L) - 1) - \left((1 + r)^2\,B - P_1\right). \end{aligned} \tag{17.10}$$

As $P_1 \leq (1 + r_B)^2\,B$ from our result in equation (17.4), we see that $\Pi_R \geq \Pi$ and hence repurchase agreements are desirable to selling the bond on to a third party. This is because the higher price paid to the seller of the bond, banks A, allows this bank to invest a larger amount into the loan, which yields a high return and thus increases the overall surplus.

Using equations (17.4) and (17.6), we can get the implied interest rate for this repurchase agreement, the Repo rate, as

$$1 + r_R = \frac{P_1}{P_0^R} = \frac{\pi\,(1 + r_L)}{\pi\,(1 + r_L) - 1}, \tag{17.11}$$

where $\pi\,(1 + r_L) > 1$ to ensure positive surpluses and a viable loan is given. Interestingly, the risk-free rate $1 + r$ is not affecting the Repo rate directly, only the loan conditions of bank A, $\pi\,(1 + r_L)$. In reality, of course these loan conditions, especially the loan rate r_L would be affected by the risk-free rate.

Thus we see that repurchase agreements are preferred over the sale of long-term securities by banks that seek liquidity for a short-term investment opportunity, such as a loan. The advantage of repurchase agreements can be found in the fact that the selling bank obtains a larger price when (temporarily) selling the security, allowing it a larger investment and hence a higher profitability. While this reduces the profits of the bank temporarily purchasing the security, engaging in a repurchase agreement is nevertheless profitable for banks holding excess cash reserves and thus repurchase agreements are entered.

Repurchase agreements are a way to finance short-term loans for which banks do not have the requisite cash reserves, but they hold other less liquid assets. They therefore help with the efficient allocation of resources towards banks with the best investment opportunities.

Reading Tomura (2016)

17.2 Repo runs

Repurchase agreements are often, although not always, short-term arrangements and thus not directly suitable to finance long-term investments. However, similar to deposits are not withdrawn and hence able to finance long-term loans, repurchase agreements can be rolled over and thus are able to finance long-term investments. In the same way that deposits can be withdrawn and cause the banks a liquidity short fall, repurchase agreements might not be rolled over anymore, having a very similar effect. While the sudden withdrawal of deposits is often referred to as a bank run, the failure of repurchase agreements being rolled over is known as a repo run.

Let us now assume that banks seek to provide loans L_t at time t that are to be repaid in two time periods, where banks charge a loan rate of r_L for these two time periods and loans are repaid with probability π. These loans are financed by repurchase agreements that are agreed for a single time period only, but are rolled over with probability p and attract a repo rate of r_R. The reason why a repurchase agreement is not rolled over is exogenous and not based on the risks of the banks involved; it might be the result of the banks purchasing the security facing a liquidity shortage itself. Loans, and the associated repurchase agreements, are provided every in time period.

The profits of the bank at time t from a repurchase agreement are given by the revenue generated from the proceeds by the repurchase agreement used to provide a loan of size B_{t-2} and repaying this loan, provided it has been rolled over. In addition, the repurchase agreement entered the previous time period, B_{t-1}, may not be extended required repayment. We thus have bank profits given by

$$\Pi_B^t = \pi (1 + r_L) B_{t-2} - p (1 + r_R)^2 B_{t-2} - (1 - p)(1 + r_R) B_{t-1}. \qquad (17.12)$$

If the bank were to use their cash reserves rather than a repurchase agreement to provide loans, they would use the amount B_{t-2} to provide a loan and finance this through their cash reserves, which we assume would yield no return to the bank. Thus the profits of the bank would be

$$\hat{\Pi}_B^t = \pi (1 + r_L)^2 B_{t-2} - B_{t-2}. \qquad (17.13)$$

We now require that in equilibrium $\Pi_B^t = \hat{\Pi}_B^t$ such that financing loans directly from the banks own resources or through repurchase agreements yield the same profits. If using repurchase agreements was more profitable than using their own resources, all banks would rely on seeking repurchase agreements and no bank would be willing to provide these. Similarly, if the direct financing of loans was more profitable than using repurchase agreement, no bank would seek repurchase agreements but be willing to offer them.

We now focus our analysis on the steady state in which repurchase agreements are stable over time such that $B_{t-2} = B_{t-1} = B_t = B$. Inserting this steady state into equations (17.12) and (17.13) and setting these equal we obtain $(1 - p)(1 + r_R)^2 + p(1 + r_R) = 1$, for which one solution is

$$1 + r_R = 1. \qquad (17.14)$$

Let us now assume that for exogenous reasons, a repo run occurs such that no repurchase agreements are rolled over. In this case the cash flow of banks changes to a cash deficit, neglecting that any new loans might be given, of

$$\begin{aligned} M &= p (1 + r_R)^2 B + (1 + r_R) B - \pi (1 + r_L)^2 B \\ &= (1 + p - \pi (1 + r_L)) B. \end{aligned} \qquad (17.15)$$

The first term consists the repurchase agreement that was due to be repaid regularly, provided it was rolled over previously and the second term the repurchase agreement taken out in the previous time period and which is now not rolled over. The final term consists of the cash generated from the loan being repaid. The final equality emerges if we insert from equation (17.14) that $1 + r_R = 1$.

If a new repurchase agreement were to be advanced to the bank to cover this cash shortage, it could use the loan provided in the previous time period and due to be repaid in the coming time period, as collateral. Let us assume that a repurchase agreement for a fraction λ of the value of this loan can be obtained using it as collateral. In order to obtain a repurchase agreement large enough to cover the liquidity shortfall M, this collateral has to exceed the repayment of the necessary repurchase agreement, thus $\lambda\pi (1 + r_L) B \geq (1 + r_R) M$. Inserting for the liquidity shortfall from equation (17.15) this condition becomes

$$p \leq \frac{(1 + r_R + \lambda)\,\pi\,(1 + r_L) - (1 + r_R)}{1 + r_R}. \tag{17.16}$$

Hence, if the likelihood of a repurchase agreement to be rolled over is sufficiently low, the bank should able to secure a fully collateralised repurchase agreement and thus overcome its liquidity shortage. A repo run might occur, but the bank can secure a new repurchase agreement, using the outstanding loan as collateral, and thus no adverse effects on the bank are observed.

If the likelihood of repurchase agreements being rolled over is low, the cash shortage will be lower as we can see from equation (17.15). The lower demand for cash from the bank affected by a repo run, will allow them to provide sufficient collateral in order to secure a new repurchase agreement and secure sufficient cash to avoid a failure. Once repurchase agreements are rolled over frequently enough, the cash demands exceed the ability of the bank to provide collateral and it will not obtain a new repurchase agreement; this will lead to the failure of the bank.

We have thus seen that repo runs can occur, but as long as the cash demands of banks are sufficiently low to be covered by loans they have already provided, they can secure new repurchase agreement to avert a liquidity shortage. This is only possible if the repurchase agreements are not too routinely rolled over and thus banks do not rely on them too much for the provision of liquidity. It is only then that banks are able to secure new repurchase agreements to overcome the liquidity shortfall.

Reading Martin, Skeie, & von Thadden (2014)

Conclusions

Repurchase agreements are the preferred way to raise cash, compared to the sale of assets. This is driven by the ability of the bank selling the asset to negotiate a higher price and thus obtain higher proceeds from a repurchase agreement compared to an outright sale of the asset. The loss of the asset in this case makes the bargaining position of the seller weaker as its future revenue from the asset is lost. Hence repurchase agreements are a cost-efficient way banks can generate cash reserves, which they can use for further loans or to cover their own liquidity shortfall, provided they are holding acceptable assets on which to base the repurchase agreement.

With repurchase agreements being preferred to the sale of assets and the often short-term nature of repurchase agreements, banks may rely on these to obtain cash reserves. With the need for repurchase agreements having to be rolled over to ensure the more long-term liquidity needs of banks are met, banks expose themselves to this rollover risk. Similar to deposits being withdrawn at short notice, repurchase agreements might not extended, leading to a repo run. Banks can provide the more long-term investments they have made with the proceeds of the repurchase agreement as collateral to obtain new repurchase agreements and over come the resulting liquidity shortage. This will work as long as the general withdrawal rate of repurchase agreements is not too high and thus the liquidity demands by the bank can be covered by these investments.

Chapter 18

Deposit insurance

Banks provide loans to companies using deposits, but are supposed to repay these deposits, even if the loans they have given, default. Such defaults, if occurring in sufficient numbers, will not allow banks to repay all deposits as they will not have the funds available to do so; in such a case depositors would incur a loss. It is now possible to provide insurance against such losses. If the bank is not able to repay the deposits in full, this insurance would make payments to depositors such that they receive the full value of their deposits.

Such deposit insurance can in principle be provided by any type of insurance company who charges an insurance premium and will make these payments to depositors if needed due to banks not being able to repay deposits in full. In many cases deposit insurance is not provided by a form of common insurance based on insurance premia, but the size of deposits held by banks are often so large that having to make payment would overwhelm any commercial insurance company. For this reason conventional insurance of deposits are rarely found and instead it is the government or central bank that provide the insurance. Such insurance can be explicit and be backed by legislation, or it can be an implicit deposit insurance where government or central banks have made informal commitments to ensure depositors do not face losses. In some instances the deposit insurance might be even more implicit in that no such commitment has been given, but it is seen as politically or economically not feasible to allow a bank to fail and not repay its depositors. The consequence of such implicit deposit insurance is that no insurance premia are raised, although even explicit government guarantees most often do not require the payment of insurance premia.

Providing deposit insurance can change the incentives of banks as depositors will no longer be concerned with the risks that banks are taking. Thus, if, and how, deposit insurance premia are determined can affect bank behaviour. In chapter 18.1 we will see what impact the pricing of deposit insurance has on the behaviour of banks and how setting prices wrong can distort the incentives of banks. We will also consider how deposit insurance should be adequately priced.

Deposit insurance is in most instances not unlimited. We often find that the amount of deposits insured per depositor is limited and that not all depositors are actually covered by the deposit insurance; it is a widespread practice to limit deposit insurance to individual depositors, excluding corporate depositors, sometimes with the exception of small businesses, and limit the amount of deposits that are covered. That such limits are in the interest of banks will be explored in chapter 18.2, addressing both limitations.

Even if deposit insurance is provided free to banks through guarantees by governments, the payments that are made if the deposit insurance is called upon, will have to be funded, usually through general taxation. How to raise the necessary deposit insurance premia, either ex-ante through a conventional insurance scheme or ex-post through taxation is the subject of chapter 18.3. We will analyse whether banks, depositors or general taxation should be used to pay for deposit insurance.

18.1 The pricing of deposit insurance

Deposit insurance is in many instances provided free by governments or central banks, either based on a legal requirement to provide such insurance or an implicit guarantee based on re-assurances made to the public, often as the stability of the banking system is questioned. Of course we can interpret such a case as the deposit insurance having a price of zero to banks. In other instances, however, banks are charged a fee for this provision of deposit insurance.

While deposit insurance protects depositors against any losses the bank may make that would prevent them from causing a bank run due to the possibility of losses if retaining their deposits, its presence might affect the incentives of banks. In chapter 18.1.1 we will investigate how deposit insurance not set at the correct price can affect the risk-taking of banks and in chapter 18.1.2 we explore the optimal pricing of deposit insurance to take these risks into account and reduce the incentives to take on additional risks. Similarly, chapter 18.1.3 will explore how banks make decisions that require bailouts and how deposit insurance can influence such decisions.

18.1.1 Fixed-price deposit insurance

While deposit insurance is often provided free through government schemes, either as an explicit insurance required by law, or as an implicit insurance by either the government or central bank; such implicit guarantees might be inferred to from statements made by government or central bank officials. In other cases, however, banks have to contribute a fixed amount into a fund to finance any payouts from such an insurance scheme. This amount, representing the insurance premium to be paid by banks, will often be fixed for a bank and is not varied, apart from the size of the bank, as commonly measured by the amount of deposits that are to be insured.

Let us initially consider the case where no deposit insurance is provided. In this case, banks use their deposits and their own equity E to finance loans L, such that $L = D + E$. Then, if loans are repaid with probability π, the loan rate is set at r_L

and the deposit rate at r_D, the profits of the bank, taking into account their own investment of equity E, are given by

$$\Pi_B = \pi (1 + r_L) L - (1 + r_D) D - E. \tag{18.1}$$

In competitive markets banks make no profits, hence we have $\Pi_B = 0$, and hence after inserting $L = D + E$, we get the loan rate as

$$1 + r_L = \frac{(1 + r_D) D + E}{\pi (D + E)}. \tag{18.2}$$

If deposit insurance is provided at a fixed deposit insurance premium of P, the bank will obtain a payout from the deposit insurance if the loans are not repaid. In this case the deposit insurance reimburses depositors and banks will make no losses. Hence the bank retains all revenue from loan repayments, after having repaid depositors, if the loans are repaid, in exchange for the deposit insurance premium. Thus their profits are given by

$$\hat{\Pi}_B = \pi ((1 + r_L) L - (1 + r_D) D) - P - E. \tag{18.3}$$

If the bank sets the loan rate competitively as obtained in equation (18.2) and we use that $D = L + E$, the bank profits with deposit insurance are given by

$$\hat{\Pi}_B = (1 - \pi) (1 + r_D) D - P. \tag{18.4}$$

If we define $\kappa = \frac{E}{D}$ as the equity ratio, we can rewrite the bank profits in equation (18.4) as

$$\begin{aligned} \hat{\Pi}_B &= (1 - \pi) (1 + r_D) \frac{1}{\kappa} E - P \qquad (18.5) \\ &= (1 - \pi) (1 + r_D) \frac{1}{1 + \kappa} L - P. \end{aligned}$$

In order to maximize these profits, assuming that the premium P is fixed, we see that banks would seek to minimize the repayment rate of loans, π, thus increasing the risks banks are taking. In addition, banks would seek the lowest possible equity ratio, κ, holding as little equity as possible. By increasing risks, banks benefit from the insurance payout that cover their losses if these risky loans are not repaid, while benefitting from these loans being repaid. Having less equity increases the loan rate as we can see from equation (18.2), increasing profits from the repayment of loans further. Thus having deposits insurance with a fixed premium provides incentives for banks to increase the they take, increasing moral hazard in their decisions.

The deposit insurance premium should take into account this behaviour of banks and it can be set such that bank profits are zero, equivalent to the case of having no deposit insurance, $\hat{\Pi}_B = 0$. This would then give us a deposit insurance premium of

$$P = (1 - \pi) (1 + r_D) D = (1 - \pi) (1 + r_D) \frac{1}{1 + \kappa} L. \tag{18.6}$$

In this case the deposit insurance premium would take into account the risk the bank is taking by requiring a higher premium if the repayment rate of loans reduces, π, or the equity ratio κ reduces. With the profits of banks remaining unchanged as they increase risks, banks have no incentive to do so and the moral hazard from the introduction of deposit insurance at a fixed premium is eliminated.

If deposit insurance charges a fixed premium, including no premium at all as in many government-banked deposit insurance schemes, banks have an incentive to increase their risks and benefit from the insurance payout should they not be able to repay depositors, while obtaining all benefits if they are profitable. Hence if deposit insurance is not priced according to the risks banks take, other regulatory measures are required to limit the risk taking of banks, such as capital requirements.

Readings Furlong & Keeley (1989)

18.1.2 Deposit insurance as a put option

Deposit insurance pays the depositors if the value of the assets of the bank are insufficient to make full payment to all depositors; in this case deposit insurance pays the difference between the claims of depositors, consisting of the deposits D and interest r_D, and the value of loans the bank holds. If loans have a repayment rate of π and banks charge a loan rate r_L, these loans have a value of $\pi\left(1+r_L\right)L$. Hence with deposits repaid at time T, the payment of the deposit insurance is given by

$$P_T = \max\left\{0; \left(1+r_D\right)D - \pi\left(1+r_L\right)L\right\} \tag{18.7}$$

This expression represents the payoff of a put option with maturity of time T, where $\left(1+r_D\right)D$ represents the strike price and $\pi\left(1+r_L\right)L$ the value of the underlying asset, the loans given by the bank. If the payout at time T represents a put option, the value of receiving these payouts prior to this time can be valued as a put option, too.

Let us assume that there is only one time period until deposits have to be repaid, thus $T = 1$ and the risk-free rate is zero, such that cash holdings do not attract any interest. Banks have provided a portfolio of N otherwise identical loans L_i, where $L = NL_i$, each loan with a repayment rate of π and the actual repayments being independent of each other. Variance of this portfolio of loans is then given by $\sigma^2 = N\pi\left(1-\pi\right)L_i^2 = \pi\left(1-\pi\right)\frac{L^2}{N}$.

We can now use option pricing theory to determine the value of this deposit insurance interpreted as the value of a put option. We might use the Black-Scholes valuation of a European put option, which gives us, when using that the risk-free rate is zero and we only consider a single time period, a value of

$$P_0 = \left(1+r_D\right)D\Phi\left(d_2\right) - \pi\left(1+r_L\right)L\Phi\left(d_1\right), \tag{18.8}$$

$$d_1 = \frac{1}{\sigma}\ln\frac{1+r_D}{\pi\left(1+r_L\right)}\frac{D}{L} - \frac{1}{2}\sigma, \tag{18.9}$$

$$d_2 = d_1 + \sigma. \tag{18.10}$$

Here Φ denotes the standard normal distribution function. Standard option pricing theory suggests that the value of this option increases in the risk to the loans, represented by the variance of loans, σ^2, and the leverage $\frac{D}{E}$, where we assume that loans are given using the deposits obtained as well as any equity the bank holds, such that $L = D + E$. It has to be noted that the risk σ will also depend on the repayment rate π and with very low repayment rates, the value of the deposit insurance will be small. Realistically, we assume that $\pi > \frac{1}{2}$ and hence the lower the repayment rate, the higher the variance.

If the value of the deposit insurance is given by the value of this put option, the insurance premium should reflect this value to the bank; the insurance premium will reflect the risks the bank takes through its influence on the variance σ_2. The more loans are available to repay deposits, the lower the payments of the deposit insurance will be, and hence the lower the deposit insurance premium should be. If a bank holds more equity, it will have more loans to repay deposits as $L = D + E$ and hence a lower leverage will reduce this premium. As is obvious from the value of the deposit insurance in equation (18.8), the more deposits are to be insured, the higher the deposit insurance premium will be; this is achieved without having to know the profits of banks to extract any surplus.

We can thus interpret deposit insurance as a put option on the value of the loans the bank has given and determine the insurance premium as the value of this put option. This will allow us to take into account the risks associated with the deposit insurance and charge the bank a fair premium. Being charged a deposit insurance premium that takes into account the risks the bank takes, will reduce any moral hazard that might arise from being able to take on additional risks without having to pay higher deposit rates due to being isolated from these risks because of deposit insurance.

Reading Merton (1977)

18.1.3 The impact of deposit insurance on bailouts

Banks make decisions not only about the risks they are taking, but their individual decisions can have an impact on the social costs of a bank failing. If a single bank fails, the social costs are usually low, but multiple banks failing will impose significantly higher costs. If bank are making decisions that increases the correlation of such failures, they may not increase the risks they are individually taking, but the risks the banking system poses. If deposit insurance takes such risks into account adequately, it may provide a mechanism to internalise these social costs.

Let us assume that banks face a liquidity shortage with probability p and that there are two banks. A bank facing such a liquidity shortage will have to sell their loans in order raise additional cash reserves and sell it either to the other bank if it faces no liquidity shortage or to outside investors. Banks have provided loans L at a loan rate r_L and we assume that loans are repaid with probability π. Outside investors pay a fraction λ of the value of the loan, $\pi(1 + r_L) L$, such that banks obtain $\lambda\pi(1 + r_L) L$, while other banks are willing to pay a fraction $1 > \hat{\lambda} > \lambda$ of the loan value. The purchase price is assumed to be higher as banks are more familiar with the loan portfolio they are purchasing than an outside investor and will therefore be willing to make a better offer.

If the two banks are operating in the same market, we assume that they both face the same liquidity shortage and cannot sell the loans to each other, thus obtaining $\lambda\pi\,(1 + r_L)\,L$ from an outside investor. If the total deposits D that are withdrawn are attracting interest r_D, the total amount that is withdrawn is $(1 + r_D)\,D - \lambda\pi\,(1 + r_L)\,L$, which would have to be covered by deposit insurance. As this liquidity shortage occurs with probability p, the expected payment the deposit insurance has to make is

$$P = p\,((1 + r_D)\,D - \lambda\pi\,(1 + r_L)\,L)\,, \tag{18.11}$$

where we assume that $(1 + r_D)\,D > \lambda\pi\,(1 + r_L)\,L$ such that selling loans does not cover the liquidity required. If the market for deposit insurance is competitive, this would be the premium the bank has to pay for its deposit insurance.

If banks are operating in different markets, we assume that liquidity shortages of banks are independent of each other and if only one bank faces a liquidity shortage, the other will be able to purchase the loans it has to sell, in which case they obtain $\hat{\lambda}\pi\,(1 + r_L)\,L$. The deposit insurance thus has to pay the liquidity shortfall after selling to an outside investor if both banks fail and after selling to the other bank if only one of them fails. Thus the deposit insurance pays the amount of

$$\begin{aligned} \hat{P} &= p^2\,((1 + r_D)\,D - \lambda\pi\,(1 + r_L)\,L) \\ &\quad + p\,(1 - p)\left((1 + r_D)\,D - \hat{\lambda}\pi\,(1 + r_L)\,L\right) \\ &= P - p\,(1 - p)\left(\hat{\lambda} - \lambda\right)\pi\,(1 + r_L)\,L. \end{aligned} \tag{18.12}$$

As the other bank is willing to pay a higher price for the loans that have to be sold, the insurance payout will be reduced. We again assume that the market for deposit insurance is competitive and the insurance premium is identical to these expected payments.

If only one bank fails, the bank selling their loans will make a loss as they are selling these below their value, but the bank purchasing these loans will make a profit of the same size, thus the position of the banking system as whole is unchanged. As one bank will be surviving we assume that there are no costs associated in liquidating one of the banks and a bailout of the failing bank is not required.

If both banks fail, the loans are sold to outside investors imposing a loss of $(1 - \lambda)\,\pi\,(1 + r_L)\,L$ on the bank that is not recovered within the banking sector. If we further assume that outside investors do not value these loans more highly due to their unfamiliarity with the loan market, this imposes social costs of this magnitude. As both banks fail, the lack of banks will impose social costs of C. Bailing out a bank would necessitate to recapitalise it with its looses of $(1 - \lambda)\,\pi\,(1 + r_L)\,L$. Hence a bailout would be desirable if $C \leq (1 - \lambda)\,\pi\,(1 + r_L)\,L$.

Let us now assume that the deposit insurance takes into account these social costs and the premium reflects this accurately. We thus have for the insurance premia if banks operate in the same an different market

$$\begin{aligned} P^* &= p\left((1+r_D)\,D - \lambda\pi\,(1+r_L)\,L \right. \\ &\quad \left. + \min\{C; (1-\lambda)\,\pi\,(1+r_L)\,L\}\right), \\ \hat{P}^* &= p^2\left((1+r_D)\,D - \lambda\pi\,(1+r_L)\,L \right. \\ &\quad \left. + \min\{C; (1-\lambda)\,\pi\,(1+r_L)\,L\}\right) \\ &\quad + p\,(1-p)\left((1+r_D)\,D - \hat{\lambda}\pi\,(1+r_L)\,L\right). \end{aligned} \tag{18.13}$$

If banks are bailed out, they are fully recapitalised and they continue their operation, making profits of $\Pi_B^* = \pi\,(1+r_L) - (1+r_D)\,D$ and if banks are not bailed out, they fail and $\Pi_B^* = 0$. Hence bank profits, if the two banks are operating in the same market, are given by the profits it makes if no liquidity shock occurs and if a liquidity shock occurs the bank will obtain Π_B^*. Thus we have

$$\Pi_B = (1-p)\,(\pi\,(1+r_L)\,L - (1+r_D)\,D) + p\Pi_B^* - P^*. \tag{18.14}$$

If banks operate in different markets, they will obtain their operating profits if both banks do not face a liquidity shock and if one bank faces a liquidity shock, they will obtain their operating profits if it is the other bank facing this shock and they will obtain profits from the purchase of the loans at a discount, which is financed by additional deposits. If they are the only bank facing the liquidity event, they will be liquidated and obtain no profits. If both banks face a liquidity shortage, they only make profits if they are bailed out. We thus have the bank profits given as

$$\begin{aligned} \hat{\Pi}_B &= (1-p)^2\,(\pi\,(1+r_L)\,L - (1+r_D)\,D) \\ &\quad + p\,(1-p)\,(\pi\,(1+r_L)\,L - (1+r_D)\,D \\ &\quad + \left(\pi\,(1+r_L) - (1+r_D)\,\hat{\lambda}\pi\,(1+r_L)\,L\right)\big) \\ &\quad + p^2\Pi_B^* - \hat{P}^*. \end{aligned} \tag{18.15}$$

Banks operate in different markets if it is more profitable to do so, hence we need $\hat{\Pi}_B \geq \Pi_B$, which solves for

$$\begin{aligned} P^* - \hat{P}^* \geq p\,(1-p)\left(\Pi_B^* - (\pi\,(1+r_L)\,L \right. \\ \left. - (1+r_D)\,\hat{\lambda}\pi\,(1+r_L)\,L)\right). \end{aligned} \tag{18.16}$$

We can now see that if $C > (1-\lambda)\,\pi\,(1+r_L)\,L$, then no bail out happens and hence $\Pi_B^* = 0$. As in this case $P - \hat{P} > 0$ and $\pi\,(1+r_L)\,L - (1+r_D)\,\hat{\lambda}\pi\,(1+r_L)\,L > 0$, this condition is fulfilled. Hence, if bailouts cannot occur, banks seek to minimize insurance premia and obtain additional profits from buying assets by covering different markets.

If $C < (1-\lambda)\,\pi\,(1+r_L)\,L$ and hence a bailout happens such that $\Pi_B^* = \pi\,(1+r_L)\,L - (1+r_D)\,D$, we get from inserting this value into (18.16) and using the expressions for the deposit insurance from equation (18.13), that for $\Pi_B \leq \hat{\Pi}_B$ we need

$$C \leq (1 + r_D) D - \left(r_D \hat{\lambda} + \lambda\right) \pi (1 + r_L) L. \tag{18.17}$$

Thus if the cost of a bailout, C, are sufficiently small, banks may operate in different markets. Banks here exploit the possibility to obtain a bailout and if this is sufficiently likely to happen, they will operate in the same markets. If the bailout costs are high, the higher costs of the deposit insurance induces them to operate in different markets.

In the case that deposit insurance is provided for free, $P^* = \hat{P}^* = 0$, or at a fixed price that does not reflect the risks banks are taking, $P^* = \hat{P}^*$, we see from equation (18.16) that we require $\hat{\lambda}\pi (1 + r_L) L \leq D$ if inserting for Π_B^*, thus if the purchase price of the assets is sufficiently low, banks operate in different markets. With fairly priced deposit insurance, as obtained in equation (18.13), the right-hand side of the condition in equation (18.16) becomes positive and hence the condition is less restrictive, meaning that for a wider range of parameters banks will operate in different markets.

If banks operate in different markets bailouts happen with probability p^2, while if bank operate in the same markets they occur with probability p, which is higher and thus imposes higher social costs and thus it is socially optimal for banks to operate in different markets. If bailouts of banks can be ruled out, banks optimally choose to operate in different markets. However, if bailouts can happen, the deposit insurance has an influence on the choice of banks. Pricing deposit insurance accurately makes it more likely that banks operate in different markets and minimize social costs. If the risks of the banks' choices are not fully taken into account and deposit insurance is not priced accurately, banks may make decisions that require bailouts more frequently. The pricing of deposit insurance can therefore affect decisions of banks that impose social costs due to their possible failure.

The pricing of deposit insurance should take into account the social costs of banks failing and if doing so, it can be used to provide incentives to banks such that they avoid making decisions that impose such high social costs. For example they might deliberately make decisions to enter different markets to existing banks such that risks in the banking system are better diversified. While deposit insurance premia cannot achieve this aim completely, it can provides incentives that make such decisions more likely.

Reading Acharya, Santos, & Yorulmazer (2010)

Résumé

We have seen that if the price of deposit insurance is not taking into account the risks banks are taking, it will provide then with a strong incentive to increase the risk by providing more risky loans and reducing the amount of equity they put at risk. As any losses the bank makes are covered by the deposit insurance, banks will never make losses, but obtain any profits if bank makes. With higher risks these profits are bigger and if the bank engages less equity, the loss of equity will also be lower. Banks may also make decisions that require bailouts more often if the deposit insurance premium does not adequately take into account the risks their decisions impose.

Such risks might not be risks banks themselves take, but which are imposed on the banking sector as a whole and impose additional social costs. If deposit insurance takes into account any such social costs, it can be used as a tool to induce bank to take decisions that reduce social costs.

It is thus important that deposit insurance is offered at a price that fully reflects the risks the bank is taking; that way the moral hazard which results in the bank taking higher risk can be reduced or eliminated. We can interpret deposit insurance as a put option on the value of the loans a bank has provided and determine the insurance premium accordingly.

18.2 Limits to deposit insurance coverage

Deposit insurance does in most cases not cover all deposits. It is quite common for certain deposits to be completely excluded and for others to impose a limit on how much deposits are insured. Deposit insurance normally only extends to the deposits made by individuals and not companies and other organisations, although sometimes small businesses are included in the deposit insurance scheme. The aim of deposit insurance here is to protect individuals from bank failure in the assumption that they are not reliably able to assess the risks of banks, while companies are deemed to be able to make such assessments. But even individual depositors are not protected for deposits of any size; usually only deposits up to a certain amount are protected and any deposits in excess of this coverage limit will be unprotected. The argument used for such limits is similar as for the exclusion of companies from deposit insurance, namely that wealthy individuals should be able to make their own assessment of the risks a bank might pose.

While limits to deposit coverage are often imposed by regulators, we will explore here who such limitations might be optimal for banks as well. In chapter 18.2.1 we will investigate whether a deposit insurance is optimal to provide full or partial coverage of deposits, and chapter 18.2.2 then explores how much of their deposits should be covered.

18.2.1 The optimality of deposit insurance limits

In many cases the amount of deposits insured is limited. Such a limit is typically applied to the deposits of each individual at a single and those that have larger deposits will either have to divide their deposits between a number of banks or any deposits in excess of the deposit insurance limit will not be covered in case the bank is not able to repay them in full. Of course, if deposit rates at one bank are more attractive than at other banks, it might be optimal for depositors to retain all deposits at a single bank. Thus banks will compete for these large deposits and will have to decide whether it is actually optimal for them to limit their deposit insurance.

Let us assume that deposit insurance is available to a single depositor up to the amount of D. We further assume that we have two types of depositors, one type deposits an amount of D with a single bank, which is thus covered by its deposit insurance. The other type of depositor has deposits of $2D$ available, which they can deposit with a single bank, where only the amount of D would be covered by deposit insurance, or they divide the deposit up by providing deposits of D each to two banks and are thus fully covered. A fraction λ of depositors are able to make large deposits of $2D$ and a fraction $1-\lambda$ make small deposits of D. The deposit insurance here is not provided by banks, but it can best be described as a government guarantee for which no deposit insurance premium is charged; from the bank's perspective, deposit insurance is free.

We consider two banks who offer differentiated banking services to depositors. Such difference might be in the range or type of services they offer, for example the availability and ease of use of online banking facilities but also access to cash and a branch network. Using the Hotelling model, we assume that these two banks are located at a distance of 1 along a straight line, which will represent the preferences of depositors. A depositor will have distance $0 \leq d_j \leq 1$ from bank j. We can interpret d_i as the location of the depositor relative to bank i and the distance to this bank imposes costs onto depositors; a distance of 1 to a bank would imply costs of c, such that the costs at distance d_j are given by cd_j. Hence if a bank having its deposits at bank i, moving deposits to bank j will result in additional costs of cd_j.

We can now investigate the competition for depositors between these two banks and will consider a situation where no deposit insurance is offered, deposit insurance covers the full amount of deposits, including the large deposits of $2D$, and then we will look into the case where only deposits up the amount of D are covered by deposit insurance.

No deposit insurance Let us assume a depositor is currently having their deposits with bank j, which pays a deposit rate of r_D^j. Banks invest these deposits fully into loans on which interest r_L is payable and these loans are repaid with probability π. Thus, deposits cannot be repaid with probability $1-\pi$. Hence with $\hat{D} = D$ for small depositors and $\hat{D} = 2D$ for large depositors, we get the repayments to depositors when staying with their current bank and moving to the other bank, bank j as

$$\begin{aligned} \Pi_D^{jj} &= \pi\left(1+r_D^j\right)\hat{D} - \hat{D} - (1-\pi)\hat{D}, \\ \Pi_D^{ji} &= \pi\left(1+r_D^i\right)\hat{D} - \hat{D} - (1-\pi)\hat{D} - cd_i. \end{aligned} \tag{18.18}$$

The depositor would move to bank j if this is more profitable. Requiring that $\Pi_D^{ij} \geq \Pi_D^{ii}$ will give us that a depositor will move to bank j if

$$d_i \leq d_i^* = \pi\frac{\left(1+r_D^i\right) - \left(1+r_D^j\right)}{c}\hat{D}, \tag{18.19}$$

where we assume that the constraint that $0 \leq d_i \leq 1$ is fulfilled. As this condition is fulfilled independent of the size of the size of the deposit or how much of its deposit is moved to the other bank, large depositors would move their entire deposits rather than dividing the deposit between banks. Any depositor that is closer than d_i^* to this bank will switch their deposits.

Thus the total deposits for bank i, assuming it charges the higher deposit rate, will consist of the large and small deposits it currently holds, $2\lambda D$ and $(1-\lambda)\,D$, as well as the new deposits that have been switched from bank j by those who are closer enough to bank j. Thus the total deposits are given by

$$\begin{aligned} D_i &= \lambda\left(1 + 2\pi\frac{\left(1+r_D^j\right) - \left(1+r_D^i\right)}{c}\right)2D \\ &\quad + (1-\lambda)\left(1 + \pi\frac{\left(1+r_D^j\right) - \left(1+r_D^i\right)}{c}D\right)D \\ &= (1+\lambda)\,D + \pi\,(1+3\lambda)\,\frac{\left(1+r_D^j\right) - \left(1+r_D^i\right)}{c}D^2, \end{aligned} \tag{18.20}$$

where we used in the first equation the distance d_i^* with the respective large and small deposit size.

As banks invest their deposits fully into loans, their profits are given by

$$\Pi_B^i = \pi\left((1+r_L) - \left(1+r_D^i\right)\right)D_i. \tag{18.21}$$

This allows us to obtain the deposit rate that maximizes these profits by solving the first order condition $\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} = 0$ after inserting for the deposits from (18.20) as

$$1 + r_D^* = (1+r_L) - \frac{1-\lambda}{\pi\,(1+3\lambda)}\frac{c}{D}. \tag{18.22}$$

If we restrict ourselves to symmetric equilibria, where $r_D^i = r_D^j$, we can easily show that for bank j, who is loosing these deposits and hence $d_j^* < 0$, the same first condition emerges.

Inserting the deposit rate from equation (18.21) back into the profits of the bank in equation (18.21), we easily get that bank profits are given by

$$\Pi_B^* = \frac{(1+\lambda)^2}{1+3\lambda}D. \tag{18.23}$$

Having obtained the profits of banks in the absence of deposit insurance, we can now explore how deposit insurance affects bank profits.

Full deposit coverage If deposit insurance covers the full amount of deposits, including large deposits, then deposits are always repaid and the profits for depositors not switching banks and switching banks, respectively, are given by

$$\Pi_D^{jj} = \left(1 + r_D^j\right) \hat{D} - \hat{D}, \tag{18.24}$$
$$\Pi_D^{ji} = \left(1 + r_D^i\right) \hat{D} - \hat{D} - cd_i.$$

The depositor would move to bank j if this is more profitable. Requiring that $\Pi_D^{ij} \geq \Pi_D^{ii}$ will give us that a depositor will move to bank j if

$$d_i \leq d_i^{**} = \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{c} \hat{D}. \tag{18.25}$$

Following similar steps to the case of no deposit insurance, we get the total amount of deposits at bank i, comparable to equation (18.20), as

$$D_i = (1 + \lambda) D + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{c} (1 + 3\lambda) D^2. \tag{18.26}$$

Maximizing bank profits in the same way as without deposit insurance gives the optimal deposit rate as

$$1 + r_D^{**} = (1 + r_L) - \frac{1 + \lambda}{3 + \lambda} \frac{c}{D} \tag{18.27}$$

and inserting this deposit rate into equation (18.21) we obtain the bank profits as

$$\Pi_B^{**} = \pi \frac{(1 + \lambda)^2}{1 + 3\lambda} D = \pi \Pi_B^*. \tag{18.28}$$

Hence with deposit insurance covering all deposits, banks make lower profits and would thus prefer that no deposit insurance is provided. This result emerges because the increased competition in the absence of risk to depositors, reduces the banks' profits; the absence of risk, π, will increase the scope for depositors to switch banks as $d_i^{**} > d_i^*$, thus more depositors would switch, given a deposit rate, increasing the competition for these depositors, which will affect the profits of banks.

While the deposit insurance will increase competition for depositors which reduces profits, the deposit insurance makes deposits risk-free for depositors, allowing banks to reduce deposit rates, increasing their profits. The former effect dominates here, making full deposit insurance less attractive than no deposit insurance.

We can now explore how deposit insurance only covering deposits up to the size of D will affect competition between banks.

Partial deposit coverage Let us now assume that deposit insurance would only cover deposits of size D, and any large deposit of $2D$ would only be covered up to that amount and the remainder might be lost if the bank is not able to repay its deposits. Hence large depositors obtain when staying at bank j, switching the amount of D to bank i, and switching the full amount to bank i, respectively, as

$$\begin{aligned}
\Pi_D^{jj} &= \left(1 + r_D^j\right) D + \pi \left(1 + r_D^j\right) D - (1 - \pi) D \\
\Pi_D^{jij} &= \left(1 + r_D^j\right) D + \left(1 + r_D^i\right) D - cd_j \\
\Pi_D^{ji} &= \left(1 + r_D^i\right) D + \pi \left(1 + r_d^i\right) D - (1 - \pi) D - cd_j
\end{aligned} \tag{18.29}$$

Considering the case that a large depositor would move the amount of D to the other bank. They would do so if $\Pi_D^{jij} \geq \Pi_D^{jj}$, from which we obtain

$$d_i \leq d_i^{***} = \frac{\left(1 + r_D^j\right) - \pi \left(1 + r_D^i\right) + (1 - \pi)}{c} D. \tag{18.30}$$

Moving the full amount of deposits to

Banks will now gain some deposits from their competitor as large deposits are moved to them, but will also lose some deposits from large depositors moving the amount of D to the other bank. As small depositors are fully insured, they behave like in the case of full deposit insurance. We get the deposits of bank i as

$$\begin{aligned}
D_i &= \lambda \left(2D - \frac{\left(1 + r_D^j\right) - \pi \left(1 + r_D^i\right) + (1 - \pi)}{c} D^2 \right. \\
&\quad \left. + \frac{\left(1 + r_D^i\right) - \pi \left(1 + r_D^j\right) + (1 - \pi)}{c} D^2 \right) \\
&\quad + (1 - \lambda) \left(D + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{c} D^2 \right) \\
&= (1 + \lambda) D - (1 + \lambda\pi) \left(\left(1 + r_D^j\right) - \left(1 + r_D^i\right) \right) \frac{D^2}{c}.
\end{aligned} \tag{18.31}$$

Here the first term denotes the large depositors, with losses to the other bank and then gains from the other bank and the second term denotes the effect of the small depositors and the gains made from them.

Inserting this expression into the profits of the bank in equation (18.21) and maximizing profits by choosing the optimal deposit rate, we obtain, again using only symmetric equilibria, that

$$1 + r_D^{***} = (1 + r_L) - \frac{1 + \lambda}{1 + \pi\lambda} \frac{c}{D} \tag{18.32}$$

and hence when inserting this expression back into the profits of the banks in equation (18.21), we obtain

$$\Pi_B^{***} = \pi \frac{(1 + \lambda)^2}{1 + \pi\lambda} D. \tag{18.33}$$

We can now easily see that $\Pi_B^{***} > \Pi_B^*$ if $\pi > \frac{1}{1+2\lambda}$ and hence the bank profits for deposit insurance covering deposits partially are highest and banks whose loan repayment rate is sufficiently high would prefer would prefer such an arrangement. In all cases full deposit coverage is the least favoured arrangement.

The partial cover through deposit insurance increases competition for deposits that large depositors will partially switch to another bank to benefit from the deposit insurance. This will reduce the profits of banks. But on the other hand, deposits become risk-free for depositors and hence banks can pay lower deposit rates, which will increase profits. This latter effect will dominate and increase the profits of banks, making this arrangement of partial deposit insurance optimal for banks.

Summary Banks prefer the provision of deposit insurance that covers deposits up to a certain level only. This allows banks to pay lower deposit rates due to small deposits becoming risk-free, while not increasing competition between banks to such an extent that this advantage is fully eroded. We here assumed that deposit insurance was provided free, such a through government guarantees for which no charge is made, and hence there was no need to consider the impact any deposit insurance premia might have on the profits of banks. It is thus that banks will be content with governments only providing deposit insurance to smaller depositors and would not advocate that deposit insurance is made available more widely.

Reading Shy, Stenbacka, & Yankov (2016)

18.2.2 Optimal coverage limits

A bank obtains deposits from a variety of sources, private individuals as well as companies and other institutional depositors. Deposit insurance is usually only extended to private individuals and not corporate of institutional depositors and hence we can distinguish these two types of deposits by whether they are covered by deposit insurance or not. Banks may freely determine the definition of deposits that are covered by the deposit insurance scheme and thus can ascertain how much of their deposits are actually covered.

Let us assume that the total deposits D of a bank consist of insured deposits, D_I, and uninsured deposits, D_U, where obviously $D = D_I + D_U$; the interest paid on these deposits are r_D^I and r_D^U, respectively. Banks use their deposits to provide loans, on which they charge a loan rate r_L, and which are repaid to the bank with probability π. Banks obtain deposit insurance on which they pay a premium P, which is paid up-front such that the amount that banks can lend out is given by $L = D - P$.

Insured deposits are always repaid to depositors, either by the bank directly or, if they are not able to do so, by the deposit insurance. In contrast, uninsured deposits can only be repaid if sufficient funds are available at the bank; if the loans the bank has provided are not repaid, then no funds are available and hence uninsured deposits receive no payout. If the loan is repaid, however, uninsured deposits are repaid in full if the amount received from the loans exceeds the amount due to depositors, thus we require $(1+r_L)L \geq \left(1+r_D^I\right)D_I + \left(1+r_D^U\right)D_U$ for deposits to be repaid in full to all depositors.

In this case, the uninsured depositors obtain their deposits with probability π, while the insured depositors are always repaid due to the existence of deposit insurance. In equilibrium, the expected returns of uninsured and insured deposits has to be equal in order to avoid depositors switching between insured and uninsured deposits. Thus we require $\pi\left(1+r_D^U\right) = 1+r_D^I$, from which we easily obtain that the deposit rate on uninsured deposits is given by

$$1+r_D^U = \frac{1+r_D^I}{\pi}. \tag{18.34}$$

If on the other hand $(1+r_L)L < \left(1+r_D^I\right)D_I + \left(1+r_D^U\right)D_U$, deposits cannot be repaid fully, but we assume that all deposits are of equal priority and hence repaid proportionally such that uninsured deposits obtain a fraction $\frac{\left(1+r_D^U\right)D_U}{\left(1+r_D^I\right)D_I+\left(1+r_D^U\right)D_U}$ of the complete proceeds the bank receives. Again, the expected return of uninsured and insured deposits have to be equal, $\frac{\left(1+r_D^U\right)D_U}{\left(1+r_D^I\right)D_I+\left(1+r_D^U\right)D_U}\pi(1+r_L)L = \left(1+r_D^I\right)D_U$, which gives us a deposit rate for uninsured deposits of

$$1+r_D^U = \frac{\left(1+r_D^I\right)^2 D_I}{\pi(1+r_L)L - \left(1+r_D^I\right)D_U} \geq \frac{1+r_D^I}{\pi} \tag{18.35}$$

For a viable solution we require $\pi(1+r_L)L - \left(1+r_D^I\right)D_U > 0$, which also implies that if uninsured deposits obtain all proceeds from the loans banks have provided, they must earn at least the return of insured deposits. Using that $D = D_I + D_U$ and $L = D - P$, this becomes

$$\begin{aligned} D_I = D - D_U &> D - \frac{\pi(1+r_L)}{1+r_D^I}(D-P) \\ &= \frac{\pi(1+r_L)}{1+r_D^I}P - \frac{\pi(1+r_L) - \left(1+r_D^I\right)}{1+r_D^I}D. \end{aligned} \tag{18.36}$$

In a competitive market for deposit insurance, the premium will cover the expected payout of the insurance. If the loans are repaid, the deposit insurance has to repay the difference between the claim of the insured deposits, $\left(1+r_D^I\right)D_I$, and the repayments received from the bank directly. If $(1+r_L)L \geq \left(1+r_D^I\right)D_I + \left(1+r_D^U\right)D_U$, then the bank is able to repay all deposits and the deposit insurance has nothing to pay, but if $(1+r_L)L < \left(1+r_D^I\right)D_I + \left(1+r_D^U\right)D_U$ the bank allocates a fraction $\frac{\left(1+r_D^I\right)D_I}{\left(1+r_D^I\right)D_I+\left(1+r_D^U\right)D_U}L$ of the proceeds $(1+r_L)L$. If the loan is not repaid, the deposit insurance has to repay all insured deposits. Hence the expected payout and thus the deposit insurance premium are given by

$$P = \pi \max \left\{0; \left(1 + r_D^I\right) D_I - \left(1 - \frac{\left(1 + r_D^U\right) D_U}{\left(1 + r_D^I\right) D_I + \left(1 + r_D^U\right) D_U}\right) (1 + r_L) L\right\} + (1 - \pi) \left(1 + r_D^I\right) D_I. \tag{18.37}$$

We can easily see that for $(1 + r_L) L \leq \left(1 + r_D^I\right) D_I + \left(1 + r_D^U\right) D_U$ the first term vanishes and hence

$$P = (1 - \pi) \left(1 + r_D^I\right) D_I. \tag{18.38}$$

Inserting for the deposit rate of uninsured deposits from equation (18.34) and the deposit insurance premium from equation (18.38) into this constraint, while noting that $L = D - P$ and $D = D_I + D_U$, this solves for

$$D_I \leq \frac{\pi (1 + r_L) - \left(1 + r_D^I\right)}{(1 - \pi) \left(1 + r_D^I\right) (\pi (1 + r_L) - 1)} D. \tag{18.39}$$

In the case that $(1 + r_L) L < \left(1 + r_D^I\right) D_I + \left(1 + r_D^U\right) D_U$, we can insert for the deposit rate of uninsured deposits from equation (18.35) into the deposit insurance premium from equation (18.37) and obtain

$$P = \left(1 + r_D^I\right) D - \pi (1 + r_L) L, \tag{18.40}$$

which using that $L = D - P$ becomes

$$P = \frac{\pi (1 + r_L) - \left(1 + r_D^I\right)}{\pi (1 + r_L) - 1} D \tag{18.41}$$

Inserting this deposit insurance premium and the deposit rate for uninsured deposits from equation (18.35) into the constraint, we obtain that

$$(1 - \pi) \left(1 + r_D^I\right) D_I < \frac{1 + r_D^I}{\pi (1 + r_L) - 1} (\pi r_L - 1) D. \tag{18.42}$$

We can now easily see that this condition cannot be fulfilled; the right-hand side will be negative and if we assume that $\pi (1 + r_L) > 1$, then either $D_I < 0$ or $D < 0$, which is impossible. Thus we find that for $(1 + r_L) L < \left(1 + r_D^I\right) D_I + \left(1 + r_D^U\right) D_U$ deposit insurance is not supplied and we focus on the case that $(1 + r_L) L \geq \left(1 + r_D^I\right) D_I + \left(1 + r_D^U\right) D_U$.

In the case that $(1 + r_L) L \geq \left(1 + r_D^I\right) D_I + \left(1 + r_D^U\right) D_U$, the bank profits are given by

$$\begin{aligned}\Pi_B &= \pi\left(1+r_L\right)L-\left(1+r_D^I\right)D_I-\left(1+r_D^U\right)D_U \\ &= \left(\pi\left(1+r_L\right)-\frac{1+r_D^I}{\pi}\right)D \\ &\quad+\frac{\left(1-\pi\right)\left(1+r_D^I\right)}{\pi}\left(1-\pi^2\left(1+r_L\right)\right)D_I\end{aligned} \tag{18.43}$$

when inserting from equation (18.34) for the deposit rate of uninsured deposits and equation (18.38) for the deposit insurance, which is paid upfront and reduces lending through $L = D - P$. We see that the bank's profits are increasing in D_I if $\pi^2\left(1+r_L\right) < 1$ and decreasing otherwise. Thus, using the constraint on insured deposits from equation (18.39), we get

$$D_I = \begin{cases} \frac{\pi(1+r_L)-\left(1+r_D^I\right)}{(1-\pi)\left(1+r_D^I\right)(\pi(1+r_L)-1)} & \text{if } \pi^2 < \frac{1}{1+r_L} \\ 0 & \text{if } \pi^2 > \frac{1}{1+r_L} \end{cases}. \tag{18.44}$$

In the case that $\pi^2 = \frac{1}{1+r_L}$, any $D_I \in \left[0; \frac{\pi(1+r_L)-\left(1+r_D^I\right)}{(1-\pi)\left(1+r_D^I\right)(\pi(1+r_L)-1)}\right]$ would be optimal for the bank. Hence if the bank provides loans with high repayment rates, π, thus low-risk loans, the deposit insurance would not be used as the increased costs of uninsured deposits are less than the costs of the insurance premium. Banks taking higher risks through providing loans with lower repayments rates will seek to insure deposits to the maximum feasible, subject to $D_I \leq D$. They will seek such deposit insurance because the cost of insurance is lower than the higher costs of uninsured deposits through higher deposit rates.

Hence we see that low-risk banks do not seek to insure deposits as the low risk of not repaying them ensures a low deposit rate and obtaining deposit insurance is more expensive than providing this risk premium on uninsured deposits. For more risky banks, the savings on the deposit rate when providing insurance outweighs these insurance costs. In general, deposit insurance will not cover the full amount of deposits, $D_I < D$, as long as the deposit rate on uninsured deposits is not too low. This limit on insured deposits is such that deposit insurance is available and the possible payout they have to make not too high, given the possible loan repayments. As deposit insurance is paid upfront it reduces the provision of loans and hence the revenue available to repay depositors.

While low-risk banks may not seek deposit insurance, those banks which provide loans of higher risk will want to insure their deposits as the lower deposit rate increases their profits. We can therefore expect banks with higher risk borrowers to be supportive of deposit insurance schemes and that its extent is as wide as possible, while less risky banks will not be concerned bout the existence of such a scheme.

Reading Dreyfus, Saunders, & Allen (1994)

Résumé

Deposit insurance is typically not provided to all depositors and banks would not find it optimal if coverage would be extended to all deposits. The increased competition between banks for deposits increases if risks are eliminated, while at the same time allowing banks to reduce deposit rates due to the lack of risks for depositors. Providing a limit on the amount of deposits that are covered by deposit insurance, limits competition between banks, while they still enjoy some reduction in deposit rates, making this arrangement optimal for banks. It is therefore that banks are not seeking to extend the amount of coverage that deposit insurance provides.

Not only is the amount of coverage deposit insurance provides for each depositor limited, but in most cases only individual depositors, and maybe small businesses, are given cover at all; other depositors remain uninsured, even if their deposits are below the coverage limit. Such an arrangement is also optimal as it limits the payments the deposit insurance has to make in case of the bank failing to repay deposits and hence the deposit insurance. This preserves the profitability of banks. Actually, banks that only take very low risks would prefer to not provide any deposit insurance as its costs outweigh its benefits, while more risky banks would benefit from the lower deposit rate they can offer, outweighing the costs of deposit insurance.

18.3 The financing of deposit insurance

In many cases it is assumed that deposit insurance is paid for by the bank and they are charged a premium by the provider of the deposit insurance. In other cases it is assumed that deposit insurance is provided by government guarantee without banks being charged. The way deposit insurance is financed can affect the incentives of banks, and of any government providing such deposit insurance, hence it is important to assess who should provide the funds for such deposit insurance.

Let us assume that a company obtaining a loan L at interest r_L can choose between two investments, one has a success rate of π_H and is considered low risk, and the other investment has a success rate of $\pi_L < \pi_H$ and is thus considered risky. The incentives are such that companies would choose the more risky investment if they could make their decision freely, but we assume that the bank can spent an amount of cL and through them monitoring the company can ensure they choose the low-risk investment with a success rate π_H. Banks use deposits D and their own equity E to provide loans. We denote by κ the leverage of a bank and define $\kappa = \frac{L}{E}$ and hence $L = D + E = \kappa E$ and $D = (\kappa - 1) E$.

Depositors can decide to invest their wealth W into deposits, such that $D = \rho W$, or they provide a loan directly to companies to the amount of $(1 - \rho) W$, but they are not able to monitor companies and they will therefore conduct the risky investment by choosing the success rate π_L.

We now establish as a benchmark the optimal decisions of banks and depositors in the absence of a deposit insurance, before then considering the optimal way such a deposit insurance should be financed.

No deposit insurance With limited liability for banks, they will only repay deposits if the loans they have provided are repaid. If banks do not monitor companies, they choose the high-risk investment such that bank profits are given by

$$\begin{aligned}\Pi_B^L &= \pi_L\left(\left(1+r_L\right)L-\left(1+r_D\right)D\right) \\ &= \pi_L\left(\left(1+r_L\right)+\left(r_L-r_D\right)\left(\kappa-1\right)\right)E,\end{aligned} \tag{18.45}$$

where r_D denotes the deposit rate and we used that $L=\kappa E$ and $D=\left(\kappa-1\right)E$. If the bank monitors the company it will choose the low-risk investment and the bank incurs additional costs cL. In this case its profits are given by

$$\begin{aligned}\Pi_B^H &= \pi_H\left(\left(1+r_L\right)L-\left(1+r_D\right)D\right)-cL \\ &= \left(\pi_H\left(\left(1+r_L\right)+\left(r_L-r_D\right)\left(\kappa-1\right)\right)-c\kappa\right)E.\end{aligned} \tag{18.46}$$

In order to induce the bank to monitor companies, it must be more profitable to do so, $\Pi_B^H \geq \Pi_B^L$, which solves for

$$r_L-r_D \geq \frac{c\kappa-\left(\pi_H-\pi_L\right)\left(1+r_L\right)}{\left(\pi_H-\pi_L\right)\left(\kappa-1\right)}. \tag{18.47}$$

Individuals investing a fraction ρ of their wealth into deposits, and with the high success rate of the investments, the repayment rate of deposits will also be high. The remainder of their wealth is invested directly with the company, who would seek the more risky investment. Their profits are thus given by

$$\begin{aligned}\Pi_D &= \pi_H\left(1+r_D\right)\rho W+\pi_L\left(1+r_L\right)\left(1-\rho\right)W-W \\ &= \pi_H\left(\left(1+r_L\right)-\left(r_L-r_D\right)\right)\rho W+\pi_L\left(1+r_L\right)\left(1-\rho\right)W-W.\end{aligned} \tag{18.48}$$

If individuals we to invest all their wealth into the company directly, they would obtain

$$\hat{\Pi}_D = \pi_L\left(1+r_L\right)W-W. \tag{18.49}$$

It is more profitable to invest some of their wealth into deposits if $\Pi_D \leq \hat{\Pi}_D$, which easily becomes

$$r_L-r_D \leq \frac{\left(\pi_H-\pi_L\right)\left(1+r_L\right)}{\pi_H}. \tag{18.50}$$

Combining equations (18.47) and (18.50), we easily see that for a viable solution, we need the leverage ratio to be not too high as we obtain that we require

$$\kappa \leq \kappa^* = \frac{\pi_L\left(\pi_H-\pi_L\right)\left(1+r_L\right)}{c\pi_H-\left(\pi_H-\pi_L\right)^2\left(1+r_L\right)}. \tag{18.51}$$

We can easily see from equation (18.46) that bank profits are increasing in the leverage, provided $r_L > r_D + c$, which we assume to be the case here. In this case the bank would choose the highest possible leverage, κ^*, such the inequalities in equations (18.47) and (18.50) become equalities and hence solving equation (18.50) we obtain the deposit rate to be

$$1 + r_D = \frac{\pi_L}{\pi_H}(1 + r_L). \tag{18.52}$$

With the requirement that $r_L > r_D + c$, we need the monitoring costs c to be sufficiently small. We easily obtain when inserting for r_D that we require

$$c < \frac{\pi_H - \pi_L}{\pi_H}(1 + r_L). \tag{18.53}$$

From equation (18.51) we also require that

$$c \geq \frac{(\pi_H - \pi_L)^2}{\pi_H}(1 + r_L). \tag{18.54}$$

Combining these two requirements we obtain that

$$\frac{(\pi_H - \pi_L)^2}{\pi_H}(1 + r_L) \leq c < \frac{\pi_H - \pi_L}{\pi_H}(1 + r_L). \tag{18.55}$$

As a further constraint, we need that the leverage is exceeding 1, $\kappa^* > 1$ as otherwise banks do not use deposits. Hence from equation (18.51) we require

$$c < (\pi_H - \pi_L)(1 + r_L), \tag{18.56}$$

which is more restrictive than the upper constraint in equation (18.55) and this constraint becomes

$$\frac{(\pi_H - \pi_L)^2}{\pi_H}(1 + r_L) < c \leq (\pi_H - \pi_L)(1 + r_L). \tag{18.57}$$

We can now continue by introducing deposit insurance and will consider who pays the deposit insurance premium.

Optimal financing sources Deposit insurance may be paid for by those benefitting from it directly, depositors, by charging a fee of τ_D on deposits; it may also be paid by the banks themselves and we assume that for that reason banks would be charged a fee of τ_E on their equity. Finally, the premium might be raised from the general public, similar to taxation. A government guarantee would be similar as the payment from this guarantee will have to be covered through taxation by the general public. In order to avoid a duplication of payment, we here consider that a fee τ_W is levied on the fraction of wealth that is not invested into deposits. Hence the total premium raised from all sources combined is

$$P = \tau_D D + \tau_E E + \tau_W (1 - \rho) W. \tag{18.58}$$

After paying the premium, the deposits available from individuals are $(1-\tau_D)\, D = \rho\,(1-\tau_D)\, W$. As the bank also has to pay its share of the deposit insurance premium, and assuming the leverage ratio is help constant, its deposits are equal to $D = (\kappa - 1)\,(1-\tau_E)\, E$, and setting these two expressions equal gives us

$$\rho = (\kappa - 1)\,\frac{1-\tau_E}{1-\tau_D}\frac{E}{W}. \tag{18.59}$$

Inserting this expression into equation (18.58) gives us for the deposit insurance premium

$$P = \left(\tau_E + (\kappa - 1)\,\frac{1-\tau_E}{1-\tau_D}\,(\tau_D - \tau_W)\right) E + \tau_W W. \tag{18.60}$$

As a fair insurance, this premium is paid out to depositors to cover their losses if the bank is not able to repay their deposits. Hence depositors obtain

$$\begin{aligned}\Pi_D &= \pi_H\,(1+r_D)\,(1-\tau_D)\,D + \pi_L\,(1+r_L)\,(1-\tau_W)\,(1-\rho)\,W \\ &\quad + \pi_L\,(1+r_L)\,P \\ &= (\kappa-1)\,(1-\tau_E)\,((\pi_H - \pi_L)\,(1+r_L) - \pi_H\,(r_L - r_D))\,E \\ &\quad + \pi_L\,(1+r_L)\,(W + \tau_E E)\,.\end{aligned} \tag{18.61}$$

The first expression gives the expected repayments of the deposits, which are only repaid if the monitored bank loan is repaid, and the second term represents the loan given directly to the company, which remains unmonitored. The final term shows the value of the deposit insurance premium, which we assume is invested by the deposit insurance scheme into the company directly. The second equality is obtained when inserting for the deposit insurance premium from equation (18.58).

If depositors invest directly into the company, they obtain

$$\hat{\Pi}_D = \pi_L\,(1+r_L)\,(1-\tau_W)\,W. \tag{18.62}$$

Again, deposits are provided if $\hat{\Pi}_D \geq \Pi_L$, which solves for

$$r_L - r_D \leq \frac{(\kappa-1)\,(\pi_H - \pi_L)\,(1+r_L)\,(1-\tau_E)\,E + \pi_L\,(\tau_W W + \tau_E E)}{\pi_H\,(\kappa-1)\,(1-\tau_E)\,E}. \tag{18.63}$$

The incentives for the bank are unchanged, merely the equity reduces to $(1-\tau_E)\,E$, thus the constraint in equation (18.47) remains valid and combining this with equation (18.63), we obtain a viable solution is available if the leverage of the bank is below

$$\kappa \leq \kappa^{**} = \frac{\tau_W W + E}{(1-\tau_E)\,E}\kappa^{*}, \tag{18.64}$$

where κ^* is the leverage ratio banks choose in the absence of deposit insurance, it was defined in equation (18.51). We see immediately that $\kappa^{**} > \kappa^*$ as the paid-out deposit insurance makes deposits more attractive. Thus with deposit insurance, banks will choose a higher leverage as we can easily show that banks profits are again increasing with leverage.

As bank lending is socially beneficial due to the monitoring of banks that reduces the risks companies take, it would be optimal to maximize bank-lending. The maximum bank lending is achieved if $\rho = 1$ such that $D = W$, giving total lending of $L = (1 - \tau_E) E + (1 - \tau_D) W$, which is clearly maximized for

$$\tau_E = \tau_D = 0, \tag{18.65}$$

implying $L = E + W$. Thus any deposit insurance would be paid for by non-depositors, what is often referred to as general taxation. Therefore, optimally, depositors and banks are subsidized by the general public. This is overall optimal because banks provide additional benefits in the form of monitoring companies' investments; these benefits can only be realised if banks obtain deposits. Thus ensuring the provision of deposits is maximized, through not requiring them to contribute to the deposit insurance, generates the highest social surplus.

As we also find that $L = \kappa^{**} (1 - \tau_E) E$, we have from setting $L = E + W = \kappa^{**} (1 + \tau_E) E$ that the optimal fee the general public contributes to the deposit insurance is given by

$$\begin{aligned} \tau_W &= \frac{\pi_H c - (\pi_H - \pi_L)^2 (1 + r_L)}{\pi_L (\pi_H - \pi_L)(1 + r_L)} \frac{E + W}{W} - \frac{E}{W} \\ &= \frac{1}{\kappa^*} \frac{E + W}{W} - \frac{E}{W}. \end{aligned} \tag{18.66}$$

We have $\tau_W \geq 0$ if $\kappa^* E \leq E + W$, which means that in the absence of deposit insurance, the total loans $(\kappa^* E)$ cannot exceed the total resources available, $E + W$, which is trivially true. Hence the fee paid the general public is positive.

Using the constraint to ensure that deposits are provided to banks, equation (18.63), as an equality and inserting the leverage ratio κ^{**}, the fee $\tau_E = 0$ and τ_W from equation (18.66) and noting that $(\kappa^{**} - 1) E = D = W$, we get the deposit rate as

$$1 + r_D = \frac{\pi_L}{\pi_H} (1 + r_L)(1 + \tau_W). \tag{18.67}$$

This deposit rate in higher than without deposit insurance, as comparison with equation (18.52) easily shows. This is done to increase the attractiveness of deposits such that depositors do not invest directly into the company.

As $\rho = 1$, $\tau_E = \tau_D = 0$, there is no deposit insurance premium, as given in equation (18.58), that is actually raised, thus no deposits can be insured and it is only the thread of taxation of non-deposit wealth, combined with the higher interest rate on deposits, that more deposits are achieved. We finally find that when inserting all variables, the leverage ratio in equation (18.64) will be given by $\kappa^{**} = 1 + \frac{W}{E}$.

We thus find that it is optimal for deposit insurance premia to be paid by the general public, those not providing deposits to banks, so as to maximize the benefits of banks' monitoring effort. However, as all wealth is deposited no actual premia are raised and thus deposit insurance would not be provided. Such a result is unsatisfactory as it cannot explain the financing of deposit insurance. However, it depends on the assumption that banks are able to accept all wealth as deposits, implying a very high leverage ratio. It is much more common that regulators will impose a maximum leverage that banks are allowed. We will consider such a constraint next.

Leverage limits Let us now consider a regulator that imposes a maximum leverage ratio of $\overline{\kappa} < \kappa^{**}$ on a bank, which thus restricts its ability to obtain deposits. The loans the bank can provide are this given by

$$L = (1 - \tau_E) E + (1 - \tau_D) D = (1 - \tau_E) \overline{\kappa} E \tag{18.68}$$

if we insert for $D = \rho W$ and for ρ from equation (18.59) as well as $D = (\overline{\kappa} - 1) E$. Maximizing bank loans therefore implies again that $\tau_E = 0$ and banks should not contribute to the deposit insurance premium as that would reduce the amount of bank lending due to reduced equity.

Setting $\kappa^{**} = \overline{\kappa}$, we obtain from equation (18.66) that the fee charged on non-deposits is given by

$$\tau_W = \frac{\overline{\kappa} \pi_H c - (\pi_H - \pi_L)(1 + r_L)((\pi_H - \pi_L)\overline{\kappa} + \pi_L)}{\pi_L (\pi_H - \pi_L)(1 + r_L)} \frac{E}{W} > 0. \tag{18.69}$$

Suppose now that we want to raise deposit insurance premia sufficient to cover deposits fully; these losses are arising if the bank loans are not repaid. The deposit insurance premium raised will be invested into the companies directly, such that a full coverage of the losses requires

$$(1 - \pi_H)(1 + r_D) D = \pi_L (1 + r_L) P. \tag{18.70}$$

Inserting all expressions for deposits D and deposit insurance premium P, this expression solves for

$$\frac{E}{W} = \frac{(1 - \tau_D) \tau_W \pi_L (1 + r_L)}{(1 - \pi_H)(1 + r_D) - \pi_L (1 + r_L)(\tau_D (1 - \tau_D) - \tau_W)} \frac{1}{\overline{\kappa} - 1}. \tag{18.71}$$

We can now insert for τ_W from equation (18.69) and solve the resulting expression for τ_D. It is apparent that the solution for τ_D will be positive. We thus find that if a regulator restricts the leverage, full coverage of deposits can be ensured by collecting a fee from depositors and the general public; it is never optimal for banks to contribute to the deposit insurance. The reason is that this fee on banks will reduce their equity and hence given the constraints on leverage reduce lending by a factor $\overline{\kappa}$, which in turn reduces the revenue from banks available to repay depositors, increasing the possible payout required from the deposit insurance. Raising revenue from depositors has a much lower impact on the amount that can be lent; while a raising the deposit insurance premium from non-depositors would not affect lending at all, but charging a too high fee would provide incentives to deposit their wealth instead, which banks could not accept due to the constraints on their leverage. It is thus a balance between fees charged to depositors and non-depositors that allow to raise the necessary deposit insurance premium.

Summary We have see that deposit insurance should optimally be financed by depositors and general taxation, but not by banks itself. If a leverage constraint has been imposed by regulators, the effect on the ability of banks to provide loans can be profound as the amount charged to the bank will reduce their lending by a multiple of this fee, increasing the costs of the deposit insurance through less revenue from lending to repay depositors. We found that this leverage restriction is central to be able to finance deposit insurance.

The common observation that deposit insurance is paid for by taxpayers rather than banks or depositors is consistent with the results we obtained here. Depositors are also taxpayers, as are those who are not depositing their wealth with banks, with most taxpayers are also being depositors, and while different fees might be charged to different groups of taxpayers in our model, it recovers the observation that very few deposit insurance schemes are funded directly by banks. Most such deposit insurance schemes take the form of government guarantees of deposits, either explicit through legislation or through implicit guarantees, which implies that if the deposit insurance has to pay out, taxation will be used to recover these payments.

Reading Morrison & White (2011)

Conclusions

Deposit insurance might increase the incentives of banks to increase the risks they are taking when providing loans. They can do so by either directly providing loans to companies that are more likely to default or by increasing their leverage, providing more loans given the amount of equity they hold and thereby possibly increasing the losses on depositors. If deposit insurance is priced incorrectly and does not take into account the risks banks are taking appropriately, they may take more risks, either individually, or as part of a banking system. Banks may take additional risks if the deposit insurance is offered too cheaply as in this case, they can obtain higher returns resulting from their more risky behaviour without facing higher costs. Without deposit insurance, depositors would take into account these higher risks bank are taking and the deposit rate would increase such that banks have no incentive to increase risks. With deposit insurance, the deposit rate will be unaffected as deposits are deemed to be safe and hence bank can increase profits from taking higher risks. Banks might also increase risks by aligning their businesses more, resulting in a situation where multiple banks fail at the same time, increasing the likelihood of a bailout compared to a situation where only a single bank would fail. If deposit insurance takes adequately into account these additional risks, the incentives for such strategic decisions by banks can be reduced.

The coverage of deposit insurance is limited to specific groups of depositors, in most cases individual depositors and sometimes small businesses, as well is there often a limit on the amount of deposits that are insured. If deposit insurance is not paid for, banks benefit from deposit insurance through their ability to pay lower deposit rates as the risks the bank takes does not need to be accounted for. However, deposit insurance will increase the competition between banks as the surplus that banks have are higher, which will erode the benefits of paying lower deposit rates. Banks will seek to limit such competition by applying an upper limit on the amount of deposits that are insured. That way banks can benefit from lower deposit rates on those deposits that are insured and at the same time limit competition for uninsured deposits, allowing them increase their profits. Similarly, banks would not want all types of deposits insured. Banks that face low risks may not benefit from adequately priced deposit insurance at all as the deposit rate was not much higher without deposit insurance, while d more risky banks would benefit from the reduced deposit rates. They would, however not want to insure all deposits; this would imply a high deposit insurance premium which would divert resources from other profitable investments. Thus they would like to limit the deposit insurance coverage.

The deposit insurance premium, whether charged in a conventional insurance model as an upfront fee or with government guarantees after any payout has been made, needs to be paid for. If banks have limits on their leverage, it would not be optimal for them to pay the deposit insurance premium as the resources this requires will not be available to contribute to the capital on which the leverage ratio is based, reducing the amount of lending that os can be conducted. It would therefore be optimal for depositors to be charged a fee and the general public to contribute through taxation. As depositors are commonly also taxpayers and most taxpayers are depositors, we thus see that any deposit insurance should be paid for out these two overlapping groups.

Chapter 19

Payment services

Deposit accounts are not only used to invest excess funds depositors have with the aim of gaining interest. Banks offer a wider range of services associated with such accounts, most notably they allow payments between accounts to be made, negating the need to settle any debt or invoices using cash. Depositors conduct a large number of such payments between accounts and hence significant amounts of payments are flowing between different banks. In chapter 19.3 we will explore the implications of banks organising such payments between them and what potential implications of this are.

However, deposit accounts are also used to access cash. However, such cannot only be accessed at the bank holding the account, but at cash machines belonging to different banks and also in the form of cashback in retail stores. How banks can agree such arrangements is discussed in chapter 19.1, which also includes the consideration for access to other banking services through online services. Finally, the question of the optimal fee for deposit accounts to pay for these additional services is considered.

Customers do not only rely on making payments in cash or by bank transfer, but the use of payment cards accounts for an increasing fraction of such payments within the retail sector. In chapter 19.2 we will therefore investigate the issue and management of this payment method.

19.1 Account services

While deposits are mostly seen as a form of investment for depositors and a source of funding for banks, there are many other services provided to depositors. Banks provide their depositors with the ability to withdraw cash from their account, either in their branches or through cash machines. Typically such withdrawals are not limited to facilities provided by the bank itself, but it is common for depositors to be able to withdraw cash from cash machines operated by other banks or even from non-banks such as retailers. In chapter 19.1.1 we will look into the incentives for banks to cooperate with each other in providing these services and under which conditions banks may allow the customers all or only selected competitors access to their services. How access to such services should be compensated for between banks is explored in chapter 19.1.2 when discussing the so-called interchange fee.

It has become common for banks to offer account services online, although not all banking activities can be accessed this way, with some requiring personal attendance at a branch, even though the types of services in this category constantly reducing. In chapter 19.1.3 we will investigate which type of services bank will offer remotely and which ones are retained as being available in branches only.

Finally, in many cases deposit accounts and their services are provided free, while in other cases a fee is charged. It seems that whether fees are charged or not is a question of the market segment the bank is operating in with most banks operating a similar model in that specific market segment. How banks can determine whether a fee is charged and if so how much is analysed in chapter 19.1.4. We finally determine in chapter 19.1.5 why account fees are waived if the deposits made are large, while accounts with smaller deposits are charged a fee.

19.1.1 Bank cooperation for cash access

Depositors often can withdraw cash, pay in cash or cheques, check their account balance, and access other services at cash machines. While a wide range of services are offered at cash machines operated by the bank the depositor maintains his account with, some of these services are also available to those holding their account at another bank. Accessing services using the cash machines of another bank requires an agreement between those two banks to allow such access. Of course, by being able to access the cash machines of another bank, it becomes easier for depositors to access any services, making their bank more attractive than another bank that does not enable access; it hence affects the competition between banks.

Let us assume that banks offer account services that are different between banks in that they provide different services, such as the access to online services or product ranges offered. Each depositor will have their own preferences for the type of services and products they would want to access and we assume that these preferences can be represented by a location on a circle. Each depositor will be located at a specific point on this circle denoting their preferences and banks will establish themselves at specific points, representing a bundle of services and products. The further a depositor is away from a bank, the less it meets his preferences, reducing his utility by the amount of c per unit of distance. Depositors are located uniformly around this circle and we consider the case of three banks offering their services. The banks are located at a distance of 1 unit from each other.

Depositors seek access to banks for two types of services. One service is available only at the bank they are having their account with, and they require such services a fraction $1 - \lambda$ of times, while for a fraction λ they seek services that can also be accessed at cash machines. In addition to the regular access to theses services, depositors may seek access to cash machines while on a random location of this circle, for example if requiring services while shopping or during holidays. Such access is required with probability p and as such services are required with probability λ the expected costs are $p\lambda c$ for the distance to the bank they can access these services at.

If banks are not allowing access to each others' cash machines, depositors have to access the cash machines of their own bank; with banks being at a distance of 1 from each other on this circle, the longest distance to their own bank would be $\frac{3}{2}$ and with the location at which the services is to be accessed being equally distributed on this circle, the average distance would be $d_N = \frac{3}{4}$. If all banks allow access to each others' cash machines, the maximal distance to a bank would be $\frac{1}{2}$ and hence the average distance would be $d_C = \frac{1}{4}$. In the case that only two of the banks allow access to each other's cash machines by their respective depositors, their depositors will be either located between them with a maximum distance of $\frac{1}{2}$, giving average distance of $\frac{1}{4}$, or they are located outside of these two banks, where the maximum possible distance to one of the banks is 1, if the depositor is located at the position of the bank not allowing access to their cash machines; hence the average distance would be $\frac{1}{2}$. The probability of being located between these two bank is $\frac{1}{3}$, and hence the average distance to access a cash machine in this case would be $d_P = \frac{1}{3}\frac{1}{4} + \frac{2}{3}\frac{1}{2} = \frac{5}{12}$. If the depositor holds his account with the bank that does not cooperate with the two other banks to allow access to their cash machines, he cannot benefit from the agreement of the two other banks and his average distance will remain at $d_N = \frac{3}{4}$.

We can now assess how any such cooperations between banks to allow depositors access to each others' cash machines will affect competition between them and thus which degree of cooperation is optimal. We will initial consider the case where no fees are charged for accessing the cash machine of another bank.

Free access to cash machines Let us assume that cash withdrawals at cooperating banks are free to depositors and do not impose additional costs on the bank. We can now compare the profits banks are making when cooperating with one or both banks in the provision of cash services.

No cooperation If banks are not cooperating, depositors have to conduct all their business at their own bank. With bank i paying interest r_D^i on deposits D and customers being located at a distance d_i to their bank, the net benefits to the customer is given by

$$\Pi_D^i = \left(1 + r_D^i\right) D - c d_i D - p\lambda d_N D. \tag{19.1}$$

The first term accounts for the interest on their deposits and the second term adjusts this for the costs arising from the use bank services at their own bank and the final term the costs of accessing cash machines while at a random location.

A depositor prefers bank i over bank j if its profits are higher, thus $\Pi_D^i \geq \Pi_D^j$. Noting that $d_j = 1 - d_i$ as banks are located one unit apart, we get that the depositors prefer this bank if their location is sufficiently close to the bank. The requirement is

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}. \tag{19.2}$$

Similarly we get for the decision between banks i and bank k that

$$d_i \leq \hat{d}_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^k\right)}{2c}. \tag{19.3}$$

Banks will obtain deposits from all those depositors located closer than d_i^* and $\hat{d}_i^*$, giving it a market of $d_i^* + \hat{d}_i^*$. We assume that banks use their deposits to fully finance loans at a loan rate r_L, where loans are repaid with probability π. Hence their profits are given by

$$\Pi_B^i = \left(\pi\left(1 + r_L\right) - \left(1 + r_D^i\right)\right)\left(d_i + \hat{d}_i\right) D. \tag{19.4}$$

Banks choose the deposit rate optimally such that it maximizes their profits, which after inserting from equations (19.2) and (19.3) for the market share of depositors gives us the first order condition as

$$\begin{aligned}\frac{\partial \Pi_B^i}{\partial\left(1 + r_D^i\right)} = &-\left(1 + \frac{2\left(1 + r_D^i\right) - \left(1 + r_D^j\right) - \left(1 + r_D^k\right)}{2c}\right) D \\ &+ \left(\pi\left(1 + r_L\right) - \left(1 + r_D^i\right)\right)\frac{D}{c} = 0.\end{aligned}$$

As all banks are identical, we will only consider symmetric equilibria where deposit rates are identical, i.e. $r_D^i = r_D^j = r_D^k = r_D$. Inserting this requirement into the above first order condition we easily obtain

$$1 + r_D = \pi\left(1 + r_L\right) - c. \tag{19.5}$$

Inserting this result, we obtain that the market shares of banks from equations (19.2) and (19.3) becomes $d_i = \hat{d}_i = \frac{1}{2}$ and we get from equstion (19.4) that

$$\Pi_B^N = cD. \tag{19.6}$$

Full cooperation The other extreme assumption would be that all banks cooperate by allowing access to their cash machines to depositors of other banks. In this case, depositors will use the cash machine of the nearest bank, while still attending their own bank for other services. Thus the profits of depositors are given as

$$\Pi_D^i = \left(1 + r_D^i\right) D - (1 - \lambda) d_i c D - \lambda \min\left\{d_i, d_j, d_k\right\} cD - p\lambda c d_C D. \tag{19.7}$$

The first term represents the profits from the interest on deposits, the second term the costs of accessing services at their bank and the third term the costs of accessing cash services at any of the banks, while the final term takes into account access to cash services while at a random location.

A depositor prefers bank i over bank j if its profits are higher, thus $\Pi_D^i \geq \Pi_D^j$. Noting that $d_j = 1 - d_i$ as banks are located one unit apart, we get that the depositors prefer this bank if their location is sufficiently close to the bank. Following the same steps as in the case of banks not cooperating, we obtain the deposit rate and bank profits as

$$\begin{aligned} 1 + r_D &= \pi \left(1 + r_L\right) - \left(1 - \lambda\right) c, \\ \Pi_B^C &= \left(1 - \lambda\right) cD. \end{aligned} \tag{19.8}$$

We see that with banks cooperating, their profits are lower, $\Pi_B^C = (1 = \lambda)\, \Pi_B^N \leq \Pi_B^N$. The cooperation between banks increases their competition for deposits as reflected in the higher deposit rate. Competition is increased as the distance to their own bank to access cash services becomes less important, reducing the effect of differentiated accounts.

Partial cooperation In the intermediate case that two banks cooperate and allow access to cash machines for each other's depositors, their depositors will obtain profits of

$$\Pi_D^i = \left(1 + r_D^i\right) D - \left(1 - \lambda\right) c d_i D - \lambda \min\left\{d_i, d_j\right\} cD - p\lambda c d_P D, \tag{19.9}$$

The first term represents the interest gained on their deposits and the second term the costs of accessing bank services at their own bank. The third term shows the costs of accessing cash services at the bank which is nearest to them, provided they are cooperating, and the final term takes into account the costs of access to cash services while at a random location. For depositors choosing between these two cooperating banks, bank i is preferred if $\Pi_D^i \geq \Pi_D^j$, which easily solves for

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}. \tag{19.10}$$

If bank k is not cooperating with the other two banks, then their depositors obtain profits as given in equation (19.1) where banks did not cooperate as the cooperation of the other banks, doe snot affect them, while the profits of depositors with the cooperating bank obtain profits according to equation (19.9). As the cooperating banks are identical and we only consider symmetric equilibria such that $r_D^i = r_D^j$ and hence from equation (19.10) we have $d_i^* = \frac{1}{2}$ and hence the depositor will be closer to its own bank than the other cooperating bank. We thus have $\min\left\{d_i, d_j\right\} = d_i$ and obtain that depositors access all services at their own bank and hence $\Pi_D^i = \left(1 + r_D^i\right) D - c d_i D - p\lambda c d_P D$. Thus for a depositor to prefer the cooperating bank i over the bon-cooperating bank k we require that $\Pi_D^i \geq \Pi_D^k$, which solves for

$$d_i \leq \hat{d}_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^k\right) - p\lambda c \left(d_N - d_P\right)}{2c}. \tag{19.11}$$

The profits of the cooperating bank is then given by

		Bank j	
		Cooperating	Not cooperating
Bank i	Cooperating	$\Pi_B^P, \Pi_B^P, \hat{\Pi}_B^P$	$\Pi_B^N, \Pi_B^N, \Pi_B^N$
	Not cooperating	$\Pi_B^N, \Pi_B^N, \Pi_B^N$	$\Pi_B^N, \Pi_B^N, \Pi_B^N$

(a) Bank k not cooperating

		Bank j	
		Cooperating	Not cooperating
Bank i	Cooperating	$\Pi_B^C, \Pi_B^C, \Pi_B^C$	$\Pi_B^P, \hat{\Pi}_B^P, \Pi_B^P$
	Not cooperating	$\hat{\Pi}_B^P, \Pi_B^P, \Pi_B^P$	$\Pi_B^N, \Pi_B^N, \Pi_B^N$

(b) Bank k cooperating

Fig. 19.1: Cooperation game for access to cash services

$$\Pi_B^i = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^i\right) D\right) \left(d_i^* + \hat{d}_i^*\right). \tag{19.12}$$

Maximizing this expression with respect to the deposit rate $1 + r_D^i$ and noting that the two cooperating banks are equal, implying $r_D^i = r_D^j$, we use the first order condition $\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} = 0$ and obtain

$$\begin{aligned} 1 + r_D^i = & \frac{(2-\lambda) c}{(3-2\lambda) c}\pi\left(1 + r_L\right) \\ & - \frac{(1-\lambda) c}{(3-2\lambda) c}\left(2c - \left(1 + r_D^k\right) + \frac{p\lambda c}{3}\right). \end{aligned} \tag{19.13}$$

For the non-cooperating bank we then have similarly their profits given by

$$\Pi_B^k = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^k\right) D\right) \left(\left(1 - \hat{d}_i^*\right) + \left(1 - \hat{d}_j^*\right)\right). \tag{19.14}$$

The first order condition $\frac{\partial \Pi_B^k}{\partial\left(1+r_D^k\right)} = 0$, when noting that $r_D^i = r_D^j$, solves for

$$1 + r_D^k = \frac{1}{2}\pi\left(1 + r_L\right) - \frac{1}{2}c + \frac{1}{2}\left(1 + r_D^i\right) + \frac{p\lambda c}{6}. \tag{19.15}$$

Combining equations (19.14) and (19.15), we solve the deposit rates as

$$1 + r_D^i = \pi\left(1 + r_L\right) - \frac{1 + 4c + \frac{p\lambda c}{3}}{5 - 3\lambda}\left(1 - \lambda\right), \tag{19.16}$$

$$1 + r_D^k = \pi\left(1 + r_L\right) - \frac{1}{2}\frac{2c + \left(7c + 1\right)\left(1 - \lambda\right) - \frac{2}{3}\left(2 - \lambda\right)\lambda pc}{5 - 3\lambda}.$$

Using this result, we can insert the deposit rates into equations (19.10) and (19.11) and then obtain the profits of the cooperating banks in equation (19.12) and of the non-cooperating bank in equation (19.15). It can be shown that the profits of the two banks cooperating are higher than the bank not cooperating, $\Pi_B^P \geq \hat{\Pi}_B^P$.

Having established the properties of each strategy, we can now continue to assess the equilibrium cooperation between banks.

Equilibrium strategies The three banks now enter a strategic game of cooperation as shown in figure 19.1, where we note that for cooperation at least two banks are needed and hence if only one bank would cooperate, the profits of non-cooperations are obtained. We denote by Π_B^P the profits of those two banks that cooperate and by $\hat{\Pi}_B^P$ the profits of the bank not cooperating. With the above, we know that $\Pi_B^P \geq \hat{\Pi}_B^P$ and noting that we had established that $\Pi_B^N \geq \Pi_B^C$, we can now analyze the equilibrium of this game.

We can now distinguish a number of cases. If $\Pi_B^P > \Pi_B^N > \hat{\Pi}_B^P > \Pi_B^C$ or $\Pi_B^P > \hat{\Pi}_B^P > \Pi_B^N > \Pi_B^C$, we can see that the only equilibria are those in which two of the banks cooperate to provide access to the cash service of each other's depositors. In the case that $\Pi_B^N > \Pi_B^P > \hat{\Pi}_B^P > \Pi_B^C$ or $\Pi_B^N > \Pi_B^P \Pi_B^C > \hat{\Pi}_B^P >$ the equilibrium is for all banks to not cooperate and if $\Pi_B^N > \Pi_B^P > \Pi_B^C > \hat{\Pi}_B^P$ all banks provide access to cash services for each others' depositors.

An analytical expression for the parameter constellations corresponding to the different equilibria is difficult to obtain and interpret; however, figure 19.2 illustrates this relationship. If the costs c are small, the benefits from offering differentiated accounts are small as the losses for depositors not obtaining their preferred services are small, leading to a high degree of competition. In this case, banks cooperating to provide access to cash services to depositors at other banks will erode the small degree of market power they have retained, making such cooperation not profitable. As the costs of depositors not obtaining their preferred account services increases, banks will find it increasingly difficult to compete for depositors that are more inclined to the services of other banks. Cooperating with another bank will give them an advantage in attracting depositors due to their ability to access cash services at lower costs if at a random location. While the cooperation with another bank increases the degree of competition between those two banks, the competitive advantage they gain over the excluded bank and hence the ability to attract additional depositors will compensate for this effect. If the costs to depositors are not too high, the competition from all three banks cooperating will be too high and one bank will remain excluded; its profits are higher than when all banks were to cooperate. As the costs to depositors increase, the advantage cooperating banks have over the excluded bank increases and the excluded bank will seek to join the cooperation so that it is able to compete with the other banks on an equal footing and increase its profits, leading to the full cooperation of all banks. The more important cash services are, λ, the more depositors benefit from the cooperation of banks and are willing to accept lower deposit rates in return for being able to access these services at other banks. Thus, the degree of cooperation is increasing in the importance of cash services.

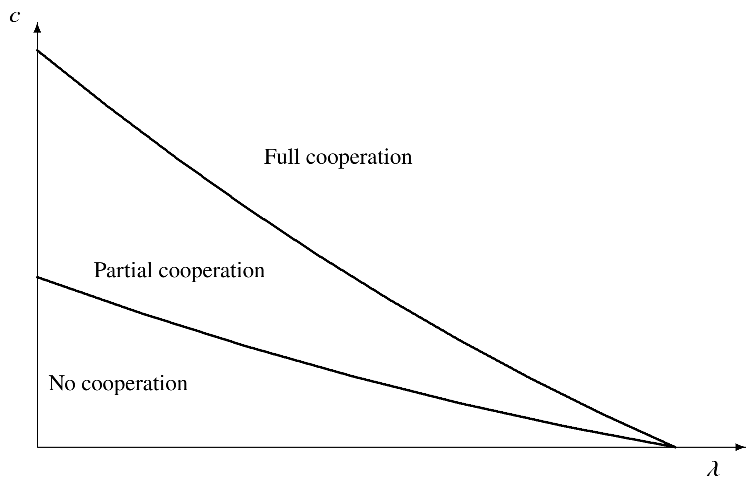


Fig. 19.2: Equilibrium cooperation for cash services without access fees

We thus see that we should expect banks to cooperate in the provision of cash services if depositors have strong preferences for the services offered by a bank and cash services are important to them. As either these preferences or the importance of cash services declines, the cooperation between banks reduces until cooperation seizes fully.

Access fees Thus far we have assumed that banks not only cooperate in providing access to cash services for depositors of their competitors, but provide these services for free. Let us now assume that access to cash services at a bank other than a depositor's own bank is charged a fee f. Again, we investigate cases with different levels of cooperation between banks.

No cooperation If banks are not cooperating in the provision of cash services, there cannot be any withdrawals at other banks and the profits this generates to banks will be identical to the case of no access fees, hence

$$\Pi_B^N = cD. \tag{19.17}$$

Full cooperation If all banks cooperate and provide cash services to depositors of all other banks, then the benefits to depositors are given similarly to equation (19.7) by

$$\Pi_D^i = \left(1 + r_D^i\right) D - (1 - \lambda)\, d_i c D - \min\left\{\lambda c d_i, \lambda c d_j + f, \lambda c d_k + f\right\} D - p \lambda c d_C D. \tag{19.18}$$

The access fee f is here added to the costs for the withdrawal at any of the other banks. If we consider a depositor located between banks i and j, we can ignore the final term of $\lambda c d_k + f$ as this bank would be too car away to be considered. Similarly, we get for the same depositor the benefits from choosing bank j as

$$\Pi_D^j = \left(1 + r_D^j\right) D - (1 - \lambda)\, c d_j D - \min\left\{\lambda c d_i + f; \lambda c d_j\right\} D - \lambda c d_C D.$$

Noting that $d_j = 1 - d_i$, we assume that $f \geq \lambda c$ such that the withdrawal cost is always larger than the cost to use the bank, such that the depositor would always use cash services at their own bank, unless in a random location. Thus we have $\min\left\{\lambda c d_i; \lambda c d_j + f\right\} = \lambda c d_i$ and $\min\left\{\lambda c d_i + f; \lambda c d_j\right\} = \lambda c d_j$. Inserting these relationships, we get that a depositor prefers bank i if $\Pi_D^i \geq \Pi_D^j$, which solves for

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}. \tag{19.19}$$

If in a random location, the depositors that might withdraw from bank i and pay an access fee f are all those that are not its own depositors, which are $3 - d_i - \hat{d}_i$, and as depositors were assumed to be uniformly distributed around the circle and all banks are identical, a fraction $\frac{1}{3}$ of these will go to bank i. Hence the total additional revenue will be $p\frac{3-d_i-\hat{d}_i}{3} f D$, where p denotes the probability of depositors wanting to withdraw cash. Hence bank profits consist of the profits generates through their own depositors as well as those depositors from other banks using their cash services, which gives us

$$\Pi_B^i = \left(\pi (1 + r_L) L - \left(1 + r_D^i\right) D\right)\left(d_i + \hat{d}_i\right) + p\frac{3 - d_i - \hat{d}_i}{3} f D. \tag{19.20}$$

Noting that $\hat{d}_i = \frac{1}{2} + \frac{\left(1+r_D^i\right)-\left(1+r_D^k\right)}{2c}$ in analogy to equation (19.19), we get the first order condition for a profit maximum as $\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} = 0$. As all banks are alike, we only consider symmetric equilibria such that $r_D^i = r_D^j = r_D^k - r_D$ and thus $d_i = \hat{d}_i = \frac{1}{2}$, which then gives us

$$1 + r_D = \pi (1 + r_L) - c - \frac{pf}{3}. \tag{19.21}$$

Inserting this deposit rate into the bank profits of equation (19.20) gives us

$$\Pi_B^C = (c + pf)\, D > \Pi_B^N. \tag{19.22}$$

Hence the profits banks make when fully cooperating to provide access to cash services to all depositors of other banks for a fee f, their profits are increased compared to the case of no bank providing any access. While competition between banks will increase as in the case without an access fee and thus reduce bank profits, this is compensated here with the fee income. The fee also reduces competition between banks compared to the case without such an access fee, thus reduces the lost profits from competition.

Partial cooperation If only two banks are cooperating to provide access to cash services, the profits of banks are given as above in equations (19.10) and (19.11). As was explained there, depositors will use cash services only at their own bank and hence no additional revenue if created for bank. Bank i is preferred over banks j and k, the non-cooperating bank, respectively, if

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}, \tag{19.23}$$
$$d_i \leq \hat{d}_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^k\right) - p\lambda\left(d_N - d_P\right)}{2c}.$$

Here, however d_P will change as the benefits of cooperation are reduced due to the access fee. If the need to withdraw cash emerges when located between the two cooperating banks, the depositor will go to its own bank as we assumed that $f \geq \lambda c$. The distance between the two cooperating banks is 2 and the customers goes to the other bank if $\lambda c d_i \geq \lambda c\left(2 - d_i\right) + f$, or $d_i > \frac{2\lambda c + f}{2\lambda c}$. As the total length of the circle is 3, the probability of paying this fee is $\frac{2\lambda c + f}{6\lambda c}$, thus

$$d_P = \frac{5}{12} - \frac{2\lambda c + f}{6\lambda c} f. \tag{19.24}$$

The profits of the cooperating banks are then given by

$$\Pi_B^i = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^i\right) D\right)\left(d_i + \hat{d}_i\right) + p\frac{2\lambda c + f}{6\lambda c} f D, \tag{19.25}$$

where the last term accounts for the fee income. Using equation (19.23), we easily solve the first order condition $\frac{\partial \Pi_B^i}{\partial\left(1 + r_D^i\right)} = 0$ as

$$1 + r_D^i = \frac{2}{3}\pi\left(1 + r_L\right) - \frac{2}{3}c + \frac{1}{3}\left(1 + r_D^k\right) - \frac{p\lambda c}{3}\left(d_N - d_P\right). \tag{19.26}$$

where we again restricted ourselves to symmetric equilibria with $r_D^i = r_D^j - r_D$.

Similarly for the non-cooperating bank, we obtain the same result as in equation (19.15)

$$1 + r_D^k = \frac{1}{2}\pi\left(1 + r_L\right) - \frac{1}{2}c + \frac{1}{2}\left(1 + r_D^i\right) + \frac{p\lambda c}{2}\left(d_N - d_P\right). \tag{19.27}$$

Solving equations (19.26) and (19.27) gives us the deposit rates of the cooperating and non-cooperating banks, respectively, as

$$
\begin{aligned}
1 + r_D^i &= \pi\,(1 + r_L) - c + \frac{4p\lambda c}{15}\,(d_N - d_P)\,, \\
1 + r_D^k &= \pi\,(1 + r_L) - c + \frac{p\lambda c}{5}\,(d_N - d_P)\,.
\end{aligned}
\tag{19.28}
$$

Having established the properties of each strategy, we can now continue to assess the equilibrium cooperation between banks.

Equilibrium strategies We know that all banks cooperating in providing cash access to their competitors is more beneficial than not providing such access, $\Pi_B^C > \Pi_B^N$, and when inserting equation (19.28) into equation (19.23) and subsequently equation (19.25), we can show that $\Pi_B^P > \Pi_B^N$. Determining the equilibrium of the strategic game in figure 19.1, we see that the equilibrium is for all banks to fully cooperate.

Thus, if an access fee is charged to depositors from other banks for accessing their cash services, banks will fully cooperate by allowing access to all depositors. This result arises because the access fee on the one hand limits competition between banks as depositors seek to access cash services with their own bank due to the increased costs of other banks, and on the other hand the access fees generate additional revenue for banks. This provides banks with sufficient additional revenue to overcome the slightly increased competition from cooperating.

Summary Banks will in general cooperate in providing access to cash services for depositors of other banks. Cooperation between banks will allow their depositors to access services more easily and thus any competitive advantage a bank had, such as more cash machines, a more physical branches, or more services being available through remote access, will be eroded, weakening their market position and increasing competition between banks. While access to services at other banks may make a bank itself more attractive, these two aspects will have to be balanced and if the market position of banks is weak, the increased competition eroding the small degree of market power will be the stronger effect, making cooperation between banks less likely. This may lead to a situation where groups of banks cooperate with each other, but the cooperation does not extend to the all banks, giving rise to groups of banks cooperating with each other.

On the other hand, if banks are able to charge an access fee, cooperation between banks would be complete. The access fee limits the increase in competition between banks and generates additional income top banks, making such cooperation profitable. The access fee does not have to be levied directly on depositors accessing the service at another bank, it can also be charged to the bank of that depositor; the effect is identical as banks will adjust the deposit rate to account for the additional costs depositors incur and whether the revenue originates with the depositor or another bank, is irrelevant for the profits of the bank providing access. It is thus that access fees encourage a wider cooperation between banks, which benefit depositors, but will impose additional costs on them.

Reading Matutes & Padilla (1994)

19.1.2 Interchange fees for cash services

Banks often allow depositors of other banks to use some of their facilities for cash withdrawals, enquiries on account balances, or even the depositing of cash or cheques. Access to such services is usually through the use of cash machines and depositors are not charged any fees for this access. However, the bank of the depositor accessing these services is charged a fee by the bank providing this service a so-called interchange fee.

Let us assume that we have M banks offering their cash services to depositors of all other banks free of charge and each bank has N_i access points for their services, for example cash machines, such that in total there are $N = \sum_{i=1}^{M} N_i$ access points. If depositors access cash services randomly, each bank will provide a fraction $\frac{N_i}{N}$ of these services. As an interchange fee f is only charged to depositors of the other $M-1$ banks, a bank charges for a fraction $\frac{M-1}{M}$ of cash services accessed, they remaining being by their own depositors. Thus their income from the interchange fee will be $f \frac{N_i}{N} \frac{M-1}{M} D$, where we assume that the frequency with which the services are accessed by each depositor depends on the deposit size D. At the same time, banks have to pay interchange fees yo other banks. These fees are paid on the $1 - \frac{N_i}{N}$ service accesses that are made at other banks and the deposits held at each bank is $\frac{D}{M}$. This will give the bank revenue of $f \left(1 - \frac{N_i}{N}\right) \frac{D}{M}$. Offering these cash services incurs fixed costs of C for each of the access points. This allows us to determine the bank profits as

$$\begin{aligned} \Pi_B^i &= \pi \left(1 + r_L\right) L - \left(1 + r_D^i\right) \frac{D}{M} \\ &\quad -f \left(1 - \frac{N_i}{N}\right) \frac{D}{M} + f \frac{N_i}{N} \frac{M-1}{M} D - CN_i \\ &= \left(\pi \left(1 + r_L\right) - \left(1 + r_D^i - f\right)\right) \frac{D}{M} + f \frac{N_i}{N} D - CN_i, \end{aligned} \tag{19.29}$$

where π denotes the probability of loans being repaid, r_L, the loan rate, r_D^i the deposit rate bank i charges and L the amount of loans, assuming that $L = \frac{D}{M}$ as bank finance their loans entire through deposits.

Banks are offering homogeneous accounts and as access is free to all cash services at any bank for their depositors and therefore competition will be driven by attracting deposits through deposit rates. Perfect competition will imply that banks make no profits from offering deposits, such that $\pi \left(1 + r_L\right) - \left(1 + r_D^i\right) - f = 0$, or

$$1 + r_D^i = \pi \left(1 + r_L\right) - f. \tag{19.30}$$

While banks compete for deposits, they do not necessarily compete with interchange fees; we rather assume that banks can jointly agree thee interchange fee. Using the deposit rate from equation (19.30), we can rewrite the profits of banks in equation (19.29) as

$$\Pi_B^i = f \frac{N_i}{N} D - C N_i. \tag{19.31}$$

Noting that $N = \sum_{j=1}^{M} N_j$, we get the first order condition for the optimal number of cash access points as

$$\frac{\partial \Pi_B^i}{\partial N_i} = f \frac{N - N_i}{N^2} D - C = 0, \tag{19.32}$$

from which we easily obtain that

$$N_i = N - \frac{C}{fD} N^2. \tag{19.33}$$

As all banks are equal, all banks will have the same number of access points such that $N = M N_i$. Multiplying equation (19.33) by M and solving for the total number of access points, N, we obtain

$$N = \frac{M-1}{M} \frac{f}{C} D. \tag{19.34}$$

Not surprisingly, the number of access points increases the higher the interchange fee, the lower the costs of providing access is, and the higher the demand from depositors. Also, the more banks are in the market competing for income from interchange fees, the more access points are available.

Using that $N_i = \frac{N}{M}$, we obtain from equation (19.31) that bank profits are given by

$$\Pi_B^i = \frac{f}{M^2} D \tag{19.35}$$

and banks seek to charge the highest possible interchange fee.

However, we require depositors willing to provide deposits for banks to make profits, thus the provision of deposits has to be profitable to them. Depositors will access cash services randomly as they need them, and we assume that the more access points exist, the lower the costs to do so are; let us assume that the distance to a service point on average is $d_i = \frac{1}{2N}$ and the costs of reaching these service points are c. We thus have the profits of depositors given by

$$\begin{aligned} \Pi_D^i &= \left(1 + r_D^i\right) D - c d_i D - D \\ &= r_D^i D - \frac{M}{M-1} \frac{c}{f} C \\ &= (\pi (1 + r_L) - 1) D - fD - \frac{M}{M-1} \frac{c}{f} C, \end{aligned} \tag{19.36}$$

where the final equality merges when inserting for the deposit rate from equation (19.30). For deposits to be provided we require that $\Pi_D^i \geq 0$, which solves for

$$f \leq f^* = \frac{1}{2}\left(\pi\left(1+r_L\right)-1\right)+\sqrt{\frac{1}{4}\left(\pi\left(1+r_L\right)-1\right)^2-\frac{M}{M-1}\frac{cC}{D}}. \quad (19.37)$$

Banks can charge higher interchange fees if they were to extract all surplus from depositors by adjusting the number of access points such that

$$\Pi_D^i = \left(\pi\left(1+r_L\right)-1\right)D - fD - cd_iD = 0. \quad (19.38)$$

Hence banks would have to provide

$$N = \frac{c}{2\left(\pi\left(1+r_L\right)-1-f\right)} \quad (19.39)$$

such access points, where we used that $d_i = \frac{1}{2N}$.

Using again that $N_i = \frac{N}{M}$, the bank profits in equation (19.31) become

$$\Pi_B^i = f\frac{D}{M} - \frac{cC}{2M\left(\pi\left(1+r_L\right)-1-f\right)} \quad (19.40)$$

and hence the optimal interchange fee is given by the first order condition

$$\frac{\partial \Pi_B^i}{\partial f} = \frac{D}{M} - \frac{cC}{2M\left(\pi\left(1+r_L\right)-1-f\right)^2} = 0. \quad (19.41)$$

This solves for the optimal interchange fee to be

$$f^{**} = \pi\left(1+r_L\right) - 1 - \sqrt{\frac{cC}{2D}} \quad (19.42)$$

and inserting this into equation (19.39), the optimal number of access points is given by

$$N = \sqrt{\frac{cD}{2C}} \quad (19.43)$$

and bank profits are then

$$\Pi_B^i = \frac{f^{**}}{M}D - \frac{1}{M}\sqrt{\frac{cC}{2D}}. \quad (19.44)$$

Whether a bank would choose interchange fee f^* or interchange fee f^{**} will depend on the profits these generate, as determined by equation (19.35) and equation (19.44), respectively.

The social optimum would seek to minimize the costs associated with access to cash services by choosing the optimal number of access points. These costs consist of the costs faced by depositors to reach these services, cd_iD, and the banks to provide access points, CN. The interchange is not relevant for the social optimum as they are only a re-distribution of wealth between banks. Thus the total costs are given by

$$\Pi_W^i = cd_i D + CN = \frac{c}{2N} D + CN, \tag{19.45}$$

which leads to the first order condition for minimum costs that

$$\frac{\partial \Pi_W^i}{\partial N} = -\frac{c}{2N^2} D + C = 0, \tag{19.46}$$

from which we easily obtain that

$$N = \sqrt{\frac{cD}{2C}}. \tag{19.47}$$

Thus the number of access points associated with an interchange fee of f^{**} would be socially optimal. For the social optimum to be chosen by banks, we would require that the profits in equation (19.44) exceed those in equation (19.35).

Solving equation (19.42) for $\pi(1 + r_L) - 1$ and inserting the resulting expression into equation (19.37), we get the relationship between the two interchange fees as

$$f^* = \frac{1}{2} f^{**} + \frac{1}{2}\sqrt{\frac{cC}{2D}} + \sqrt{\frac{1}{4}\left(f^{**} + \sqrt{\frac{cC}{2D}}\right)^2 - \frac{M}{M-1}\frac{cC}{2D}}. \tag{19.48}$$

From this relationship we can easily determine that

$$\begin{aligned}\frac{\partial f^*}{\partial f^{**}} &= \frac{1}{2} + \frac{1}{4}\frac{f^{**} + \sqrt{\frac{cC}{2D}}}{\sqrt{\frac{1}{4}\left(f^{**} + \sqrt{\frac{cC}{2D}}\right)^2 - \frac{M}{M-1}\frac{cC}{2D}}} \\ &= \frac{1}{2} + \frac{1}{4}\frac{\pi(1 + r_L) - 1}{\sqrt{\frac{1}{4}\left(f^{**} + \sqrt{\frac{cC}{2D}}\right)^2 - \frac{M}{M-1}\frac{cC}{2D}}} \\ &> 1,\end{aligned} \tag{19.49}$$

where the final inequality arises when using equation from (19.42) that $\pi(1 + r_L) - 1 = f^{**} + \sqrt{cC}2D$. We thus see that if we change the interchange fee f^{**} changes more than the interchange fee f^*.

To obtain the social optimum we require that the profits in equation (19.44) exceed those in equation (19.35), this requires

$$f^* \leq M f^{**} - \frac{M}{D}\sqrt{\frac{cC}{2D}}. \tag{19.50}$$

As $\frac{\partial f^{**}}{\partial c} = -\frac{1}{2}\sqrt{\frac{C}{cD}} < 0$ and $\frac{\partial f^{**}}{\partial C} = -\frac{1}{2}\sqrt{\frac{c}{CD}} < 0$, we see that as we increase the costs to depositors accessing cash services, c, or the costs of banks to offer an access point, C, the interchange fees reduce, but due to equation (19.49), the interchange fee f^* reduces more. This makes it more likely that the condition in equation (19.50) is fulfilled as these costs increase. As for $c = 0$ or $C = 0$ we have $f^* = f^{**} = \pi(1 + r_L) - 1$, the condition in equation (19.50) is only fulfilled for sufficiently large costs. If these costs are small, the bank will choose an interchange fee of f^*, associated with the number of access points $N = \frac{M-1}{M}\frac{f^*}{C}D$ and hence the number of access points is lower than the social optimum.

We thus see that while interchange fees give an incentive to banks to provide access points for depositors to use, the costs of these prevents them from setting up a sufficiently large number, unless the costs of not having access points close-by is high to depositors or the costs of providing these access points is high. In these cases, the banks can charge interchange fees that are sufficiently high for banks to provide a socially optimal number of access points.

Reading Donze & Dubec (2006)

19.1.3 Remote banking access

Banks offer a variety of account services and use different ways to access these. Some services can only be accessed at a bank branch as they are best conducted in person, for example if advice is sought or a meeting is arranged to discuss the specific needs of the depositor. Other services can be used either from home using remote access to accounts by seeking advice through online chats or video calling, or they might be accessed while on a mobile device at any location. Banks need to invest making services accessible remotely and will do so only if it is profitable to do so.

Let us assume we have two banks that offer differentiated accounts, which differ in a range of services the banks offer. Depositors have different preferences for such services and these are expressed by their locations on a circle, which we assume has a uniformly distributed density of depositors with their respective preferences. The two banks are located at equal distances on this circle and their locations indicate the type of services they offer. For simplicity we assume that the distance between the two banks is 1.

Depositors have three different types of interactions with banks. Firstly they have N interactions that require them to attend a bank branch and will incur costs c to attend such meetings for each unit of distance; secondly there are N^* interactions that could be conducted remotely, but only from their home location. In addition, depositors will have $\hat{N}$ interactions with their banks that can be accessed remotely at the location they are currently at. We assume that depositors may want to access these services at any location they might be, for example when travelling and thus assume they are located randomly on the circle. With a circle of total length 2, due to the two banks being 1 unit part, their average distance to their bank is 1 and if they were not able to have remote access they would face costs cD to access these services for deposits of size D. Those services that cannot accessed remotely will only be demanded if the depositor is not travelling, and his position will have a distance d_i from banks i, such that his costs will be cd_iD.

We can now assess the profits of banks providing remote access to depositors for some of their services and compare these when no such access is provided.

No remote access With banks paying a deposit rate of r_D^i, the profits of depositors whose bank does not allow remote access are given by

$$\Pi_D^i = \left(1 + r_D^i\right) D - (N + N^*)\, cd_iD - \hat{N}cD. \tag{19.51}$$

Depositors will prefer bank i over the other bank, bank j, if $\Pi_D^i \geq \Pi_D^j$. Using that $d_j = 1 - d_i$ as the distance between the two banks is 1, we get that depositors choose bank i if their distance if less than

$$d_i = \hat{d}_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c\,(N + N^*)}, \tag{19.52}$$

where $\hat{d}_i$ denotes the market share on the other side of the bank. Using that banks profits are given by

$$\begin{aligned} \Pi_B^i &= \pi\,(1 + r_L)\, L - \left(1 + r_D^i\right) D_i \\ &= \left(\pi\,(1 + r_L) - \left(1 + r_D^i\right)\right) \left(d_i + \hat{d}_i\right) D, \end{aligned} \tag{19.53}$$

using that loans are fully financed by deposits to the banks, D_i and the deposits are determined from $D_i = \left(d_i + \hat{d}_i\right) D$ as the market share of those depositors that chose banks i. Here π denotes the probability of the loan being repaid and r_L the loan rate. Inserting from equation (19.52) for the market share of the bank, we can obtain the optimal deposit rate the bank will offer through the first order condition $\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} = 0$, which solves for

$$1 + r_D^i = \frac{1}{2}\left(\pi\,(1 + r_L) - c\,(N + N^*) + \left(1 + r_D^j\right)\right). \tag{19.54}$$

Similarly we get for the other bank with profits $\Pi_B^j = \left(\pi\,(1 + r_L) - \left(1 + r_D^j\right)\right)\left((1 - d_j) + \left(1 - \right.\right.$ that

$$1 + r_D^j = \frac{1}{2}\left(\pi\,(1 + r_L) - cN + \left(1 + r_D^i\right)\right). \tag{19.55}$$

Combining equations (19.54) and (19.55) we get that

$$1 + r_D^i = 1 + r_D^j = \pi\,(1 + r_L) - c\,(N + N^*). \tag{19.56}$$

Inserting equations (19.52) and (19.56) into equation (19.53), we get the bank profits without remote access as

$$\Pi_B^B = c\left(N + N^*\right) D. \tag{19.57}$$

We see that the deposit rate and bank profits are not affected by the interactions that could be conducted remotely, although this facility is not offered, as all depositors and banks are affected equally by its costs and the competition between banks remains unaffected.

Both banks offering remote access If both banks offer remote access to the $\hat{N}$ and N^* interactions where this is possible, depositors do not need to travel to the bank for these and will incur no costs. Thus depositor profits are given by

$$\Pi_D^i = \left(1 + r_D^i\right) D - Ncd_i D. \tag{19.58}$$

Again, depositors prefer using banks i over banks j if $\Pi_D^i \geq \Pi_D^j$, from which we obtain

$$d_i = \hat{d}_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2cN}. \tag{19.59}$$

The profits of banks remain given by equation (19.53) and after inserting from equation (19.59), the first order condition for the optimal deposit rate, $\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} = 0$ solves for

$$1 + r_i^D = \frac{1}{2}\left(\pi\left(1 + r_L\right) - cN + \left(1 + r_D^j\right)\right) \tag{19.60}$$

and similarly for the other bank

$$1 + r_D^j = \frac{1}{2}\left(\pi\left(1 + r_L\right) - cN + \left(1 + r_D^i\right)\right) \tag{19.61}$$

such that when combining these two deposit rates we obtain that

$$1 + r_D^i = 1 + r_D^j = \pi\left(1 + r_L\right) - cN. \tag{19.62}$$

Inserting all results from equations (19.59) and (19.62) into the banks profits of equation (19.53), we obtain

$$\Pi_B^R = cND \leq \Pi_B^B. \tag{19.63}$$

The profits here are lower as competition between banks is increased due to depositors having to rely less on accessing bank branches, which is costly for depositors. Thus banks would prefer to not offer remote access to depositors as this increases competition between the banks.

One bank offering remote access Finally let us assume that bank i offers remote access and bank j only branch access. The profits of depositors with bank i offering remote access are given by equation (19.58) and for those with bank j not offering remote access are given by equation (19.51). Depositors prefer bank the bank offering remote access of $\Pi_D^i \geq \Pi_D^j$, from which we obtain that

$$d_i = \hat{d}_i \leq d_i^* = \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right) + c\left(N + N^* + \hat{N}\right)}{c\left(2N + N^*\right)}. \tag{19.64}$$

The bank profits are again given by $\Pi_B^i = \left(\pi\left(1 + r_L\right) - \left(1 + r_D^i\right)\right)\left(d_i + \hat{d}_i\right) D$ and $\Pi_B^j = \left(\pi\left(1 + r_L\right) - \left(1 + r_D^j\right)\right)\left(\left(1 - d_i\right) + \left(1 - \hat{d}_i\right)\right) D$, for the bank enabling remote access and the bank not enabling remote access, respectively. The first order conditions from banks maximizing their profits over their optimal deposit rates yield these deposit rates as

$$\begin{aligned} 1 + r_D^i &= \pi\left(1 + r_L\right) - \frac{1}{3}c\left(3N + 2N^* + \hat{N}\right), \\ 1 + r_D^j &= \pi\left(1 + r_L\right) - \frac{1}{3}c\left(3N + N^* + \hat{N}\right). \end{aligned} \tag{19.65}$$

We clearly see that $1 + r_D^j \leq 1 + r_D^i$ as the ability of remote access that bank i offers reduces the costs of their depositors and hence they are willing to accept a lower deposit rate.

Inserting these expressions into the bank profits, we obtain

$$\begin{aligned} \hat{\Pi}_B^R &= 2\frac{\left(N + \frac{2}{3}N^* + \frac{1}{3}\hat{N}\right)^2}{2N_B + N_O}cD, \\ \hat{\Pi}_B^B &= 2\frac{\left(N + \frac{1}{3}N^* - \frac{1}{3}\hat{N}\right)^2}{2N_B + N_O}cD. \end{aligned} \tag{19.66}$$

As $\hat{\Pi}_B^B \leq \hat{\Pi}_B^R$ we see that the profits of the bank offering remote access are increased compared to a bank not offering this access due to the lower deposit rate they can pay and the additional deposits their remote access attracts.

Having explored the profits bank make from providing remote access to depositors and not providing such access, we can now examine the equilibrium decision of the two banks.

Equilibrium decisions on remote access Banks are involved in a strategic game to introduce remote access. We know from our results above that both banks not offering remote access to depositors provides them with higher profits than both banks offering remote access due to increased competition. However, we have also seen, that the bank introducing remote access to its depositors will make higher profits than the bank not doing so, as they gain a competitive advantage. This might provide them with an incentive to introduce remote access and increase profits at the expense of the other bank, which would then react by introducing remote access too.

Let us first derive the condition that the profits of the bank offering remote access as the only bank exceeds the profits of the bank if both banks do not offer remote access, $\hat{\Pi}_B^R \leq \Pi_B^B$. This is the case if

$$\hat{N} \leq \hat{N}^* = 3\left(\sqrt{\frac{1}{2}\left(2N + N^*\right)\left(N + N^*\right)} - \left(N + \frac{2}{3}N^*\right)\right). \tag{19.67}$$

Similarly we obtain that the bank not offering remote access, while the other bank does, exceeds those when both banks offer remote access, $\hat{\Pi}_B^B \leq \Pi_B^R$. This solves for

$$\hat{N} \leq \hat{N}^{**} = 3\left(\left(N + \frac{1}{3}N^*\right) - \sqrt{\frac{1}{2}N\left(2N + N^*\right)}\right), \tag{19.68}$$

where we easily see that $\hat{N}^* \leq \hat{N}^{**}$.

Figure 19.3 shows the profits banks obtain in this strategic game to introduce remote access to their depositors. We observe that the equilibrium is for both banks to offer remote access if $\hat{\Pi}_B^R > \Pi_B^B$ and $\hat{\Pi}_B^B > \Pi_B^R$, i.e. from equations (19.67) and (19.68) if $\hat{N} > \hat{N}^{**}$. In the case that $\Pi_B^B > \hat{\Pi}_B^R$ and $\hat{\Pi}_B^B > \Pi_B^R$, both banks will rely only on branch-based banking and offer no remote access. These conditions from equations (19.67) and (19.68) imply that $\hat{N} < \hat{N}^*$. Finally, if $\hat{\Pi}_B^R \geq \Pi_B^B$ and $\hat{\Pi}_B^B \geq \Pi_B^R$, corresponding to $\hat{N}^* \leq \hat{N} \leq \hat{N}^{**}$, one bank will offer remote access, while the other bank will not.

		Bank j	
		No remote access	Remote access
Bank i	No remote access	Π_B^B, Π_B^B	$\hat{\Pi}_B^B, \hat{\Pi}_B^R$
	Remote access	$\hat{\Pi}_B^R, \hat{\Pi}_B^B$	Π_B^R, Π_B^R

Fig. 19.3: Strategic game to provide remote access

Hence we see that for low demand of remote transaction while travelling, $\hat{N}$, banks will not offer this facility. Even though we assume here that there are no costs in giving remote access, the increased competition for depositors reduces the profits of banks. Once the demand for remote interactions increases, only one bank will offer remote access initially such that competition does not increase too much. Only with high demand will the loss in market share of the bank not offering remote access, induce them to also offer remote access, despite the increased competition.

Even though giving remote access to depositors is increasing competition between banks as it reduces the level of differentiation in branch services that banks can provide, it may be introduced due to competitive pressures. If there is sufficient demand by depositors for such access, some bank will introduce remote access with the aim of securing additional market share, which compensates them for the increased competition. Those banks that have not introduced remote access will face a reduced market share and increased competition, leaving them with lower profits, but they will not introduce remote access as this would increase competition between banks further and them increasing their market share would not compensate for this effect on competition. It is only once the demand for remote access is sufficiently high that the losses to the bank offering remote access are that significant, that the remaining banks would offer remote access, too. Regaining some of the market share from the banks having introduced remote access earlier, will compensate them for the increased competition.

Remote access to depositors is not introduced because it benefits banks, on the contrary, it will increase competition and thus reduce their profits, but as the consequence of banks seeking a competitive advantage over other banks. The result of such attempts to gain market share is that all banks introduce remote access once the demand is sufficiently high, but would make higher profits if not doing so.

It will be smaller banks that introduce new services and innovations, such as remote access to accounts, in order to gain market share from the more established banks. These larger banks will initially not react to the emergence of new services as they seek to not increase competition unduly; as long as the demand for such services is low, their loss of deposits will be very limited. It is only once demand increases that the potential loss of market share becomes relevant and they will introduce these services themselves. While this will increase the overall competition, it will enable these banks to stop the loss of market share to more innovative banks, and even regain some of the deposits they have lost.

Reading Bouckaert & Degryse (1995)

19.1.4 Account fees

Banks provide a wide range of services with their deposit accounts, most notably the ability to make payments through the transfer of funds to other accounts, the withdrawal of cash, or the ability to use payment cards. Other ancillary services may include the provision of insurance for purchases made using payment cards, or the ability to access financial advice. In many cases such services are provided free, while in other cases banks may charge a fee for maintaining an account.

Let us assume that depositors have preferences for specific services and these preferences are identified by its position along a line of length one, at whose ends each a bank is located offering differentiated account services; depositors are distributed uniformly along this line. By using the accounts offered by either bank, the depositor will lose utility proportional to its distance from the bank and it will lose c when at a distance of 1 unit.

Banks charge a fee f_i for maintaining the account and companies may need to obtain a loan L with probability p. If they obtain a loan, it is repaid with probability

π and the investment gives a return R, while the loan rate the bank charges is r_L^i. Hence company profits, when using bank i, are given by

$$\Pi_C^i = \left(p\pi\left((1+R)-\left(1+r_L^i\right)\right)-cd_i-f_i\right)L, \tag{19.69}$$

where d_i measures the distance of the depositor to bank i and $d_j = 1-d_i$ denotes the distance to the other bank. A company will prefer bank i over bank j if it generates them higher profits, thus $\Pi_C^i \geq \Pi_C^j$, which solves for

$$d_i \leq d_i^* = \frac{1}{2} + \frac{f_j - f_i + p\pi\left(\left(1+r_L^j\right)-\left(1+r_L^i\right)\right)}{2c}. \tag{19.70}$$

The bank will provide all loans to depositors closer to them than d_i^*, representing their market share of all loans L. With a deposit rate r_D and the assumption that deposits fully finance loans, we obtain the bank profits as

$$\Pi_B^i = \left(p\left(\pi\left(1+r_L^i\right)-(1+r_D)\right)+f_i\right)d_i^*L, \tag{19.71}$$

Inserting from equation (19.70) for the market share of the bank, we get the first order conditions for the optimal fee and loan rate as

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial f_i} &= d_i^*L - \frac{p\pi\left(1+r_L^i\right)-p\left(1+r_D\right)+f_i}{2c}L = 0 \\ \frac{\partial \Pi_B^i}{\partial\left(1+r_L^i\right)} &= p\pi d_i^*L - \frac{p\pi\left(1+r_L^i\right)-p\left(1+r_D\right)+f_i}{2c}p\pi L = 0, \end{aligned} \tag{19.72}$$

which are both identical conditions once we divide the second equation by $p\pi$. Hence we will not be able to identify a single solution for the optimal loan rate and account fee, but only a relationship between them.

For bank j we get the same result, noting that $d_j = 1-d_i$, hence we need to solve the first order conditions

$$\begin{aligned} 2cd_i^* - p\pi\left(1+r_L^i\right)-f_i+p\left(1+r_D\right) &= 0, \\ 2c\left(1-d_i^*\right)-p\pi\left(1+r_L^j\right)-f_j+p\left(1+r_D\right) &= 0. \end{aligned} \tag{19.73}$$

Setting these two expressions equal and inserting for d_i^* from equation (19.70), we can obtain a relationship between the two account fees, which becomes

$$f_j = f_i - p\pi\left(\left(1+r_L^j\right)-\left(1+r_L^i\right)\right). \tag{19.74}$$

Using this expression in equation (19.70) for the market share of bank i and inserting this into the first line of the first order condition in equation (19.73), we obtain

$$f_i = c - p\left(\pi\left(1+r_L^i\right)-\left(1+r_D\right)\right). \tag{19.75}$$

The equivalent strategy is available for the other bank.

We observe a close trade off for banks between charging high loan rates and lower account fees. A bank may charge a low loan rate r_L, which then allows them to charge a higher account fee f_i. Furthermore, the two banks can employ different pricing strategies as in all cases that equation (19.75) is fulfilled, we have $d_i^* = \frac{1}{2}$, giving both banks equal market shares, and the bank profits are maximal. We also see easily that higher deposit rates will increase the fee charged, as would a higher market power, as measured by the costs imposed on depositors not obtaining their preferred account services, c. If companies are more likely to demand a loan, as measure by p, the fees are lower and more risky borrowers, those with a lower chance of success π, will increase the account fee. Overall we observe that if the bank has higher income from providing loans, either because they are more frequently demanded, p, more likely to be repaid, π, or a higher loans rate charged, r_L, the fee will be lower. This is to attract more depositors who might provide banks with such income. Increased market power, c will allow banks to generate higher profits, allowing it to charge higher account fees and higher deposit rates, r_D, reduce the profits of banks, for which they are compensated through a higher account fee.

Thus, in an economy with mostly safe borrowers, that borrow frequently and where deposit rates are low, we should observe lower account fees than in economies where loans are more risky and there is less demand for such loans. Of course, individual banks can vary the account fee by changing the loan rate, if we assume that companies cannot easily switch banks if they require a loan. Hence we might see different account fees charged by competing banks, but also economies with different characteristics might have similar account fees, because banks decide on a different allocation between account fees and loan rates.

Reading Thanassoulis & Vadasz (2021)

19.1.5 Minimum account balances

A common feature of many accounts is that banks do not charge fees if the account balance is above a certain threshold; for any smaller deposits banks in many cases charge a fee. We will investigate here why such a practice exists and why it is beneficial for banks to charge only the holders of accounts with small fees.

We assume that there are two type of accounts, a fraction λ of the N accounts holds a large deposit, D_L, while a fraction $1 - \lambda$ of the accounts hold a small deposit $D_L < D_S$. Banks are engaged in monopolistic competition by offering differentiated accounts and are positioned at the end of line with length 1; their depositors are uniformly distributed along this line and their position indicates their preferences for services by banks. the further away they are from a bank, the less does the bank's services meet their demands. This distance is used as a measure for the loss in utility as the bank they choose does not offer the services they seek. We assume that the marginal costs of deviating from their optimal services are c. If banks the two banks are charging a fee of F_i, the loss in utility of a depositor using banki is given by

$$\Pi_C^i = -cd_i - F_i, \tag{19.76}$$

where d_i denotes the distance to bank i. we note that as the banks are having a distance of 1 and depositors are located between them that $d_1 + d_2 = 1$. A depositor is indifferent between choosing the two banks if $\Pi_C^1 = \Pi_C^2$, which solves for

$$d_1^* = \frac{1}{2} + \frac{F_2 - F_1}{2c}. \tag{19.77}$$

Any depositor closer than d_1^* to bank 1 will choose this bank and all other depositors will choose bank 2. Thus the demand for deposits at bank 1 is given by $D_L^1 = d_1^* \lambda N D_L$ and $D_S^1 = d_1^* \lambda N D_S$ for large and small deposits respectively. We see that bank 1 has a fraction d_1^* of the market, where a fraction λ $(1-\lambda)$ of deposits is large (small); each deposit if D_L (D_S) and there are a total of N depositors.

We can now assess the fees charged and profits generated to banks for different ways banks can charge fees. We commence by firstly considering the case where all depositors are charged a fee.

All accounts charged fees If we assume that banks generate profits μ from deposits, having taken into account any costs, then their profits are given by this revenue, in addition to the fee income. This fee income consists of the fee R_i, which is paid by all those $d_i^* N$ depositors that choose this bank. Focussing on bank 1, its profits are given by

$$\Pi_B^1 = \mu \left(D_L^1 + D_S^1\right) + d_1^* F_1 N. \tag{19.78}$$

We can now insert for all expressions from the terms determined above and then obtain the first order condition for the optimal fee that bank 1 should charge as

$$\begin{aligned} \frac{\partial \Pi_B^1}{\partial F_1} = &-\frac{\mu\left(\lambda D_L + (1-\lambda) D_S\right)}{2c} N \\ &+ \left(\frac{1}{2} + \frac{F_2 - 2F_1}{2c}\right) N = 0. \end{aligned} \tag{19.79}$$

If we focus only on symmetric equilibria such that $F_1 = F_2$, this solves for the optimal fee to be

$$F_i = c - \mu\left(\lambda D_L + (1-\lambda) D_S\right). \tag{19.80}$$

Inserted into the profits of the bank, we easily obtain that

$$\Pi_B^i = \frac{1}{2} cN. \tag{19.81}$$

These profits for the banks we can now compare tho that where not all depositors are charged a fee. We initially consider the case where only large depositors are charged a fee.

Only large deposits charged fees If a bank charges only its large depositors a fee, we obtain for small deposits with $\hat{F}_1 = \hat{F}_2 = 0$ that $d_1^* = \frac{1}{2}$ and hence the demand for small deposits for bank 1 is given by $D_S^1 = \frac{1}{2}(1-\lambda) D_S N$, while the threshold for large depositors remains as given in equation (19.77) and hence the large deposits at bank 1 are given by $D_L^1 = d_1^* \lambda D_L N$. The fee income is now only generated by the large depositors and hence it is given by $d_1^* \hat{F}_1 \lambda N$.

Inserting these finding into the profits for bank 1, we obtain in analogy to equation (19.78) that

$$\hat{\Pi}_B^1 = \mu \left(d_1^* \lambda D_L + \frac{1}{2}(1-\lambda) D_S \right) N + d_1^* \hat{F}_1 \lambda N. \tag{19.82}$$

After inserting for d_i^* from equation (19.77), we obtain the first order condition for the optimal account fee as

$$\frac{\partial \hat{\Pi}_B^1}{\partial \hat{F}_1} = -\frac{\mu \lambda D_L}{2c} N + \left(\frac{1}{2} + \frac{\hat{F}_2 - 2\hat{F}_1}{2c} \right) \lambda N = 0. \tag{19.83}$$

Again focussing on symmetric equilibria with $\hat{F}_1 = \hat{F}_2$, we can solve this condition for the optimal fee

$$\hat{F}_i = c - \mu D_L, \tag{19.84}$$

giving the bank a profit of

$$\hat{\Pi}_B^i = \frac{1}{2}\lambda c N + \frac{1}{2}\mu (1-\lambda)(D_S - D_L) N < \Pi_B^i. \tag{19.85}$$

The inequality follows from the observation that we assumed $D_S < D_L$ and $\lambda < 1$. It is therefore that it is never optimal for banks to charge a fee only to large depositors.

We can now consider the case where banks charge a fee only to small depositors.

Only small deposits charged fees If a bank charges only its small depositors a fee, we obtain for large deposits with $\hat{F}_1 = \hat{F}_2 = 0$ that $d_1^* = \frac{1}{2}$ and hence the demand for large deposits for bank 1 is given by $D_L^1 = \frac{1}{2}\lambda D_S N$, while the threshold for small depositors remains as given in equation (19.77) and hence the small deposits at bank 1 are given by $D_L^1 = d_1^* (1-\lambda) D_L N$. The fee income is now only generated by the small depositors and hence it is given by $d_1^* \hat{\hat{F}}_1 (1-\lambda) N$.

Inserting these finding into the profits for bank 1, we obtain in analogy to equation (19.78) that

$$\hat{\hat{\Pi}}_B^1 = \mu \left(d_1^* (1-\lambda) D_S + \frac{1}{2}\lambda D_L \right) N + d_1^* \hat{\hat{F}}_1 (1-\lambda) N. \tag{19.86}$$

After inserting for d_i^* from equation (19.77), we obtain the first order condition for the optimal account fee as

$$\frac{\partial \hat{\hat{\Pi}}_B^1}{\partial \hat{F}_1} = -\frac{\mu (1-\lambda) D_S}{2c} N + \left(\frac{1}{2} + \frac{\hat{\hat{F}}_2 - 2\hat{\hat{F}}_1}{2c} \right) (1-\lambda) N = 0. \tag{19.87}$$

Again focussing on symmetric equilibria with $\hat{\hat{F}}_1 = \hat{\hat{F}}_2$, we can solve this condition for the optimal fee

$$\hat{\hat{F}}_i = c - \mu D_S, \tag{19.88}$$

giving the bank a profit of

$$\hat{\hat{\Pi}}_B^i = \frac{1}{2} (1-\lambda) cN + \frac{1}{2} \mu \lambda D_L N. \tag{19.89}$$

We can now determine whether the bank prefers to charge only the small depositors, giving it profits $\hat{\hat{\Pi}}_B$ or charging all depositors and obtaining Π_B^i. The profits of charging only small depositors are higher if

$$c < \mu D_L. \tag{19.90}$$

As we can interpret c as the degree of market power a bank has, we see that as long as the market power of banks is not too strong, it is optimal for banks to charge a fee only for small depositors. It is straightforward to see that $\hat{\hat{F}}_i > F_i$ and hence smaller depositors are charged a higher fee than if all depositors were charged. The reason for the higher fee here is that banks seek to charge fees from only fewer depositors and thus in order to increase their profits, will have to charge a higher fee from each depositor.

With a high degree of competition between banks, small c, the fees will be relatively low and hence the loss of fee income from large deposits is not important for their overall profits; however, the attraction of large deposits, increasing their revenue from deposits μ. With higher market power, banks can extract more fees from depositors and therefore the loss of fees from large depositors weights more than the loss of income from the deposits themselves. Small depositors are too small to provide sufficient revenue from their deposits and hence charging only large depositors is never optimal.

Summary we have seen that it is optimal for banks to waive account fees for large deposits; this allows banks to attract these deposits and generate profits by financing loans, which in a situation where fees are low due to intense competition between banks are low and thus cannot compensate for any lost deposits due to competition. This only changes in less competitive markets where the fee income become more important to banks. However, the consequence is that small depositors are charged a higher fee than they would be if all depositors were paying fees,

Reading Shy (2024)

Résumé

Banks provide additional services to their depositors, beyond paying interest on their deposits. The account held by depositors can be used to make payments and access cash, but may also be the gateway to other form of advice. Often it can be beneficial for depositors if they can access these services not only at their own bank, but also at other banks. A prime example for this facility is the withdrawal of cash from cash machines, regardless of which bank operates the specific machine, as this saves effort in locating a cash machine of the depositors' own bank and then seeking to attend this machine, rather than the nearest. Thus depositors may benefit from banks cooperating with each other to allow depositors such access. However, such access agreements increase competition between banks as it makes the distinctive elements of a bank, for example its extensive network of cash machines, less important, leading to more direct competition through deposit rates. A similar argument can be made for banks allowing remote access to their services, such as through online banking. However, competition to attract depositors may induce banks to offer such services, even though it is detrimental to their profits.

While depositors are rarely required to pay for accessing such services, banks charge often charge each other a fee for providing the services to depositors of other banks. Such interchange fees can also affect the provision of services to depositors as the fee income provides an incentive to expand these, improving the welfare of depositors. Without interchange fees, or very low interchange fees, the costs in providing services would limit the extend of providing services to depositors from other banks, causing their provision to be well below the social optimum.

Often deposit accounts are provided free of charge or fees are charged that do not cover the costs of providing the range of services. Banks can afford to do so by recovering their costs through higher loan rates, or lower deposit rates. Overall, depositors will be indifferent in what combination of account fees and interest charges they pay the bank; this may lead to different pricing strategies of banks where some banks may offer accounts at low costs but also charge high loan rates or pay low deposit rates, while other banks would charge higher account fees, but are more attractive on interest rates.

19.2 Payment cards

Payments can traditionally be made by cash or bank transfer and bank transfer is the main payment for between businesses and increasingly also between individuals. While cash transactions have been used mainly between individuals and companies, especially in the retail industry, the use of payment cards has become the most widely used form for such payments. Payment cards, more commonly known as debit and credit cards, are issued by banks to their customers and are operated by organisations that ensure the payments made by a customers reaches its intended recipient. In order to be able to use a payment card, the merchant must subscribe to the specific network that operates the card held by their customer.

With debit cards, the transaction is instantly taken out of the deposit account and if the balance is not sufficient, the transaction would be declined. With credit cards the transaction is not immediately taken out of the deposit account, but the purchaser is granted a loan by the card issuer until payment is due. In chapter 19.2.1 we will look at incentives for banks to provide such credit cards and merchants to accept them rather then rely on cash payment. We we do not look at the competition between different payment cards, but focus on a single card and the incentives for its adoption only. How the organisation administering the payment card charges banks for their services is analysed in chapter 19.2.2, focussing the interchange fee that banks are charged.

19.2.1 Issuing credit cards

While we can interpret debit cards as a close substitute for cash payments, credit cards provide the purchaser effectively with a loan until the payment on his card becomes due. At least until the upcoming monthly payment of the credit card is due, this loan is typically provided free of any interest charges, it is only in cases where the balance is not repaid in full that interest will be charged. Banks issuing such cards are normally not charging a fee for the use of credit or debit cards, except when they add additional benefits to the card such as insurance cover, but only charge merchants a fee for the ability to accept them as payment.

Let us assume that depositors seeking to use a credit card have an income W_t in time period t and they seek to purchase goods to the value of P in each time period from which they gain value V. We assume that the income of consumers is uncertain and has a distribution $G(W_t)$. Depositors finance their purchases from their income in each of the two time periods or they can use a credit card for payment in time period 1, where the balance has to be repaid in time period 2. Depositors are not charged interest on their purchases using the credit card and for simplicity assume that deposits do not pay any interest.

Credit card use by depositors If the depositor were to use cash as payment in time period 1, they could purchase the good if $W_1 \geq P$ and would retain the amount of $W_1 - P$ for purchases in time period 2, in addition to their income W_2. Hence they are able to make a purchase in time period 2 if $(W_1 - P) + W_2 \geq P$ if they purchased the good in time period 1, $W_1 \geq P$, and if $W_1 + W_2 \geq P$ if no purchase was made in time period 1, $W_1 < P$; these two conditions are identical. This gives us a utility of the depositor using cash of

$$\begin{aligned}
\Pi_D &= \text{Prob}\,(W_1 \geq P)\,(V - P) \\
&\quad +\text{Prob}\,(W_1 + W_2 \geq 2P)\,(V - P) \\
&= (2 - G(P) - G(2P - W_1))\,(V - P)\,.
\end{aligned} \tag{19.91}$$

If the depositor uses a credit card, he can consume with certainty in time period 1 and in time period 2 he can consume if his income from both time periods, $W_1 + W_2$ allows him to repay the credit card balance, P, and make a purchase in time period 2. Hence the utility of the depositor is given by

$$\begin{aligned} \hat{\Pi}_D &= (V - P) + \text{Prob}\,(W_1 + W_2 \geq 2P)\,(V - P) \\ &= (2 - G\,(2P - W_1))\,(V - P)\,. \end{aligned} \tag{19.92}$$

It is now easy to see that $\Pi_D < \hat{\Pi}_D$ and therefore depositors strictly prefer using credit cards over cash payments.

Merchants accepting credit cards Let us propose that banks issue credit cards to any depositor with a sufficiently high income of $W_1 > W^*$ and they charge merchants a fee of F to accept credit cards for a payment. For simplicity assume that merchants face no costs of selling their goods such that they retain the purchase price P fully.

If merchants do not accept credit cards, they can only sell their goods in time period 1 if $P < W_1$, which happens with probability $1 - G\,(P)$. If he accepts cards, all those depositors that can obtain a credit card will purchase the good, that is all those with an income of $W_1 > W^*$; thus a purchase happens with probability $1 - G\,(W^*)$. As we will see below when discussing the bank's decision to issue a credit card to depositors, we will have $W^* < P$ and hence when accepting credit cards, the merchant is more likely to make a sale.

If the depositor makes a purchase in time period 1, he will be able to make a purchase in time period 2 if the combined income from both time periods is sufficient for two purchases. If he did not make a purchase in time period 1, the combined income only needs to be sufficient for this single purchase. We assume that in time period 2, purchases are made using cash payments only; as any credit card bills have to be settled in time period 2 there is no inherent advantage to using credit cards. The profits of the merchant, taking into account that for credit card payments he is charged a fee F, for cash and credit card payments, respectively, are given by

$$\begin{aligned} \Pi_C &= (1 - G\,(P))\,P \\ &\quad + (1 - G\,(P))\,(1 - G\,(2P - W_1))\,P \\ &\quad + G\,(P)\,(1 - G\,(P - W_1))\,P, \\ \hat{\Pi}_C &= (1 - G\,(W^*))\,(P - F) \\ &\quad + (1 - G\,(W^*))\,(1 - G\,(2P - W_1))\,P \\ &\quad + G\,(W^*)\,(1 - G\,(P - W_1))\,P. \end{aligned} \tag{19.93}$$

Merchants will accept credit card payments if it is more profitable to do so, $\Pi_C \geq \hat{\Pi}_C$, which solves for

$$F \leq F^* = \frac{G\,(P) - G\,(W^*)}{1 - G\,(W^*)}\,(1 - G\,(2P - W_1) + G\,(2P - W_1))\,P. \tag{19.94}$$

If the fee charged by banks for being able to accept credit card payments is not too high, merchants are accepting these as it will increase their sales. If the income in time period 1 is low, depositors cannot make cash payments and the merchant would lose these sales, while when accepting credit card payments, the sale could commence. Even if the income in time period 2 would be high, the depositor would not purchase two units of the good to compensate for not making purchase in time period 1, resulting in lost sales to the merchant. If the income does not recover sufficiently in time period 2, the depositor would not be able to make another purchase, but the merchant has already secured a sale in time period 1. Thus the acceptance of credit card payments increases the merchant's sale and as long as the fee the bank charges is not too high, he will increase his profits.

Banks issuing credit cards Banks issuing credit cards will be concerned about the ability of depositors to repay their purchase. Having made a purchase P, the balance of the credit card can be repaid as long as $W_1 + W_2 \geq P$.

As the credit card is only issued after the bank observes the income in time period 1, W_1, banks can use this information when deciding who is issued a credit card. Let us define

$$G\left(W|W_1\right) = Prob\left(W_1 + W_2 \leq W\right) = G\left(W - W_1\right) \tag{19.95}$$

and we then obtain

$$G\left(W|W_1 > W^*\right) = \int_W^{+\infty} G\left(W - W_1\right) \frac{dG\left(W_1\right)}{1 - G\left(W\right)} \tag{19.96}$$

from the definition of conditional probability.

If we assume that banks provide depositors with credit cards if their income in the first time period exceeds a certain threshold W^*, thus $W_1 > W^*$, we can determine their profits. As when using a car the purchase in time period 1 commences with certainty, the bank will obtain the fee income F and make a payment P to the merchant. The bank then recovers the purchase price fully from the depositor if $W_1 + W_2 \geq P$, which occurs with probability $1 - G\left(P|W_1 > W^*\right)$; it recovers the purchase price partially if $W_1 + W_2 < P$. In this case the bank would obtain the full income of the depositor, noting that credit cards were only issued for incomes of $W_1 > W^*$. We thus have the profits of banks given as

$$\begin{aligned} \Pi_B = {} & \left(1 - G\left(W^*\right)\right)\Big(\left(F - P\right) + \left(1 - G\left(P|W_1 > W^*\right)\right)P \\ & + \int_{\min\{W^*;P\}}^{P} \left(W_1 + W_2\right) dG\left(W_1 + W_2|W_1 > W^*\right)\Big), \end{aligned} \tag{19.97}$$

Firstly we see that for the fee the bank charges merchants, we find

$$\frac{\partial \Pi_B}{\partial F} = \left(1 - G\left(W^*\right)\right) > 0 \tag{19.98}$$

and hence in order to maximize their profits, banks charge the maximal fee possible by extracting all surplus from the merchant, making the constraint on merchants accepting credit card payments in equation (19.94) an equality and the fee charged is F^*. We then get after inserting from equation (19.94) for the fee, that the profits are changing with the threshold for issuing credit cards according to

$$\frac{\partial \Pi_B}{\partial W^*} = g\left(W^*\right)\Bigg(-P + PG\left(P - W^*\right) + \int_{\min\{W^*;P\}}^{P} \left(W_1 + W_2\right) dG\left(W_1 + W_2 - W^*\right)\Bigg). \tag{19.99}$$

If $P \leq W^*$ then the income of depositors is always sufficient to repay the credit card balance as the income from time period 1 alone would be sufficient, given we require that $W_1 > W^*$. In this case, the final term in equation (19.99) becomes zero and $G\left(P - W^*\right) = 0$, hence

$$\frac{\partial \Pi_B}{\partial W^*} = -G\left(W^*\right) P < 0, \tag{19.100}$$

and hence the optimal threshold for providing credit cards would be the lowest possible threshold, $W^* = 0$. However, as we had assumed that $P \leq W^*$ and the price would reasonably be positive, the condition that $P \leq W^*$ cannot be met. Hence we need $P > W^*$ and depositors may not always be able to repay their credit card balances. Such defaults by depositors are paid out of the fee income of the banks.

In equation (19.99), at $W^* = P$ the expression is negative if $P > 0$, hence we decrease W^* until equation (19.99) is equal to zero. We also see that the larger the purchase price P is, the smaller the first two terms become and hence the lower the threshold W^* will become. Thus higher purchase prices will lower the standards for issuing cards. This is because from (19.94) we have

$$\frac{\partial F^*}{\partial W^*} = -g\left(W^*\right) \frac{1 - G\left(P\right)}{\left(1 - G\left(W^*\right)\right)^2} P < 0, \tag{19.101}$$

implying that with lower standards for issuing credit cards, a lower threshold W^*, increases the fee income. This is arising from the merchant making more profits from higher sales to low-income depositors; this increased fee income is then used to offset the loses from defaulting depositors. The higher price of a good, relative to the income of depositors, reduces sales with cash payments and a lower threshold for issuing credit cards allows merchants to increase their sales, which enables them to pay higher fees to banks, which in turn compensates them for any losses from depositors not being able to repay their credit card balance.

Summary We have thus seen that banks are issuing credit cards to depositors and employ a threshold in terms of their income that allows for defaults on their purchases. As banks do not charge interest on these loans, they recover their losses by charging merchants a fee for each purchase made using a credit card. Merchants benefit from being able to sell to depositors that have a too low income to afford their goods in time period 1 and using credit card payments will increase their sales.

Not charging interest on the loan they provide depositors with, despite defaults on these loans occurring, is nevertheless optimal for banks as merchants compensate them for these risks through the fee they are paying.

The provision of credit cards overcomes the inefficiency of depositors having a temporarily low income that would ordinarily impact negatively on the ability to purchase goods. If the consumption of goods is limited per time period, depositors cannot compensate for this shortage by increasing their consumption at a later stage. Here credit cards can increase welfare by allowing depositors a steady consumption; any risks arising from depositors not being able to pay their credit card balance will be covered by fees charged to merchants benefitting from higher sales, allowing banks to make profits from the issue of credit cards.

Reading Chakravorti & To (2007)

19.2.2 Interchange fees for card payments

Payment card issuers, for debit as well as credit cards, commonly do not only charge merchants, or much less commonly consumers, a fee for their use, but they also charge fees for processing payments between the bank of the merchant and that of their customers, the bank issuing the card. This fee for processing the payments between banks is referred to as the interchange fee. Such interchange fees are levied by the bank of the depositor on the bank of the merchant, thus the bank making the payment charges a fee to the bank receiving the payment. Thus by issuing payment cards, banks make profits from fees charged to merchants, and potentially depositors, and from administering payments between depositors and merchants.

Let us assume that banks issuing a payment card to its depositors charge them a fee F_D when using their payment card and for making payments to the bank of the merchant, they charge an interchange fee F_I. We assume issuing banks face no costs when making payments. Hence the issuing bank makes profits per transaction of

$$\Pi_B^D = F_D + F_I. \tag{19.102}$$

For the bank of the merchant, it charges them fee of F_C and has to pay the interchange fee F_I to the bank issuing the payment card. Therefore its profits are given by

$$\Pi_B^C = F_C - F_I, \tag{19.103}$$

assuming again that the bank faces no costs from making payments. Merchants and depositors face benefits from card payments of B_C and B_D, respectively, such as avoiding to handle cash for merchants and depositors, but for depositors also the ability to purchase goods with a credit card without having the funds available at the time of purchase. These benefits are assumed to be different across merchants and depositors, having a distribution of $G(B_C)$ and $H(B_D)$, respectively, but each individual knows its own benefits. Merchants will only accept payment cards if the benefits of doing so are sufficiently large to cover their costs F_C, thus we require $B_C \geq F_C$; similarly depositors will only use payment cards if $B_D \geq F_D$.

Hence the fraction of merchants accepting payment cards are then $1 - G(F_C)$ and the fraction of consumers using payment cards are $1 - H(F_D)$. Assuming that depositors do not select merchants strategically on whether they accept payment cards, the fraction of transactions using card payment would be $T = (1 - G(F_C))(1 - H(F_D))$.

The fees charged to consumers and merchants, F_D and F_C, will also depend on the interchange fee F_I. If we wanted to maximize the number of card transactions, the first order condition is given by

$$\frac{\partial T}{\partial F_I} = -\frac{\partial F_C}{\partial F_I} g(F_C)(1 - H(F_D)) - \frac{\partial F_D}{\partial F_I} h(F_D)(1 - G(F_C)) = 0. \quad (19.104)$$

Looking at equations (19.102) and (19.103), we can use the implicit function theorem such that for a given level of bank profits, we have $\frac{\partial F_D}{\partial F_I} = -1$ and $\frac{\partial F_C}{\partial F_I} = 1$. Hence the fees to depositors and the interchange fees are perfect substitutes while the fees to merchants and the interchange fee are perfect complements. This result implies from the first order condition in equation (19.104) that we require

$$\frac{g(F_C)}{1 - G(F_C)} = \frac{h(F_D)}{1 - H(F_D)} \quad (19.105)$$

and the distribution of fees between merchants and depositors will depend on the respective distribution of benefits in the population. If the benefits to merchants were to increase, thus the distribution shifts upwards, we see that $G(F_C)$ would decrease and hence the left-hand side of this equation reduce, assuming that the density $g(F_C)$ is not affected too much by this shift. This necessitates that the right-hand size also reduces and, absent a change to the distribution, that will require a reduction in the fee charged to depositors as we require that $H(F_D)$ reduces. A similar argument can be made when the benefits to depositors increase. We thus observe that those who obtain higher benefits are paying a higher fee for the use of payment cards. The interchange fee itself cannot be directly determined, but once the fees for depositors and merchants have been set, the interchange fee will be determined such that the desired profit level, as determined by equations (19.102) and (19.103), are achieved.

Let us now propose that the two banks, those issuing payment cards to depositors and of the merchants accepting such cards are forming a card issuing company that maximizes their joint profits, $\Pi = \left(\Pi_B^D + \Pi_B^C\right) T$. These profits are maximized over the interchange fee if

$$\frac{\partial \Pi}{\partial F_I} = \left(\Pi_B^D + \Pi_B^C\right) \frac{\partial T}{\partial F_I} + \left(\frac{\partial F_C}{\partial F_I} + \frac{\partial F_D}{\partial F_I}\right) T = 0. \quad (19.106)$$

If we were to set the optimal interchange fee such that it maximizes the number of transactions, the first term in this condition would be zero. We would thus require that $\frac{\partial F_C}{\partial F_I} + \frac{\partial F_D}{\partial F_I} = 0$. As the level of profits are no longer given, we cannot use our result from above that $\frac{\partial F_D}{\partial F_I} = -1$ and $\frac{\partial F_C}{\partial F_I} = 1$, even though the sum of these two derivatives must still add to zero.

If $\frac{\partial F_C}{\partial F_I} > -\frac{\partial F_D}{\partial F_I}$, the final expression in equation (19.106) would be positive, hence we would need to increase the interchange fee F_I to obtain the maximal joint profits; this condition can be interpreted as the increased interchange fee being passed on to depositors more easily than to merchants. Hence, the interchange fee is higher if banks maximize profits rather than maximize the transaction volume. If fees can be passed on to merchants more easily than to depositors, then the interchange fee will be lower than when maximizing the transaction volume.

The level of the interchange fee will depend on the sensitivity of consumers and merchants to an increase in the transaction fees. Highly competitive consumer markets imply a high sensitivity of the depositor fee F_D and hence higher interchange fees. Similarly, a high degree of competition between merchants would decrease the interchange fee due to them being more sensitive to the fees they are charged. The ability of the card issuing company to pass on any increase in the interchange fee onto the merchant, will protect its profits as we can see from equation (19.103), while increasing profits from depositors as equation (19.102) shows.

We thus see that banks can make additional profits when issuing payment cards by charging interchange fees to administer the payments between the bank of the depositor making the purchase and the bank of the merchant accepting card payments. The observation that fees charged to depositors are uncommon suggests that depositors are very sensitive to such fees and hence $\frac{\partial F_D}{\partial F_I}$ will be close to zero, implying that interchange fees will be higher than would be optimal for maximizing the transaction volume of card payments.

Reading Wright (2004)

Résumé

Credit cards have the benefit of allowing depositors to purchase goods even if their current resources do not allow them to do so. This is more attractive, at least in the short run, as credit cards do not charge interest as long as the balance is repaid at the end of a billing period, typically about four to six weeks. The possible default by depositors on this loan is paid for by merchants paying the bank a fee for being able to accept such card payments, but they in turn benefit from increased sales that finances the payment of these fees. Thus credit cards can be used to overcome short-term shortages of funds by depositors and allows them to maintain their consumption levels, with merchants benefiting from higher sales and banks making profits.

However, banks do not only benefit from the fees charged to merchants for being able to accept credit cards, they also charge the bank receiving payment from a sale for administering this payments, the so-called interchange fees. This provides banks with additional source of revenue from issuing payment cards and depositors being very sensitive to any fees being charged would lead to banks obtaining most of their fee income from merchant fees and interchange fees.

19.3 Payment settlements

A large number of payments are conducted using the transfers of deposits between different accounts, often at different banks. Throughout the day, banks are thus faced with a large number of payments they make to and receive from other banks. Banks can only make such payments if they have available sufficient cash reserves to transfer to another bank, thus the timing of payments made and received can become important to ensure cash reserves are not exhausted at any time. We will therefore look at how the transfer of funds between banks can be organised in chapter 19.3.1 using a specific format, gross settlement, before then comparing it to an alternative net settlement mechanism in chapter 19.3.2, net settlement.

While some of the payments between banks are completed directly, it would be possible for banks to use a more centralised system by routing their payments through a small number of other banks, or even a single bank. Such clearing banks would receive payments from a banks, directed to various other banks, and transfer them on to the final recipients. Chapter 19.3.3 investigates the incentives to become a clearing bank.

It is common that the settlement of payments between banks is conducted multiple times each day, often twice, once in the morning and again in the afternoon. Bank may be able to delay making payments and only complete payments in a later settlement round. Chapter 19.3.4 will investigate the incentives for banks to delay payments and thereby cause a liquidity crunch in which payments are not completed ina timely manner.

A bank not able to make payments due to a liquidity shortage can significantly impact the liquidity of other banks. Their illiquidity can cause illiquidity in other banks as they do not receive payments that are due, which in turn prevent them from making payments. These payments not being made might be the reason for the initial bank not being able to make its payments. The contagion of payment failures together with potential remedies is discussed in chapter 19.3.5 where we discuss the incentives to delay payments in the face of a bank facing a liquidity shortage.

19.3.1 Gross settlement systems

Payments between banks are often settled in a number of rounds during a day and in many cases it is at the discretion of banks to decide how much of the payments they make in each of these rounds. In a gross settlement system, banks have to make payments in each round prior to receiving payments from other banks, therefore they have to hold cash reserves sufficient to make all of their payments without being able to rely on cash reserves they obtain from payments received. Banks, however, can use cash reserves obtained from other banks in previous payment rounds. In contrast to this, in net settlement systems, banks only need to hold cash reserves for the balance of payments they need to make as they can take into account any payments they receive in the same payment round, reducing the required cash reserves to make payments.

Let us assume that we have two banks that have to make payments of M to each other, for example resulting from payments depositors make to accounts held by the other bank. These payments are conducted in two rounds and banks are free to choose the amount of payments they make in each round. Banks can obtain initial cash reserves R_i^0 through a loan from the central bank by posting collateral of the same amount. Banks will then use a fraction λ_i of these cash reserves to make payments in a first round of payment settlements. Thus the first round sees payments of

$$M_i^1 = \lambda_i R_i^0 \tag{19.107}$$

such that the amount of cash reserves held by a bank at the end of round 1 is given by the fraction of cash they have retained and the payments $M_j^1 = \lambda_j R_j^0$ received from the other bank,

$$R_i^1 = (1 - \lambda_i) R_i^0 + \lambda_j R_j^0. \tag{19.108}$$

Any remaining payments are now conducted in the second round. The most payments that can be made is the amount of cash available at this point, R_i^1, and the most that needs to be paid is the remaining balance of $M - \lambda_i R_i^0$. Hence payments in the second time period are given by

$$M_i^2 = \min \left\{ R_i^1; M - \lambda_i R_i^0 \right\}. \tag{19.109}$$

The cash reserves after the second round of payments will consist of the cash reserves initially obtained from the central bank and first round payments, less payments made, $R_i^0 + \lambda_j R_j^0 - M$, if positive; in addition they obtain payments from the other bank in round 2, consisting of the remaining payments $M - \lambda_j R_j^0$ or the cash reserves this bank has available, R_j^1, if smaller. Hence the final cash reserves are given by

$$R_i^2 = \max \left\{ R_i^0 + \lambda_j R_j^0 - M; 0 \right\} + \min \left\{ M - \lambda_j R_j^0; R_j^1 \right\}. \tag{19.110}$$

The total amount of payments conducted will be M or if not enough cash reserves are available, $R_i^0 + \lambda_j R_j^0$, representing all cash reserves the bank could raise from the central bank and other banks in round 1. The final cash reserve R_i^2 will then be used to repay the initial loan R_i^0 and if this loan is repaid in full, the collateral is returned to the bank. For simplicity we assume here that no interest is payable on the loan from the central bank.

We assume that banks charge a fee f for making payments on behalf of depositors. Opportunity costs of collateral provided to the central bank, c, may include costs arising from the inability to invest these funds into more profitable loans. Banks having not sufficient cash reserves may have to cancel payments to the amount of $\max \left\{ M - R_i^0 - \lambda_j R_j^0; 0 \right\}$ at cost $\hat{c}$. We assume that these costs are quadratic in the amount that is being cancelled, reflecting the reputation loss of the bank if it can not make all payments their depositors seek. Hence, bank profits are given by

$$\Pi_B = f \min\left\{M; R_i^0 + \lambda_j R_j^0\right\} - cR_i^0 - \hat{c}\max\left\{M - R_i^0 - \lambda_j R_j^0; 0\right\}^2. \quad (19.111)$$

The first term represents the revenue from making payments on behalf of depositors, the second term the costs of the collateral provision to the central bank and the final term the costs of cancelling payments from depositors.

Banks would only raise as much cash reserves from the central bank as they would need, and due to the costs of collateral provision to obtain such cash reserves, banks would not retain any cash reserves unnecessarily, implying that $\lambda_i = \lambda_j = 1$. We furthermore assume that banks are identical and hence $R_i^0 = R_j^0 = R^0$. Thus the bank profits in equation (19.111) can be rewritten as

$$\Pi_B = f \min\left\{M; 2R^0\right\} - cR^0 - \hat{c}\max\left\{M - 2R^0; 0\right\}^2. \quad (19.112)$$

Let us first assume that $R^0 \geq \frac{1}{2}M$. In this case we easily obtain the bank profits as $\Pi_B = fM - cR^0$ by inserting for R^0. In order to maximize profits, it is obvious that the amount of cash reserves raised from the central bank is the minimal amount, thus $R^0 = \frac{1}{2}M$, giving us bank profits of $\Pi_B^* = \left(f - \frac{1}{2}c\right)M$. The amount of cash held after the second round of payments is given by equation (19.110) and we easily obtain that $R_i^2 = \frac{1}{2}M = R^0$. Hence, the cash reserves held after both rounds of payments have been completed are sufficient to repay the central bank to have their collateral returned. Furthermore, all payments required, M, are made by the bank and no payments are cancelled.

If $R^0 \leq \frac{1}{2}M$, the bank profits in equation (19.112) become $\Pi_B = 2R^0 f - cR^0 - \hat{c}\left(M - 2R^0\right)^2$ and hence the optimal amount of cash reserves raised from the central bank, as given from solving the first-order condition $\frac{\partial \Pi_B}{\partial R^0} = 0$, is obtained as

$$R^0 = \frac{1}{2}M - \frac{c - 2f}{8\hat{c}}, \quad (19.113)$$

which then gives the bank profits of

$$\Pi_B^{**} = \frac{1}{2}(2f - c)M + \frac{(c - 2f)^2}{16\hat{c}}. \quad (19.114)$$

We can easily see that these profits exceed the profits for the case that $R^0 \geq \frac{1}{2}M$, as we have $\Pi_B^{**} > \Pi_B^*$. Thus banks will find it optimal to raise cash reserves of $R^0 = \frac{1}{2}M - \frac{c-2f}{8\hat{c}}$. If the fees charged to depositors for making these payments, f, are sufficiently small such that $f < \frac{1}{2}c$, the amount of cash reserves raised will be smaller than the cash reserves banks require to make all payments as $R^0 < \frac{1}{2}M$ and banks can make payments of at most $2R^0$. As from equation (19.110) we see that the cash reserves after round 2 are given by $R_i^2 = \frac{1}{2}M > R^0$, banks can repay the central bank and have their collateral returned, holding excess cash that they can retain.

However, banks will not be able make all required payments as $M - 2R^0 = \frac{c-2f}{4\hat{c}} > 0$ if $f < \frac{1}{2}c$. Therefore, for banks that do not charge a sufficiently high fee to depositors for making payments, will not raise enough cash reserves from the central bank and prefer being forced to cancel payments, paying compensation $\hat{c}$ to depositors; the additional costs of providing collateral will outweigh the revenue from making these payments and the penalty for not completing all payments. It is thus that the gross settlement system encompasses an inefficiency in that not all payments are completed if the fee charged to depositors for making these payments, f, is too low, relative to the funding costs c of the cash reserves.

In the case that $f > \frac{1}{2}c$, equation (19.113) would imply that banks hold cash reserves in excess of $\frac{1}{2}M$, but as a requirement to this solution was that $R^0 \leq \frac{1}{2}M$, it is easy to show that banks will hold cash reserves at the minimum level of $R^0 = \frac{1}{2}M$ as this maximizes profits; holding excess cash reserves is not beneficial as they attract no interest, but are costly to raise as we outlined above. As in this case $M - 2R^0 = 0$, the bank is able to make all payments as required and due to $R_i^2 = \frac{1}{2}M = R^0$ will also have its collateral returned after repaying the loan from the central bank. In this case of depositors paying substantial fees to make payments, there is no inefficiency in a gross settlement system and all payments will be completed.

A gross settlement system requires banks to raise substantial cash reserves as banks are required to hold cash reserves for the entirety of their payments in each round and cannot account for those payments they receive. As long as the costs of obtaining the requisite cash reserves are sufficiently high compared to the revenue generated from charging depositors to make payments, banks will not raise sufficient cash reserves and will not make all payments that have been requested. Thus gross settlement systems can be inefficient unless the costs of raising the cash reserves are sufficiently low or the fees charged to depositors making payments are sufficiently high.

Reading Buckle & Campbell (2003)

19.3.2 Comparison of settlement systems

Payments can be settled using either gross settlement or net settlement systems. While in gross settlement systems banks have to use cash reserves for all the payments they have to make to other banks, net settlement systems allow banks to offset payments they received from other banks against their own payments. It is therefore that banks will be required to hold less cash reserve in net settlement systems.

Banks have obtained deposits D from depositors who are unsure about the time at which they will withdraw them. A fraction λ of depositors will withdraw deposits after a single time period to finance their consumption, being paid interest r_D^1 and hence withdrawing $\left(1 + r_D^1\right) D$ from their bank. The remaining depositors do not seek to withdraw them until the second time period; of these a fraction $1 - \gamma$ will prepare for this withdrawal by transferring their deposits to another bank. Such depositors are paid interest $\hat{r}_D^2$, consisting of the interest earned with their initial bank and any interest they obtain once the deposit has been transferred; the repayment they receive is then $\left(1 + \hat{r}_D^2\right) D$. Such a transfer might be necessary to pay for any goods or services they seek to purchase, for example if payment is to be made from an account linked to a payment card. The remaining fraction γ of deposits may be transferred to another bank prior to being withdrawn, but do not have to. If deposits remain with the bank during both time period, the bank pays interest r_D^2, such that they obtain a repayment of $\left(1 + r_D^2\right) D$.

Bank i invests their deposits D, less any cash reserves R they hold, into loans $L = D - R$, that provide a return of $\pi_i \left(1 + r_L\right) L$ after two time periods. Banks charge a loan rate of r_L to borrowers and the loan is assumed to be repaid with probability π_i. We consider here a banking system with only two banks.

We can now compare the preferences of banks for the different settlement systems. We will distinguish the case where the risks a bank takes, represented by pi_i, is known to all banks and depositors and a case where this risk is only known to the bank itself.

Known bank risks Let us assume that we know the probability with which bank loans are repaid, π_i, and that these are identical for both banks such that $\pi_1 = \pi_2 = \pi$. We assume that when making their deposits, depositors do not know whether they will be withdrawing their deposits early and whether they will have to transfer them to another bank before withdrawing them in the second time period. They are, however, aware of the probabilities λ and γ for doing so. If depositors have a utility function $u(\cdot)$, then their overall utility is given by

$$\begin{aligned}\Pi_D = {} & \lambda u\left(\left(1 + r_D^1\right) D\right) \\ & + (1 - \lambda)\left(\gamma u\left(\left(1 + r_D^2\right) D\right) + (1 - \gamma) u\left(\left(1 + \hat{r}_D^2\right) D\right)\right).\end{aligned} \tag{19.115}$$

The amount of deposits that are withdrawn after the first time period is given by

$$R = \lambda\left(1 + r_D^1\right) D + (1 - \lambda)(1 - \gamma)\left(1 + \hat{r}_D^2\right) D, \tag{19.116}$$

consisting of those withdrawing early and those withdrawing funds to the other bank. In a gross settlement system banks need to hold cash reserves of R to make these payments to depositors and transfer deposits to the other bank. Hence banks can invest $D - R$ into loans and the repayments to those depositors that remain with their bank, $(1 - \lambda)\gamma\left(1 + r_D^2\right) D$, are financed from the re returns of these loans. Thus we have

$$(1 - \lambda)\gamma\left(1 + r_D^2\right) D = \pi\left(1 + r_L\right)(D - R). \tag{19.117}$$

We now maximize the utility of depositors, equation (19.115), by choosing optimal deposit rates for those withdrawing early, r_D^1 and those transferring their deposits to other banks, $\hat{r}_D^2$, subject to the constraints in equations (19.116) and (19.117). Once these are determined, equation (19.117) allows us to ascertain the deposit rate for those retaining their deposits with their bank, r_D^2. With Lagrange coefficients ξ_1 and ξ_2, we thus obtain

$$\frac{\partial \Pi_D}{\partial\left(1+r_D^1\right)D} = \lambda \frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} \qquad (19.118)$$
$$-\left(\xi_1 - \xi_2 \pi\left(1+r_L\right)\right)\lambda = 0,$$
$$\frac{\partial \Pi_D}{\partial\left(1+\hat{r}_D^2\right)D} = (1-\lambda)(1-\gamma)\frac{u\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}$$
$$-\left(\xi_1 - \xi_2 \pi\left(1+r_L\right)\right)(1-\lambda)(1-\gamma) = 0,$$

which solves for $\frac{\partial u\left(\left(1+r_D^1\right)D\right)D}{\partial\left(1+r_D^1\right)D} = \frac{\partial\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}$, and thus with marginal utilities of those withdrawing early and those transferring to another bank being equal, their values are equal and we have $\left(1+r_D^1\right)D = \left(1+\hat{r}_D^2\right)D$. Inserting this result into equation (19.116) we get the required cash reserves as

$$R = (1-\gamma(1-\lambda))\left(1+r_D^1\right)D. \qquad (19.119)$$

Banks, being other identical, make and obtain the same amount of payments due to the transfer of deposits. It is thus that in a net settlement system, banks can offset the payments they receive from other banks against payments they have to make. In our case this reduces net payments to the amount of deposits that are withdrawn early and the cash reserves required are given by

$$R^* = \lambda\left(1+r_D^1\right)D. \qquad (19.120)$$

The objective function for depositors, equation (19.115) and the constraint from equation (19.117) remain unchanged. Hence we can interpret a gross settlement system as one that is equivalent to a net settlement system with a higher withdrawal rate $\lambda^* = 1-\gamma(1-\lambda) > \lambda$ if we compare equations (19.119) and (19.120).

If we insert equation (19.119) into equation (19.116), and this in turn into the utility of depositors in equation (19.115), we easily get the first order condition for a maximizing this utility as

$$\frac{\partial \Pi_D}{\partial\left(1+r_D^1\right)D} = \lambda \frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} \qquad (19.121)$$
$$-\lambda\pi\left(1+r_L\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D} = 0.$$

Solving equations (19.117) and (19.119) for $\left(1+r_D^2\right)D$ and $\left(1+r_D^1\right)D$, respectively, we can obtain

$$\frac{\partial\left(1+r_D^1\right)D}{\partial\lambda} = -\frac{R}{\lambda^2}+\frac{1}{\lambda}\frac{\partial R}{\partial\lambda}, \tag{19.122}$$
$$\frac{\partial\left(1+r_D^2\right)D}{\partial\lambda} = \frac{\pi\left(1+r_L\right)}{\left(1-\lambda\right)^2}\left(D-R\right)-\frac{\pi\left(1+r_L\right)}{1-\lambda}\frac{\partial R}{\partial\lambda}.$$

Using our first order condition from equation (19.121) we then obtain

$$\begin{aligned}\frac{\partial^2\Pi_D}{\partial\left(1+r_D^1\right)D\partial\lambda} &= \frac{\partial^2 u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D^2}\frac{\partial\left(1+r_D^1\right)}{\partial\lambda} \\ &\quad -\pi\left(1+r_L\right)\frac{\partial^2 u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D^2}\frac{\partial\left(1+r_D^2\right)D}{\partial\lambda} \\ &= 0,\end{aligned} \tag{19.123}$$

which after inserting from equation (19.122) can be solved for

$$\frac{\partial R}{\partial\lambda} = \frac{\frac{\partial^2 u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D^2}\frac{R}{\lambda^2}+\left(\pi\left(1+r_L\right)\right)^2\frac{\partial^2 u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D^2}\frac{D-R}{(1-\lambda)^2}}{\frac{\partial^2 u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D^2}\frac{1}{\lambda^2}+\left(\pi\left(1+r_L\right)\right)^2\frac{\partial^2 u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D^2}\frac{1}{(1-\lambda)^2}} > 0, \tag{19.124}$$

where the positivity of this expression arises from the usual assumption that marginal utility is decreasing, $\frac{\partial^2 u(C_i)}{\partial C_i^2} < 0$. We thus observe that the cash reserves are increasing in the fraction of early deposit withdrawals λ, implying directly that the cash reserves required for the gross settlement system are higher as we had established that the gross settlement system was equivalent to the net settlement system at a higher early withdrawal rate $\lambda^* = 1-\gamma\left(1-\lambda\right) > \lambda$. Another consequence of the net settlement system is that due to the lower cash reserves held, more loans $L = D - R$ can be given, benefitting the economy overall.

Having thus established that the net settlement system is more desirable to banks, as they require less cash reserves, and socially as more loans can be given, we will now change our assumption that the risks banks are taking are commonly known and investigate which impact this change has on the optimal settlement system.

Uncertain bank risks Let us now assume that the risks bank take are not commonly known; the probability with which loans are repaid are either π_H or $\pi_L < \pi_H$, with probabilities p and $1-p$, respectively. We assume that $\pi_L\left(1+r_L\right) < 1 < \pi_H\left(1+r_L\right)$ such that in cases where the bank has taken high risks, π_L, the loans are not profitable and the returns generated by them will not allow to repay deposits in full. The risks that banks are taking is known to the bank itself, but depositors only learn the risks after one time period, prior to deciding whether to transfer deposits, provided they maintain deposits at that bank; the risks of the bank any deposit transfers are made to, remain unknown to that depositor.

We will denote the deposit rate obtained by transferred deposits by $\hat{r}_D^{2,ij}$ if the bank the deposits were originally held by is of type i and the bank the deposit is transferred to is of type j. Similarly, $r_D^{2,ij}$ will denote the deposit rate applied to those deposits retained at their original bank if it is of type i, while the other bank is of type j.

Analysing the net settlement system first, constraints (19.117) and (19.120) from the case of known risks apply if both banks are of the same type, thus both having loan repayment rates of either π_H or π_L. From equation (19.117) we then obtain the repayment to depositors retaining their deposits at their bank as

$$\left(1 + r_D^{2,ii}\right) D = \frac{\pi_i \left(1 + r_L\right)}{1 - \lambda} \left(D - R\right). \tag{19.125}$$

If banks are of different types, then the depositor at the bank taking high risks π_L knows that he will never be fully repaid as the loans are loss-making and hence transfer its deposits to the other bank. For the bank with low risks, π_H, the amount of resources available will be the return on the loans, $\pi_H \left(1 + r_L\right) \left(D - R\right)$, and the money transferred in by depositors of the high-risk bank, which is all depositors not withdrawing early, $\left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,LH}\right) D$. This includes those depositors that have to transfer deposits and those that do so due to learning that the bank has taken high risks. The repayment of deposits after two time periods consists of those that have to transfer deposits to this bank, $\left(1 - \lambda\right) \left(1 - \gamma\right) \left(1 + \hat{r}_D^{2,HL}\right) D$, those not transferring deposits as the other bank is of high risk, $\left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,HL}\right) D$, and those who transferred deposits, $\left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,HL}\right) D$. All depositors are treated equally and receive the same repayment $\left(1 + \hat{r}_D^{2,HL}\right) D$, regardless of the motivation for maintaining deposits at this bank. The repayments of deposits are determined such that the deposits the bank holds are obtaining all the assets of the bank. Thus we have for deposits at the low-risk bank

$$\begin{aligned} &\pi_H \left(1 + r_L\right) \left(D - R\right) + \left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,LH}\right) D \\ &\quad = \left(1 - \lambda\right) \left(1 - \gamma\right) \left(1 + \hat{r}_D^{2,HL}\right) D \\ &\qquad + \left(1 - \lambda\right) \gamma \left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,HL}\right) D + \left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,HL}\right) D \\ &\quad = 2 \left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,HL}\right) D. \end{aligned} \tag{19.126}$$

For the high-risk bank, we get similarly that the resources available are from the loan they have provided, $\pi_L \left(1 + r_L\right) \left(D - R\right)$, and the from those deposits that had to be transferred, $\left(1 - \lambda\right) \left(1 - \gamma\right) \left(1 + \hat{r}_D^{2,HL}\right) D$. These resources are then used to repay depositors that have transferred deposits into this bank as they were required to, $\left(1 - \lambda\right) \left(1 - \gamma\right) \left(1 + \hat{r}_D^{2,LH}\right) D$, and all of their remaining depositors leaving due to them being high risk, $\left(1 - \lambda\right) \left(1 + \hat{r}_D^{2,LH}\right) D$. Thus

$$\begin{aligned}
&\pi_L \left(1+r_L\right)\left(D-R\right)+\left(1-\lambda\right)\left(1-\gamma\right)\left(1+\hat{r}_D^{2,HL}\right)D \\
&\quad = \left(1-\lambda\right)\left(1+\hat{r}_D^{2,LH}\right)D+\left(1-\lambda\right)\left(1-\gamma\right)\left(1+\hat{r}_D^{2,LH}\right)D \\
&\quad = \left(1-\lambda\right)\left(2-\gamma\right)\left(1+\hat{r}_D^{2,LH}\right)D.
\end{aligned} \tag{19.127}$$

Solving equations (19.126) and (19.127) we get the repayments to depositors by the low-risk and high-risk banks, respectively, as

$$\begin{aligned}
\left(1+\hat{r}_D^{2,HL}\right)D &= \frac{\pi_H\left(2-\gamma\right)+\pi_L}{\left(1-\lambda\right)\left(3-\gamma\right)}\left(1+r_L\right)\left(D-R\right), \\
\left(1+\hat{r}_D^{2,LH}\right)D &= \frac{\pi_H\left(1-\gamma\right)+2\pi_L}{\left(1-\lambda\right)\left(3-\gamma\right)}\left(1+r_L\right)\left(D-R\right).
\end{aligned} \tag{19.128}$$

As $\pi_L < \pi_H$, we obtain that the implied deposit rates can be ordered as $\hat{r}_D^{2,LL} < \hat{r}_D^{2,LH} < \hat{r}_D^{2,HL} < \hat{r}_D^{2,HH}$.

If a depositor's bank is low-risk, π_H, then the other bank is low-risk with probability p and the depositor receives $\left(1+\hat{r}_D^{2,HH}\right)D$, else with probability $1-p$ the other bank high-risk, leading to a repayment of $\left(1+\hat{r}_D^{2,HL}\right)D$, if remaining with this bank. If transferring deposits, again $\left(1+\hat{r}_D^{2,HH}\right)D$ is received with probability p, but the bank the deposits are transferred into is of high risk with probability $1-p$, leading to a repayment of $\left(1+\hat{r}_D^{2,LH}\right)D$. Hence, a depositor stays with the bank if

$$\begin{aligned}
&pu\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)+\left(1-p\right)u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \\
&\geq pu\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)+\left(1-p\right)u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right),
\end{aligned} \tag{19.129}$$

from which we obtain that we require $u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \geq u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)$. As we have seen that $\left(1+\hat{r}_D^{2,HL}\right)D > \left(1+\hat{r}_D^{2,LH}\right)D$, depositors who are with a low-risk bank will always prefer to not transfer their deposits.

We assume that

$$\begin{aligned}
&pu\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)+\left(1-p\right)u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \\
&\quad > u\left(\left(1+r_D^1\right)D\right),
\end{aligned} \tag{19.130}$$

such that depositors obtain a higher utility when retaining their deposits with their bank than withdrawing them early; his assumption avoid a bank run occurring.

If a depositor's bank is high-risk, π_L, then in order for depositors to transfer deposits we need

$$\begin{aligned} & pu\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LL}\right)D\right) \\ & > pu\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LL}\right)D\right). \end{aligned} \tag{19.131}$$

This is because with probability p the bank the depositor transfers its deposits to is low-risk, giving $\left(1+\hat{r}_D^{2,HL}\right)D$ and with probability $1-p$ it is also high-risk, giving $\left(1+\hat{r}_D^{2,LL}\right)D$. When staying with the bank, the other bank is low-risk with probability p, giving rise to repayments of $\left(1+\hat{r}_D^{2,LH}\right)D$ and with probability $1-p$ it is also high risk, hence the depositor obtains a repayment of $\left(1+\hat{r}_D^{2,LL}\right)D$. As $\left(1+\hat{r}_D^{2,HL}\right)D > \left(1+\hat{r}_D^{2,LH}\right)D$, we see that the condition in equation (19.131) is always fulfilled and deposits in banks exhibiting high risks will be transferred.

We similarly to equation (19.130) assume that

$$pu\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right) > u(\left(1+r_D^1\right)D), \tag{19.132}$$

such that depositors obtain a higher utility when retaining their deposits with their bank than withdrawing them early; his assumption avoid a bank run occurring. As we easily obtain using $\hat{r}_D^{2,LL} < \hat{r}_D^{2,LH} < \hat{r}_D^{2,HL} < \hat{r}_D^{2,HH}$ that

$$\begin{aligned} & pu\left(\left(1+\hat{r}_D^{2,HLH}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \\ & > pu\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \\ & > p\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LHL}\right)D\right), \end{aligned}$$

we see that if the condition in equation (19.132) is fulfilled, then the condition in equation (19.130) will be also be fulfilled. Hence the constraint in equation (19.132) is more strict and we only its validity to ensure that depositors are not withdrawing early and a bank run is avoided.

For net settlement, the expected utility are thus

$$\begin{aligned} \Pi_D = {} & \lambda u\left(\left(1+r_D^1\right)D\right) \\ & +(1-\lambda)\left(p^2 u\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)\right. \\ & +(1-p)^2\,u\left(\left(1+\hat{r}_D^{2,LL}\right)D\right) \\ & +p\,(1-p)\,u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \\ & \left.+(1-p)\,pu\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)\right), \end{aligned} \tag{19.133}$$

reflecting the utilities if deposits have to be withdrawn early, λ, and if they are retained with banks, $1-\lambda$, the cases that both banks are low-risk, p^2, both banks are high-risk, $(1-p)^2$, and they are of different risk types, $p\,(1-p)$, wither the bank of the depositor being high-risk and the other bank being low-risk, or vice versa.

In the gross settlement systems, payments received do not affect the ability to make payments as those incoming payments cannot be accessed. Hence, if a bank is high-risk, all deposits are withdrawn, leading to utility $u\left(\left(1+r_D^1\right)D\right)$. The inability of the bank to repay deposits in full will have the consequence of a banks facing a bank run. Low-risk banks are equivalent to banks in a net settlement system with a withdrawal rate of $1-\gamma\,(1-\lambda)$, as outlined in the case of known bank risks, such that the expected utility of depositors in a gross settlement system is given by

$$\begin{aligned}\Pi_D^* = &\ (1-p)\,u\left(\left(1+r_D^1\right)D\right) \\ &+p\left((1-\gamma\,(1-\lambda))\,u\left(\left(1+r_D^1\right)D\right)\right. \\ &\left.+(1-\lambda)\,\gamma u\left(\left(1+r_D^2\right)D\right)\right),\end{aligned} \tag{19.134}$$

where $\left(1+r_D^2\right)D$ is given from equation (19.117). If we now define

$$\Delta\Pi_D = \Pi_D^* - \Pi_D \tag{19.135}$$

as the difference in the utility of depositors in the gross and net settlement systems, we see that

$$\begin{aligned}\frac{\partial\Delta\Pi_D}{\partial\pi_L} &= -\frac{\partial\Pi_D}{\partial\pi_L} \\ &= -(1-\lambda)\,(1+r_L)\,(D-R) \\ &\quad\times\left((1-p)^2\,\frac{\partial u\left(\left(1+\hat{r}_D^{2,LHL}\right)D\right)}{\partial\left(1+\hat{r}_D^{2,LHL}\right)D}\frac{1}{1-\lambda}\right. \\ &\quad+p\,(1-p)\,\frac{\partial u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)}{\partial\left(1+\hat{r}_D^{2,HL}\right)D}\frac{1}{(1-\lambda)\,(3-\gamma)} \\ &\quad\left.+p\,(1-p)\,\frac{\partial u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)}{\partial\left(1+\hat{r}_D^{2,LH}\right)D}\frac{2}{(1-\lambda)\,(2-\gamma)}\right) \\ &< 0,\end{aligned} \tag{19.136}$$

after inserting from equations (19.125) and (19.128). Similarly we can obtain that

$$\begin{aligned}\frac{\partial\Delta\Pi_D}{\partial\gamma} &> 0, \\ \frac{\partial\Delta\Pi_D}{\partial p} &< 0.\end{aligned} \tag{19.137}$$

What we see from these relationships is that a gross settlement system becomes more attractive if high-risk banks are more risky, a lower π_L. As the risk of the high-risk bank increases, the resources available to repay deposits being transferred into them reduce and hence it becomes less and less attractive to transfer deposits into this bank and more and more attractive to transfer deposits out of this bank, requiring ever higher cash reserves to be held. In a net settlement system this two developments, less transfers into the bank and increased transfers out of the bank, both increase the cash reserve requirements; in a gross settlement system, however, only the increased transfer out of the bank affects the cash reserves required, given the transfers into the bank cannot be considered and are therefore not affecting the cash reserves necessary. It is thus that the benefits of the net settlement system reduce the higher the risks of the high-risk banks become.

If more deposits can be retained at the original bank, γ, gross settlement systems benefit. This is because in net settlement systems no cash reserves are required if the transfers made into the bank and the transfer made out of the bank are balanced, while in gross settlement systems cash reserves have to be held to allow payments towards other banks. The more deposits can be retained at a bank, the less such transfers occur, benefitting the gross settlement system.

Finally, if low-risk banks become more likely, net settlement systems are benefitting. If $p = 1$ then all banks are low-risk and the risks of the banks are known, giving rise to the case of the known bank risk, which favoured net settlement systems. As high-risk banks become more likely, the benefits of the net settlement system are diminishing. This is because the increased presence of high-risk banks increases the imbalance in transfers between banks, banks identified as low-risk will have lower deposit transfers leaving their banks and more transfers into their bank; high-risk banks will have the opposite imbalance. This will reduce the amount of cash reserves the low-risk bank has to hold as its balance of transfer becomes a net inflow of deposits and deposits losses are reduced. This brings the two settlement systems closer together. For high-risk banks, the cash reserves required are also reducing. While deposits will be leaving these banks, the increased presence of high-risk banks, makes such transfers less profitable for depositors as they may transfer into another high-risk bank. Thus the transfers out of each high-risk bank will be reduced, bringing these two settlement systems closer together. If the fraction of high-risk banks becomes sufficiently high, a low p, gross settlement systems can be preferred. This is because high-risk banks having to hold higher cash reserves will prevent them from providing loss-making loans, actually improving their ability to repay deposits retained with them. This will benefit depositors and will outweigh the increased cash reserves held by low-risk banks, which are small due to the low transfer of deposits out of these banks, despite lower repayments due to less profitable loans being provided.

We thus observe that if there are sufficient high-risk banks in the market, a gross settlement system is more beneficial as it reduces the amount of loss-making lending by these banks while the reduced profitable lending of low-risk banks are less affected.

Summary If risks of banks are known, net settlement systems are preferred as the lower cash reserves that banks are required to hold allows for more loans to be provided and these enable banks to pay higher deposit rates. If, however, some banks are high-risk in the sense that on average they do not have profitable lending, it might be beneficial to restrict the lending of these banks by requiring them to hold larger amounts of cash reserves. With gross settlement systems requiring higher cash reserves, applying such a payment system would be beneficial. Of course, if banks were known at the time of making deposits they are high-risk and unlikely to be able to provide profitable lending, they would not receive any deposits, thus it must be that the quality of the bank's lending is not known in advance, but information is only obtained once deposits are already made.

In banking systems with little differences in the risks banks take, net settlement systems are preferred, while for banking systems where banks take different levels of risks, depositors are better off with gross settlement systems. This will be in particular the case where risk differences between bank take can be substantial and depositors have much discretion on transferring their deposits across banks.

Reading Freixas & Parigi (1998)

19.3.3 The emergence of clearing banks

Payments between banks can be made through a centralised settlement system in which all banks simultaneously submit the payments they have to make and these payments are then completed in a single large transaction. It is, however, also possible that banks agree to make payments bilaterally between them, but this requires that banks hold accounts with each other such that they can credit payments received and debit payments made. Such a decentralised settlement has not only the disadvantage that it requires a large number of individual transactions, but that it also requires a bank to hold an account with each of the other bank. To alleviate the problem of each bank having to hold and account with each of the other banks, clearing banks can be engaged. It is now that each bank has an account with a single bank only, the clearing bank. A payment is now made by a bank to the clearing bank, who then completes the payment by making itself a payment to the recipient bank.

Let us assume that banks have to settle a payment of M between them in two settlement rounds. Banks can make the entire payment early and are given a discount λ such that they have to pay only λM in early settlement, or they settle late at the full payment of M. However, some banks may face a liquidity shortage in late settlement and will not be able to make any such payments; a fraction $1 - p$ of banks falls into this category, implying that a fraction p of banks would make the full payment. While the bank does know its own type, the bank receiving the payment will not be aware whether the bank making the payment will face a liquidity shortage,

Banks can settle their payments directly between them at some cost c, which will not only take into account any administrative costs but also those costs arising from the requirement of cash reserves to make any such payments. The bank making the payments will commit to making a payment of $M_1 = \lambda M$ if paying early and a payment of M is paying late. With the possibility of the bank facing a liquidity shortage and not being able to make a late payment, the expected late payments are given by $M_2 = pM$. If we now assume that $\lambda > p$, then it is obvious that the receiving bank will prefer banks to make early payments.

If we assume that banks are as likely to make payments as they are to receive payments, they will receive their payment of λM with probability $\frac{1}{2}$, while the costs of holding cash reserves will be incurred regardless of the direction of payment flows. The profits of the recipient bank will thus be

$$\Pi_B^* = \frac{1}{2}\lambda M - cM, \tag{19.138}$$

taking into account the costs c of direct settlement. For banks to engage in direct settlement, it must be profitable to do so and we require $\Pi_B \geq 0$, from which we easily obtain

$$c \leq c^* = \frac{1}{2}\lambda. \tag{19.139}$$

Hence the costs of direct settlement must not be too high for banks to engage in any payment activity through direct settlement.

Rather than settling payments directly between them, banks may engage clearing banks to conduct these payments on their behalf. The bank will make payments of M_t, and the clearing bank will then make a payment of $\hat{M}_t$ to the recipient bank. We assume that the clearing bank will not face a liquidity shortage and hence makes their payments with certainty. Clearing banks will not make more payments than they have received by the original bank, thus $\hat{M}_t \leq M_t$ and as before $M_1 = \lambda M$ as the discount for early payment also applies when using clearing banks; the late payment M_2 will be either the full amount M if the bank does not have a liquidity shortage or zero otherwise. Assuming that the clearing bank knows whether the bank will face a liquidity shortage, the payments are maximized with early payment if the bank will face a liquidity shortage and late payment if it will not face a liquidity shortage. However, assume in addition that clearing banks will not engage with banks facing liquidity shortages and not being able to meet their obligations, hence payments of M are received, but only if the bank will not face a liquidity shortage. If the bank originating the payment will face a liquidity shortage, the clearing bank will not conduct the payment; a consequence of this selection of banks is that clearing bank will always make late payments. With clearing banks charging a fee $\hat{c}$ for their services, the recipient bank's profits are given by

$$\Pi_B^{**} = \frac{1}{2}pM - \hat{c}M, \tag{19.140}$$

where we take into account that the paying bank must not face a liquidity shock, which is the case with probability p. We also assume that the clearing bank charges a fee for their services regardless whether they are used.

Banks will prefer the use of clearing banks over direct settlement if they receive larger payments, net of any costs, from doing so, $\Pi_B^{**} \geq \Pi_B^*$. Inserting from equations

(19.138) and (19.140) we easily obtain that

$$c \geq c^{**} = \hat{c} - \frac{1}{2} p \left(1 - \lambda\right). \tag{19.141}$$

If the costs of direct settlement are not much higher than the fee charged by clearing banks, their use is preferred. Banks will take into account the ability of clearing banks to screen other banks and ensure that full payment can be obtained, which allows clearing banks to charge a higher fee than the costs of direct settlement.

Of course, using a clearing bank itself must be profitable, $\Pi_B^{**} \geq 0$, thus requiring

$$\hat{c} \leq \hat{c}^{*} = \frac{1}{2} p. \tag{19.142}$$

The costs of using clearing banks must not exceed the benefits from screening out banks with liquidity problems. We see that the use of clearing banks allows payment settlements to be conducted even when direct settlement is very costly, improving the efficiency of the payment system.

Banks may now be using both forms of settlement, if the originating bank is not facing a liquidity shortage, it will use the clearing bank and obtain payment M, while if the originating bank will face a liquidity shortage, the direct settlement is used and the bank obtains λM. Of course, the bank has to bear the costs of both settlement systems. Hence its profits are given as

$$\Pi_B^{***} = \frac{1}{2} pM - \hat{c}M + \frac{1}{2} \left(1 - p\right) \lambda M - cM. \tag{19.143}$$

Using both mechanisms is preferred to using the direct settlement only if $\Pi_B^{***} \geq \Pi_B^{*}$, which after inserting from equations (19.138) and (19.143) becomes

$$\hat{c} \leq \hat{c}^{**} = \frac{1}{2} p \left(1 - \lambda\right). \tag{19.144}$$

If this condition is fulfilled, we see from equation (19.141) that $c \leq 0$ and hence banks would never prefer both settlement mechanisms over the direct mechanism, making this constraint irrelevant.

Similarly, using both mechanisms is preferred to using clearing banks only if $\Pi_B^{***} \geq \Pi_B^{**}$, which after inserting from equations (19.140) and (19.143) becomes

$$c \leq c^{***} = \frac{1}{2} \lambda \left(1 - p\right). \tag{19.145}$$

Thus, if $c \leq c^{***}$, $c \leq c^{**}$ and $\hat{c} \leq \hat{c}^{*}$, banks prefer to use both mechanisms over the clearing banks alone.

Of course, using both settlement forms has to be profitable, thus we require that $\Pi_B^{***} \geq 0$, from which we obtain with equation (19.143) that

$$c \leq c^{****} = -\hat{c} + \frac{1}{2} \left(p + \lambda \left(1 - p\right)\right). \tag{19.146}$$

We can show that this constraint on the profitability of banks using both settlement mechanisms is not imposing additional restrictions, in the relevant areas, using both settlement mechanisms will always be profitable for banks.

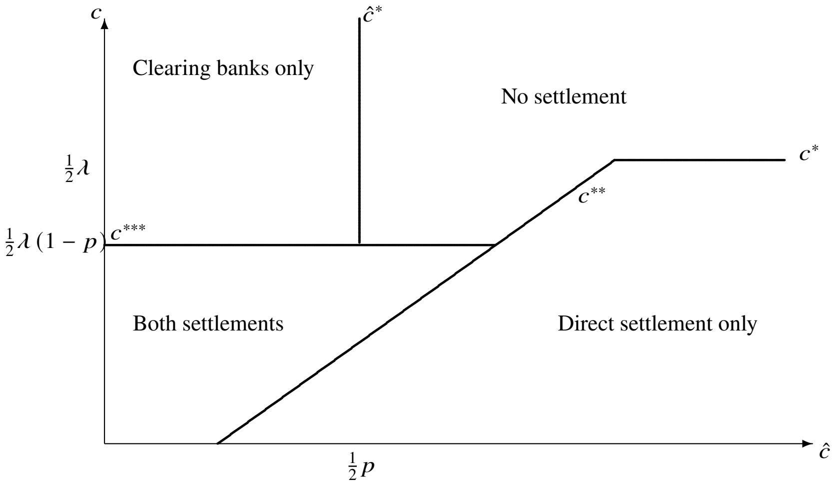


Fig. 19.4: Settlement with clearing banks

Figure 19.4 shows the preferred way of settling payments between banks. We easily note that in cases where clearing banks charge substantially higher fees than the costs of direct settlement, their use becomes unfeasible and similarly, if the costs of direct settlement is higher, only clearing banks are used for the settlement of payments. If both costs are high, banks prefer not to engage in payment settlement at all. In addition, if the costs of direct settlement are sufficiently low, banks will use both forms of settlement, despite having to bear the costs of both settlement forms. They will prefer to use clearing banks as this ensure they obtain the full payment, but will use the direct settlement if the clearing bank cannot be used due to the originating bank facing a liquidity shortage for late settlement.

If we assume that clearing banks face the same costs of settlement than other banks, c, from holding cash reserves and other administrative costs, they are profitable as long as the fee they charge for their services exceeds these costs, thus $\hat{c} \geq c$. If clearing banks are not offering their services due to this activity not being profitable, direct settlement would chosen as long as $c < c^{**}$. However, clearing banks might offer their services even at a loss for strategic reason, for example their importance for payment settlement might be recognised by regulators and they might be able to access liquidity being offered preferential conditions. Being recognised as a clearing bank may also be attractive to depositors seeking fast payments as depositing with them directly would reduce the steps required to successfully complete a payment, this might attract more deposits and benefit banks indirectly.

We have thus seen that the ability of clearing banks to identify banks that face liquidity shortages in the future and would thus not able to make payments makes their presence beneficial to banks receiving payments. Their expertise reduces the risks to banks in the payment system and this will be of particular importance when considering international payments between banks, where knowledge of the risks banks in different jurisdictions face, will often be very limited. With their knowledge of these banks due to more frequent contacts and experience from past transactions, clearing banks will be able to facilitate payments between banks where direct settlement might be more costly or even not feasible.

Reading Chapman, Chiu, & Molico (2013)

19.3.4 Liquidity shortages in settlement systems

Depositors often seek to transfer deposits at various times during the day. While settlements are not necessarily conducted real-time, that is each transfer of deposits is settled immediately, there may be multiple settlement periods each day, most commonly one settlement in the morning and a second settlement in the afternoon. Typically, transfers submitted to the bank prior to a cut-off time in the morning are expected to be completed in the first settlement period (early payment) and transfers submitted after the cut-off time will be completed in the second settlement period (late payment). It may be possible for banks to delay the completion of a transfer requested prior to the cut-off time for the first settlement, often at a cost to the bank as depositors might be concerned about any such delays. Such delays in making payments can cause liquidity shortages with banks not receiving payments in a timely manner and may also affect the ability of depositors to make their own payments if they receive funds late.

Let us assume that depositors can request transfers to be completed in the first or the second settlement period. Depositors may also not submit any transfer requests or they may submit a transfer request for both settlement periods. If we denote a request by a depositor to make a transfer M in settlement period t by p_t, we have in general that $p_1 + p_2 \neq 1$ and we assume that transfer requests are independent across the two banks we consider.

Banks receiving a transfer request for the first settlement period can delay completing the transfer to the second settlement period at some cost C to account for the loss in reputation the bank might face from not being able to complete transfers for their depositors in a timely manner, or any compensation they might have to pay any affected depositor.

We consider the decision by banks to complete transfer requests received prior to the first settlement period early or delay these to the second settlement period. If banks do not hold sufficient cash reserves, they are able to obtain a loan from the central bank that allows them to make the requisite payments, in a gross settlement system these cash reserves have to be available prior to the commencement of the settlement period, while in a net settlement system additional cash reserves are only required if not sufficient payments are received during the settlement period itself.

Gross settlement system Initially we consider the case of a gross settlement system where banks can obtain liquidity from the central bank to make payments if their cash reserves are not sufficient prior to the settlement period. Obtaining such a loan imposes costs of $\hat{C}$ onto the bank and will include the interest charged by the bank as well as other costs, such a provision of collateral and the opportunity costs of not being able to use this collateral for more profitable purposes. If the bank colds cash reserves, these costs are the opportunity costs of not being able to use these otherwise to generate profits to the bank.

The bank does have no cash reserves to make payments, thus it will have to obtain a loan from the central bank at cost $\hat{C}$ in order to make payment in the early payment round. If the other bank receives a payment request, and also pays early, the bank can repay the loan from these proceeds. If the other bank does not have a request for an early payment, which happens with probability $1 - p_1$ then it needs to extend the central bank loan to the second settlement period, again incurring costs of $\hat{C}$. The bank receives a payment request for the second settlement period with probability p_2 and, having no cash reserves, will have to obtain a loan from the central bank again at cost $\hat{C}$. Thus the costs to the bank of making the requested payment early will be

$$\Pi_B^{EE} = \hat{C} + (1 - p_1)\,\hat{C} + p_2\hat{C}. \tag{19.147}$$

If, on the other hand, the bank delays the payment, it does not require a loan from the central bank for the early payment round, but faces delay costs C. If the other bank receives an early transfer requests and makes this payment early, it will obtain sufficient cash reserves to make the late payment, otherwise it will have to take a central bank loan at costs $\hat{C}$. The bank will again face the possibility of a payment request for the second settlement period with probability p_2 and would need a loan from the central bank to make this payment. Its costs of making a late payment while the other bank makes an early payment is thus

$$\Pi_B^{LE} = C + (1 - p_1)\,\hat{C} + p_2\hat{C}. \tag{19.148}$$

A bank may make an early payment, while the other bank delays its payment. In this case the bank would need to rely on a central bank loan for the early payment and as the payment of the other bank is not received until the second settlement period, it would need to extend this loan, giving it total costs of $2\hat{C}$. In addition, the bank will again face the possibility of a payment request for the second settlement period with probability p_2 and would need a loan from the central bank to make this payment. Making payment early while the other bank delays payment gives the bank costs of

$$\Pi_B^{EL} = 2\hat{C} + p_2\hat{C}. \tag{19.149}$$

Finally, if both banks delay their payments, the bank will face the delay costs and as it has not received any payment from the other bank in the first settlement period, it will have to take a central bank loan to make the late payment and the possibility of a payment request for the second settlement period arrives with probability p_2, requiring an additional loan from the central bank to make this payment. If both banks delay their payments, the costs they face are given by

$$\Pi_B^{LL} = C + \hat{C} + p_2 \hat{C}. \tag{19.150}$$

		Bank 1	
		early	late
Bank 2	early	$(1 - p_1)\,\hat{C},\ (1 - p_1)\,\hat{C}$	$C - p_1\hat{C},\ \hat{C}$
	late	$\hat{C},\ C - p_1\hat{C}$	C, C

(a) Gross settlement systems

		Bank 1	
		early	late
Bank 2	early	$(1 - p_1)\,\hat{C},\ (1 - p_1)\,\hat{C}$	$C, \hat{C}$
	late	$\hat{C}, C$	C, C

(b) Net settlement systems

Fig. 19.5: Strategic payment delays

The two banks enter a strategic game on whether to make an early or late payment for the transfer that has been requested for the early settlement period. This game is shown in figure 19.5a, where we have eliminated the common factor $\hat{C} + p_2\hat{C}$ from the costs for clarity. We instantly see that for $C < \hat{C}$, the equilibrium is for both banks to delay payments to the second settlement period; if the costs of providing collateral is higher than the costs of delaying payments, these get delayed. If $C > \hat{C}$ and the costs of delaying payments are higher than the costs of obtaining a loan from the central bank, the equilibrium is to process payments instantly and make early payment.

Net settlement system In a net settlement system, the bank can use payments received from other banks to offset any shortages of cash reserves to make payment. It is therefore that banks are only required a take a loan if there is no payment from the other bank being made in the same settlement period.

If both banks make early payments, the bank only has to take a loan from the central bank for the early payment if the other bank has not received a request for a transfer itself. In this case, the loan does not need to be extended to the second settlement period if the bank does not obtain a transfer request for the second settlement period, but the other bank does; the payment from the other bank allows the bank to repay the loan. This scenario happens with probability $p_2(1-p_2)$ such that the loan is extended with probability $1-p_2(1-p_2)$. An additional loan from the central bank in the second settlement period is required if the bank obtains a request to transfer deposits, but the other bank does not obtain such a request.

If the bank decides to pay early and the other bank decides to delay its payment, the bank will have to take a loan for the first settlement period as it has no cash reserves, but the considerations for the second settlement period as in the previous case. If the bank decides to delay its payment, it faces delay costs, regardless of what the other bank does, with the loan requirements for the second settlement period as before.

The costs of banks from making payments for the different possibilities of banks making early and late payments are therefore given by

$$\begin{aligned} \Pi_B^{EE} &= (1-p_1)\hat{C} + (1-p_1)(1-p_2(1-p_2))\hat{C} \\ &\quad + p_2(1-p_2)\hat{C}, \\ \Pi_B^{EL} &= \hat{C} + (1-p_1)(1-p_2(1-p_2))\hat{C} + p_2(1-p_2)\hat{C}, \\ \Pi_B^{LE} &= C + (1-p_1)(1-p_2(1-p_2))\hat{C} + p_2(1-p_2)\hat{C}, \\ \Pi_B^{LL} &= C + (1-p_1)(1-p_2(1-p_2))\hat{C} + p_2(1-p_2)\hat{C}. \end{aligned} \tag{19.151}$$

The resulting strategic game is shown in figure 19.5b, where the two common final terms have been eliminated for clarity. We can easily see that if $C < (1-p_1)\hat{C}$ the only equilibrium is to delay payments as the costs of doing so are lower than obtaining a loan from the central bank if the other bank does not have a transfer request for the early settlement period. Because such a loan only need to taken if no payment from the other bank is received, this constraint is more binding than in the gross settlement system where the condition was $C < \hat{C}$.

If $C > \hat{C}$ then banks will not delay payments as the cost of doing so are too high, similar to the provision of collateral. In the intermediate case that $(1-p_1)\hat{C} \leq C \leq \hat{C}$, both banks delaying payments or both banks making early payments are equilibria.

Summary We see that late payments in net settlement systems are less common than in gross settlement systems. Delaying payments in gross settlement systems has the advantage that a loan to obtain cash reserves is not required, while with early payments such a loan would always be required; in contrast to that, in net settlement systems a loan is only required if the other bank makes no payment. This reduces the costs of making payments early, and hence payment delays are less commonly observed in net settlement systems. We therefore will see liquidity crunches less frequently in net settlement systems compared to gross settlement systems.

Reading Bech & Garratt (2003)

19.3.5 The spread of liquidity shortages

It seems obvious that a bank not receiving payments from another bank in time, might fail itself due to a liquidity shortage and not be able to make its own payments. A bank which is due to make payments will of course consider whether itself will obtain sufficient payments to avoid a liquidity shortage. If a liquidity shortage may arise, the bank has to consider whether the payments due can be made. Thus the failure of one bank to pay may well affect other banks, even if they are not directly affected by the initial failure. The failure of one bank to make a payment can spread through the payment systen and stop other payments being made.

Let us assume that we have three banks having to make payments of M to each other, holding cash reserves of M. It is, however that no payments are being made between banks 2 and 3, such that bank 2 exchanges payments only with bank 1 and so does bank 3. Consequently, bank 1 exchanges payments with both banks, banks 2 and 3, having to make payments of $2M$, while banks 2 and 3 have to make payments of M. Any payments between banks are concluded in two rounds, an early round and a late round, where payments submitted for the early round can be delayed at same cost C to banks. These costs arise from a possible reputation loss due delaying payments or any compensation being paid to depositors whose payments are delayed. A bank that cannot settle their payments in the late round faces costs $\hat{C} > M > C$ for not making these payments at all. In each round the aforementioned payments have to be made, with the option to delay payments from the early round such that the payments in the late round will then be doubled.

We now assume that bank 3 faces a liquidity shortage during the early payment round, such that it cannot make any payments and has to delay its payments. This liquidity shortage persists into the late round with probability p and the bank would not be able to make any of its payments; if the liquidity shortage does not persist, bank 3 will be able to make payments in the late round, but will not catch up on payments it was not able to make in the early round. Banks 1 and 2 have sufficient cash reserves to make one additional payment, compared to the number of payments they receive, but if all payments due are received, the cash reserves after these payments are identical to the ones they start with. It is thus that bank 3 not making payments in the early round to bank 1 would not cause the failure of this bank.

Assume now that both banks make payments in the early round. Both, banks 1 and 2 can make their payments, but bank 1 will have no cash reserves left at this stage. Therefore, in the late payment round, bank 1 will not have sufficient cash reserves to make payments to both banks 2 and 3, unless the liquidity shortage of bank 3 does not persist. Bank 1 will fail if the liquidity shortage of bank 3 persists, facing costs $\hat{C}$ and in this case the payment received from bank 2 can be used to make payments to either bank 2 or bank 3, which we assume will be split equally. Thus bank 2 will make a loss of $\frac{1}{2}M$ if the liquidity shortage of bank 3 persists. If the liquidity shortage of bank 3 does not persist, bank 2 can be paid in full, making no losses, and bank 1 will lose the payment of bank 3 from the early round, M. Thus the expected losses of bank 1 are $p\hat{C} + (1 - p)\, M$ and that of bank 2 will be $\frac{1}{2} pM$.

In the case that bank 2 delays payments and bank 1 were to make early payments, bank 1 would not have enough cash reserves to make both payments to banks 2 and 3 as it does not receive any payments itself, it would fail at costs $\hat{C}$. Bank 2 would not lose any cash reserves and thus make no losses.

If bank 1 delays payment, but bank 2 were to make early payments, then bank 2 would be deprived of any cash reserves while bank 1 will accumulate cash reserves of $2M$. In the late round bank 1 would then have to pay $2M$ to each, bank 2 and bank 3, while receiving M from bank 2. This is only possible if the liquidity shortage of bank 3 does not persist. If the liquidity shortage persists, bank 1 will fail at costs $\hat{C}$ and if it does not persist it faces the delay costs of C. Thus the costs to bank 1 are $p\hat{C} + (1 - p)\,(M + C)$, where we acknowledge that the bank will also be missing the early payment of bank 3. Bank 2 will receive their full payments if the liquidity shortage of bank 3 does not persist, but will not receive both payments in full; bank 1 will have cash reserves of $3M$ after obtaining the payments from bank 2 and will have to make payments of $4M$, thus bank 2 will receive $\frac{3}{2}M$, making a loss if $\frac{1}{2}M$. Hence the losses of bank 2 are $\frac{1}{2} pM$.

Finally, if both banks delay their payments, the initial cash positions remain unchanged until the late round of payments, where each bank is supposed to make payments of $2M$. If the liquidity shortage of bank 3 does not persist, the bank 1 can make all payments and only faces the shortage of payments from bank 3 and the delay costs. Should the liquidity shortage of bank 3 persist, bank 1 could not make all payments and fails at cost $\hat{C}$, while the payments to bank 2 are reduced to $\frac{3}{2}M$, causing a loss of $\frac{1}{2}M$. Due to the failure of bank 1, no delay costs are incurred by bank 2. Hence the losses to bank 1 are $p\hat{C} + (1 - p)\,(M + C)$ and bank 2 faces losses of $\frac{1}{2} pM$.

		Bank 1	
		early	late
Bank 2	early	$p\hat{C} + (1 - p)\, M, \frac{1}{2} pM$	$p\hat{C} + (1 - p)\,(M + C), \frac{1}{2} pM$
	late	$\hat{C}, 0$	$p\hat{C} + (1 - p)\,(M + C), \frac{1}{2} pM$

Fig. 19.6: Strategic interactions over payment delays

Figure 19.6 shows the resulting strategic interactions between banks 1 and 2 on whether to delay payments or make payments in the early round, nothing that the payoffs from the different strategies represent losses rather than profits. We see that if $\hat{C} < M + C$, the equilibrium is for both banks to delay payments and otherwise only for bank 2 to delay payments, while bank 1 pays in the early round. Thus if the costs of not making payments are sufficiently high, bank 1 will delay payments to avoid the certain failure if bank 2 decides to delay their payments, while bank 2 is indifferent between making payments early or late in this case. Even if the costs of failing to make payments are low and bank 1 would make early payments, bank 2 would delay their payments as to preserve its cash position and not incur losses if the liquidity shortage of bank 3 persists.

We thus observe that payments are delayed if one bank faces a liquidity shortage and cannot make payments, even though the missed payments by that bank can be covered with existing cash reserves held by the bank supposed to receive the payment. The payment system will observe a liquidity crunch as the result of other banks protecting their own liquidity position by making payments late. We will thus observe a breakdown or reduction of payments in the early round.

The total costs to both banks combined are minimal if they both make early payments, provided that $\hat{C} > \frac{1}{2}\frac{2-p}{1-p}M$. Thus the equilibrium of at least bank 2 making late payments is inefficient in that it imposes higher total costs. If $\hat{C} < \frac{1}{2}\frac{2-p}{1-p}M$, then bank 1 making early payments and bank 2 delaying payments would have the lowest total costs. This is an equilibrium only if $\hat{C} < M + C$ and hence this equilibrium is also minimizing the total costs if both conditions are fulfilled, requiring that $M + C \geq \frac{1}{2}\frac{2-p}{1-p}M$ as the constraint on the equilibrium needs to be stricter. Only if the delay costs are sufficiently low, $C \leq \frac{p}{1-p}M$, and the costs of not making payments are also sufficiently low, $\hat{C} < C + M \leq \frac{M}{1-p}$, would an efficient outcome be obtained. If the costs of defaulting on payments is sufficiently high, no early payments are made.

We thus see that if a bank faces a liquidity shortage that does not allow it to make payments to other banks, this will affect the behaviour of other banks who seek to preserve their liquidity, even if the missed payments can be covered by existing cash reserves and they are not directly affected by this liquidity shortage and the lack of payments received. Banks will start to delay payments so as not to expose themselves to a liquidity shortage themselves if other banks delay payments; this will result in liquidity hoarding by banks and payments being delayed. We thus observe spillovers of liquidity shortages from a single bank to affecting the ability of the payments system to effectively operate.

Reading Foote (2014)

Résumé

Payments between banks are essential to allow depositors to transfer funds between accounts at different banks and are therefore an essential part of the financial system. However, operating a payment system is costly in that banks need to hold cash reserves from which these payments are met and they seek to minimize these costs. The uncertainty about payments that are made by a bank on behalf of their depositors and received from other banks for the benefit of their depositors are uncertain, easily leading to not sufficient cash reserves being held.

Concerns about their own cash reserves and their ability to meet their obligations from making payments can lead to a situation where these payments are delayed in the anticipation of payments from other banks being received first before the bank makes any payments itself. With other banks anticipating such a move, they will also delay making payments and thus all payments are delayed, reducing the efficiency of the payment system. Any real shortages of cash reserves by banks can also spread in that banks become overly cautious in their use of cash reserves, which might even lead to failures of banks as they do not obtain any payments while making payments on behalf of their depositors. Thus a shortage of cash reserves can spread in the payment system.

Mechanisms have been developed to reduce the costs to banks, and the risks of such liquidity shortages emerging endogenously due to banks withholding payments. Most notably, a net settlement system where banks can use payments obtained from other banks to make their own payments can reduce this risk notably compared to a gross settlement system where payments from other banks can only be accessed after all payments have been made. Also if there is asymmetric information between banks on their respective liquidity positions, some banks which have superior information may act as clearing banks and overcome the possibility of payments not being conducted.

Conclusions

Payment services on deposit accounts from an important part of the benefits these accounts provide, in addition to being an investment for any excess funds. Although the importance of access to cash has reduced over time with the spread of payment cards, it nevertheless remains a concern for many depositors, individuals as well as small businesses. Banks cooperating in providing access to each other's facilities, such as cash machines will on the one hand forego some of the benefits they can provide exclusively to their own depositors, such as an extensive network of cash machines or a branch network to deposit cash into their account, but will on the other hand increase the benefits of their depositors through this reciprocal access to each other's services. This will increase their competition due to a lack of differentiation in the offerings between banks, but will also attract more depositors to their bank. In the same way does remote access to their accounts benefit depositors, while also increasing competition if banks offer ever more homogenous account features by reacting to any offering of their competitor, particularly as such remote access can be offered at relatively low costs by banks.

In many countries it is unusual for account services, such a access to cash or online banking, to be explicitly paid for through a fee. In other countries such fees, however, are more common and can be substantial; in some instances it might be that some banks offer accounts without a fee, while other accounts attract a fee. While the latter type of accounts often come with additional services, such as insurance packages or better access to other account services, another common feature is that they offer higher deposit rates or access to loan at preferential rates. Account fees and interest rates can be seen as close substitutes for depositors, who will only be concerned in the net benefits from their account. We can therefore find different types of accounts with different charging structures to co-exist and allow depositors to choose the combination that most suits their needs.

The use of many account services has considerably changed over time. One such important change has been the dominance of payment cards when paying retailers, compared to the pre-dominant use of cash previously. This development has been accelerated by the spread of online businesses who will rely on the use of payment cards. While much of the growth in the use of payment cards can be attributed to debit cards, credit cards have also become much more widespread. While these cards charge retailers a fee for each transaction, and are therefore often more expensive than the handling of cash, their use nevertheless benefits retailers as it allows depositors to make purchases even if they currently do not have the funds available. Such short-term loans can be used to smoothen consumption and can lead to an overall increase in purchases.

The increased use of non-cash payments has lead to an increase in payments being made between banks. Banks make such payments on behalf of their depositors and they might include the transfer of funds arising from transactions using payment cards, but also direct transfers between accounts; previously many of these transactions would have been settled with cash and thus seeing no role for banks beyond ensuring depositors can access cash. Payments between banks, however, can increase risks in the banking system in that these payments require banks to hold cash reserves from which these payments are taken. If banks are concerned about their cash reserves being depleted because other banks might face a liquidity shortage and curtail the payments they make, they will also reduce the amount of payments they make in order to preserve their cash reserves. This can lead to a breakdown of the payment system and any actual shortages in cash reserves can lead to the failure of banks if they are not able to delay payments any further.

Review

Deposits are the main source of funding for banks and the basis on which they are able to provide loans. Without deposits, banks would not be able to grant the volume of loans we see and their profits would also be significantly smaller. However, by providing loans through the use of deposits, banks are not only exposing themselves to high risks from this leverage, but their depositors, too. While deposit rates should reflect the risks banks expose their depositors to, the nature of deposits exposes banks to the additional risk of bank runs. With the ability to withdraw deposits instantly, which is not matched by the bank's ability to liquidate assets, any losses banks might incur, could lead to a bank run. Not o9nly does a bank run cause the bank itself to suffer significant losses from its attempt to generate liquidity, but depositors themselves would also make losses. It is not even necessary that a bank incurs losses, or that rumours to that effect are circulating, it is already sufficient that depositors think that other depositors might withdraw. With the first depositors to withdraw being repaid their deposits and those withdrawing later not obtaining any repayments as the bank has incurred losses exceeding its equity, there is a strong incentive to withdraw themselves. Thus only a change in expectations about the behaviour of other depositors, even if such a change would be unjustified by the observer's own information, can induce a bank run. This makes banks inherently vulnerable not only to actual risks, but also to the expectation formation of depositors.

Banks have developed mechanisms to withstand deposit withdrawals to some extent. With some banks having excess liquidity, they might be willing to lend their excess funds to a bank facing larger than expected deposit withdrawal. While such interbank loans would not be able to avoid a bank run, it might be sufficient to prevent expectations of depositors to change as the bank can obtain more liquidity and hence an instant withdrawal is not necessary. Depositors might be able to afford a wait-and-see approach to withdrawing and as a result no bank run materialises. A more effective way of preventing a bank run is the establishment of deposit insurance. With their deposits fully insured, depositors would not be concerned about the risks the bank takes or a change in the expectations of other depositors, they would not make a loss if retaining their deposits and hence would not withdraw and a bank run cannot emerge. However, such deposit insurance implies that banks can obtain deposits without a risk premium and can use this low-cost funding to finance more risky loans. Thus a moral hazard problem emerges that, unless banks pay an appropriate price for deposit insurance, banks will take on higher risks. The extent to which deposit insurance is provided needs to balance the protection of depositors and the incentives to take risks by banks.

In addition to deposits as investments of funds, they also provide a wide range of additional services, most notably the ability to access cash and make payments by payment cards and transfer between accounts. These services are of central importance in an economy as not to rely solely on cash payments. However, payments between accounts, in as far as these accounts are held at different banks, are also accompanied by payments between banks to enable to make and receive such payments by their depositors. Transferring funds between banks, however, requires cash reserves and limits the amount a bank a can invest into loans. With banks seeking to minimize their costs, they will hold the minimum amount of cash reserves possible to conduct these payments. This can lead to a situation where banks seek to preserve their cash position by delaying payments. This can lead a breakdown of the payment system as banks not expecting payments to be delayed may face a liquidity shortage and are not able to make payments themselves. Delaying payments can become self-fulfilling, similar to bank runs when expectations about other depositors' behaviour change, and the lack of cash reserves becomes widespread, requiring more delays in payments. The resulting lack of liquidity in the banking system will then lead to a breakdown of the ability to make payments at all. While banks have developed mechanisms to minimise such risks, for example with the use of net settlement systems, where payments made to a bank can be used simultaneously to make payments themselves, such risks cannot be avoided completely.

Often the focus is on the risks banks take when providing loans or making other investments, most notably in securities and real estate, and the implications these have on the borrowers themselves and any depositors. While such risks will affect depositors if they cannot be repaid, banks, and indirectly depositors, are also exposed to the risk of bank runs and a breakdown of the payment system. In terms of bank runs, this risk arises from the ability by depositors to withdraw instantly, which in many cases is an attraction to use deposits over other forms of investments, while assets cannot be liquidated sufficiently quickly to meet the demand of such withdrawals. For payments between banks this emerges from the desire of banks to preserve their liquidity and an incentive to delay payments to achieve this, causing liquidity shortage with other banks, who may not be able to complete their own payments.

While for most companies the risks are only associated with the assets of their organisation, banks have risks associated with their assets as well as their liabilities.

Part IV
Competition for banking services

Competition between banks for customers in the loan and deposit will of course affect the loan and deposit rates that banks will offer. It is a common assumption that banks will charge lower loan rates and offer higher deposit rates if competition between them increases, reducing their profits. We will therefore look at the impact of competition on the loan and deposit rates in an oligopolistic setting in chapter 20 and then in chapter 21 consider the case of monopolistic competition where banks have some degree of monopoly power due to their customers having preferences for a specific bank.

It is, however, not only that banks are competing with each other to provide loans and obtain deposits, they are also in competition with financial institutions that are not operating as banks, for example by only providing loans, but not accepting deposits; or they are only accepting deposits but investing these funds into the money market rather than granting loans. This latter arrangement is often referred to as shadow banking and we will discuss the reason for its emergence in chapter 23, where we will also include the willingness of banks to provide shadow banks with securities to invest in and thus enabling the existence of their own competitors.

It is not only shadow banks that compete with banks, other alternatives to banks exist for customers, for example they could obtain a loan directly from 'depositors' without involving a bank in what is known as peer-to-peer lending, a loan for large consumer and electronic goods as well as cars can also be obtained directly from finance companies that work jointly with the seller of these products. Furthermore, in many developing countries alternative lending form have been developed, most notably microfinance, where a group of borrowers is jointly responsible for all loans. All these lending forms are indirect competition to banks and we will explore their impact in chapter 24. While shadow banks mostly seek to attract deposits, these alternative provider seek to grant loans to customers. Hence banks face competition from shadow banks for deposits and for loans from such alternative providers.

Competition does not only affect the loan and deposit rates banks use, but may also influence the risks banks are willing to take when providing loans. We will therefore in chapter 25 explore how banks rect to increasing competition when granting loans not only through the loan rate they charge, but which risks they are willing to take in pursuing profits from lending. Competition between banks may not only affect the loan rate and the riskiness of the loans a banks provides, but as chapter 22 will show, it may also affect the size of a loan a bank is willing to grant.

All these aspects show the different influences competition has on bank behaviour. When conducting competition policy for the banking sector, many factors need to be considered to assess the benefits and costs of any regulatory measure. In particular, it might be that a measure has multiple effects, some that from a social welfare perspective are beneficial, such as reducing loan rates, while others might be detrimental, such as banks increasing the risks they are taking. All these often conflicting influences need to be weighed against each other to determine whether a specific policy is desirable.

Chapter 20

Oligopolistic markets

In this chapter we will look at the determination of loan and deposit rates by banks as they are competing for customers. Banks are offering a homogeneous good, deposits and loans, and a limited number of banks competing to obtain deposits from and grant loans to customers. In most cases models of granting loans or obtaining deposits assume that either banks are monopolists maximising their own profits or they are fully competitive, making no profits; the more realistic case of a small number of banks competing for customers is often not considered as such a framework is more difficult to analyse while not adding meaningful additional insights. Furthermore, most models focus on either the provision of loans or the ability to attract deposits, taking only a minimal interest in the other side of the balance sheet. For this reason we will look here at the effect of competition between banks and how the interest rates on loans and deposits are connected.

We will firstly in chapter 20.1 consider how in oligopolistic markets deposit and loan rates, charged for two different 'products', are related through the balance sheet and how this leads to deposit and loan rates influencing each other. Banks are not only setting interest rates to maximize their profits in light of competition from other banks, but also seek to manage their own balance sheet structure. Banks seek to provide loans if they have excess deposits and need to attract additional deposits or prevent deposits being withdrawn, if they have a large amount of loans outstanding. We will in chapter 20.2 evaluate how loan and deposit rates are set in such circumstances, how they relate if the position of loans and deposits change, as well as how the level of competition affects these determinants.

20.1 The impact of bank failures and deposit insurance

Let us assume that we have a market with N banks, each lending a total of L_i to M different borrowers, such that each loan is of size $\frac{L_i}{M}$ and the total lending in the market is given by $L = NL_i$. A loan is repaid with probability π, in which case the bank receives back the loan including interest r_L. If the borrower is not successful with his investments, the bank receives no repayment. The defaults of borrowers are assumed to be independent of each other.

The number of borrowers repaying their loans follows a binomial distribution such that the probability of exactly m repayments to a single bank is

$$f(m) = \binom{M}{m} \pi^m (1-\pi)^{M-m}.$$

If the revenue from the repaid loans as well as any investment banks make into cash, R_i, yielding interest r, falls short of their obligation to depositors, D_i, on which they pay interest r_D, the bank will make a loss. Thus

$$m(1+r_L)\frac{L_i}{M} + (1+r)R_i - (1+r_D) \leq 0, \tag{20.1}$$

The threshold for the number of repayments needed for the bank to break even is then given by

$$m \geq m^* = \frac{M}{(1+r_L)L_i}\left((1+r_D)D_i - (1+r)R\right). \tag{20.2}$$

In the absence of equity, a bank with limited liability would fail if making a loss. Hence $F(m^*) = \int_0^{m^*} f(m)\,dm$ can be interpreted as the probability of the bank failing as it represents the probability of the bank receiving less than m^* loan repayments and thus the probability of making a loss.

We assume the demand for total loans, NL_i, and deposits, ND_i, to be affected by the loan and deposit rate, respectively. The elasticities of demand for deposits and loans are defined as

$$\begin{aligned} \varepsilon_D &= \frac{\partial D}{\partial(1+r_D)}\frac{1+r_D}{D} > 0, \\ \varepsilon_L &= \frac{\partial L}{\partial(1+r_L)}\frac{1+r_L}{L} < 0. \end{aligned} \tag{20.3}$$

We can now consider the optimal behaviour of competing banks in this setting. We will distinguish banks with unlimited and limited liability as well as consider the impact of deposit insurance.

Unlimited liability If banks have unlimited liability, they cannot fail and hence we do not need to consider a situation where banks make losses.

With loan repayments following a binomial distribution as detailed above, the expected number of loan repayments for each bank is πM. With each loan having a size of $\frac{L_i}{M}$, we have the profits of the bank given as

$$\Pi_B^i = \pi(1+r_L)L_i + (1+r)R_i - (1+r_D)D_i. \tag{20.4}$$

Banks use their deposits to invest into loans and cash reserves, hence when ignoring the existence of equity, we have

$$D_i = L_i + R_i, \tag{20.5}$$

such that equation (20.4) becomes

$$\Pi_B^i = (\pi (1 + r_L) - (1 + r)) L_i - ((1 + r_D) - (1 + r)) D_i. \tag{20.6}$$

Maximizing these profits using optimal loan and deposit rates gives us the first order conditions

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial (1 + r_L)} &= (\pi (1 + r_L) - (1 + r)) \frac{\partial L_i}{\partial (1 + r_L)} + \pi L_i = 0, \\ \frac{\partial \Pi_B^i}{\partial (1 + r_D)} &= ((1 + r_D) - (1 + r)) \frac{\partial D_i}{\partial (1 + r_D)} + D_i = 0. \end{aligned} \tag{20.7}$$

Using $L = NL_i$ and $D = ND_i$ for the total demand for loans and deposits in the economy as well as equstion (20.3), this solves for

$$\begin{aligned} 1 + r_L &= \frac{1 + r}{\pi} \frac{N\varepsilon_L}{1 + N\varepsilon_L}, \\ 1 + r_D &= (1 + r) \frac{N\varepsilon_D}{1 + N\varepsilon_D}. \end{aligned} \tag{20.8}$$

We note that for large N the interest rates converge to competitive levels. To see this, assume that in equation (20.6) the interest rates are given, as is commonly assumed in competitive markets, and optimizing for L_i and D_i, respectively, yields the result same result as in equation (20.6) if we let the number of banks, N increase to infinity.

We furthermore see that loan and deposit rates are set independently of each other, depending solely on the properties of their respective markets, represented by the elasticities, and the return on cash as an anchor. We easily obtain that

$$1 + r_L \geq 1 + r \geq 1 + r_D. \tag{20.9}$$

We also easily see that the difference between loan and deposit rates is reducing the more banks, N, are present, thus the more competitive markets are. The ability of the bank to pay depositors are not effected by lending decisions (loan defaults) due to the unlimited liability and will thus be unaffected by conditions in the loan market. It is for this reason that the loan rate does not affect the deposit rate. Similarly, the deposit rate has no effect on the loan rate; the demand for loans is not affected by deposit rates and hence the profits banks make from loans are not directly affected by the deposit market. However, loan and deposit rates are connected through the risk-free rate, r, and any change of this interest rate will affect both in the same way, making loan and deposit rates move in the same direction.

Having seen how imperfect competition between banks with unlimited liabilities will result in loan and deposit rates being set independently, we can now continue with the more realistic case of banks having limited liability and thus being able to fail.

Limited liability If banks have limited liability, they will not be able to cover any loses but instead the bank will fail if it is not profitable. In this case, the profits to the bank are deemed to be zero as the owners of the bank do not obtain any proceeds. Hence using equation (20.2), if the number of loans repaid is sufficiently low, $m \leq m^*$, the bank fails and profits are then given by

$$\begin{aligned}\Pi_B^i &= \int_{m^*}^{M} \left(m\left(1+r_L\right)\frac{L_i}{M} + \left(1+r\right)R_i\right. \\ &\quad \left. - \left(1+r_D\right)D_i\right) f\left(m\right) dm \\ &= \left(1+r_L\right)\frac{L_i}{M}\int_{m^*}^{M}\left(m-m^*\right) f\left(m\right) dm \\ &= \left(1+r_L\right)\frac{L_i}{M}\left(M-m^*-\int_{m^*}^{M} F\left(m\right) dm\right) \\ &= \left(1+r_L\right)L_i + \left(1+r\right)R_i - \left(1+r_D\right)D_i \\ &\quad - \left(1+r_L\right)\frac{L_i}{M}\int_{m^*}^{M} F\left(m\right) dm,\end{aligned} \tag{20.10}$$

where the third line is obtained by integration in parts.

If the bank fails, depositors will not obtain all their deposits back. If $m \geq m^*$ they will receive the deposit including interest, but for $m < m^*$ seize all revenue the bank obtains to be paid partially. The effective interest $\hat{r}_D$ depositors obtain on their investment is then given by

$$\begin{aligned}\left(1+\hat{r}_D\right)D_i &= \int_{m^*}^{M}\left(1+r_D\right)D_i f\left(m\right) dm \\ &\quad + \int_0^{m^*}\left(m\left(1+r_L\right)\frac{L_i}{M} + \left(1+r\right)R_i\right) f\left(m\right) dm \\ &= \int_0^{M}\left(1+r_D\right)D_i f\left(m\right) dm \\ &\quad + \int_0^{m^*}\left(m\left(1+r_L\right)\frac{L_i}{M} + \left(1+r\right)R_i\right. \\ &\quad \left. - \left(1+r_D\right)D_i\right) f\left(m\right) dm \\ &= \left(1+r_D\right)D_i + \left(1+r_L\right)\frac{L_i}{M}\int_0^{m^*}\left(m-m^*\right) f\left(m\right) dm \\ &= \left(1+r_D\right)D_i - \left(1+r_L\right)\frac{L_i}{M}\int_0^{m^*} F\left(m\right) dm,\end{aligned} \tag{20.11}$$

using again equation (20.2) and integration by parts. The first expression denotes the deposits repaid with interest in those cases where the bank does not default and the second term denotes the case where the bank defaults, with depositors obtaining the entire proceeds of those loans that have been repaid as well as the cash reserves.

Solving for $(1 + r_D) D_i$ and inserting into equation (20.10) we get the bank profits after merging integrals as

$$\begin{aligned} \Pi_B^i = {} & (1 + r_L) L_i + (1 + r) R - (1 + \hat{r}_D) D_i \\ & - (1 + r_L) \frac{L_i}{M} \int_0^M F(m)\, dm. \end{aligned} \tag{20.12}$$

Using equation (20.5), we obtain the first order conditions for a profit maximum as

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial (1 + \hat{r}_D)} &= ((1 + r) - (1 + \hat{r}_D)) \frac{\partial D_i}{\partial (1 + \hat{r}_D)} - D_i = 0, \\ \frac{\partial \Pi_B^i}{\partial (1 + r_L)} &= \left((1 + r_L) \frac{\partial L_i}{\partial (1 + r_L)} + L_i \right) \\ & \quad \times \left(1 + \frac{1}{M} \int_0^M F(m)\, dm \right) \\ & \quad - (1 + r) \frac{\partial L_i}{\partial (1 + r_L)} = 0. \end{aligned} \tag{20.13}$$

Denoting by $\hat{\varepsilon}_D$ the demand elasticity with respect to $1 + \hat{r}_D$, we solves these first order conditions for

$$\begin{aligned} 1 + \hat{r}_D &= (1 + r) \frac{N \hat{\varepsilon}_D}{1 + N \hat{\varepsilon}_D}, \\ 1 + r_L &= \frac{1 + r}{M + \int_0^M F(m)\, dm} \frac{N M \varepsilon_L}{1 + N \varepsilon_L}. \end{aligned} \tag{20.14}$$

We again observe that the loan rate is only given by market characteristics of the loan market, represented by the elasticity of demand, as well as the probabilities of loans being repaid, which is included in $F(m)$. Hence $\frac{M}{M + \int_0^M F(m) dm}$ represents the aggregate probability of loans being repaid, adjusted for the fact that if the bank fails, the size of losses are irrelevant in our considerations.

From equation (20.11) we see that $\hat{r}_D$ will also depend on characteristics of the loan market through r_L and indirectly m^*. This is because a higher loan rate increases the resources available to depositors from those loans that are being repaid and thus reduces the risk of bank failures and thereby allows for lower deposit rates.

If competition increases, the loan rate will converge to competitive levels of the return on cash reserves, adjusted for the risks the bank faces due to loans not being repaid. The deposit rate will similarly increase to its competitive level of the return on cash reserves, again adjusted by the final term in equation (20.11) to account for the possible bank failure.

Thus the main difference to the case of unlimited liability is, apart from adjustments for bank failure, that the deposit rate is now linked to conditions of the loan market via its influence on the risk of bank failure. The loan rate is, however, not affected by the conditions in the deposit market; this is because the returns earned by the bank on loans are only affected by the risks of these loans, but not the deposits used to finance them.

We thus see the impact limited liability has if this may result in the bank failing and not being able to meet their obligations to depositors in full. In this case the deposit rate will take into account the risks of the loans the bank has provided. Given the importance of the losses to depositors, we will now investigate the impact a deposit insurance has on the setting of loan and deposit rates.

Deposit insurance If deposit insurance is implemented, deposits are covered and would not make losses if a bank fails. The deposit insurance will cover the shortfall in the assets of the bank such that depositors are repaid in full; this shortfall is given by

$$\begin{aligned} S_i &= -\int_0^{m^*} \Big(m\left(1+r_L\right)\frac{L_i}{M} + \left(1+r\right)R_i \\ &\quad - \left(1+r_D\right)D_i\Big) f\left(m\right) dm \\ &= -\left(1+r_L\right)\frac{L_i}{M}\int_0^{m^*}\left(m-m^*\right) f\left(m\right) dm \\ &= \left(1+r_L\right)\frac{L_i}{M}\int_0^{m^*} F\left(m\right) dm, \end{aligned} \tag{20.15}$$

where we used equation (20.2) to obtain the last equality integrated by parts. We assume that banks have to pay a premium for the deposit insurance which is set as

$$P_i = \gamma D_i + \alpha\frac{S_i}{1+r}. \tag{20.16}$$

Banks pay a fraction $\alpha \leq 1$ of the costs of deposit insurance, which is discounted to the present value. If deposit insurance is implicit, e.g. an unfunded government guarantee, then $\alpha = 0$. We also include the possibility that deposit insurance is funded through a levy on deposits ($\gamma > 0$), as is often the case, but not based on the risks involved.

If we assume that the premium is to be paid upfront, equation (20.5) changes to

$$D_i = L_i + R_i + P_i, \tag{20.17}$$

such that deposits are used to finance loans, cash reserved and the deposit insurance premium.

Inserting this expression and equation (20.16) into equation (20.10), we obtain the bank profits as

$$\Pi_B^i = ((1+r_L)-(1+r))\,L_i - ((1+r_D)-(1+r))\,D_i \qquad (20.18)$$
$$-(1+r_L)\frac{L_i}{M}\left(\int_{m^*}^{M} F(m)\,dm + \alpha\int_0^{m^*} F(m)\,dm\right).$$

Similarly we obtain equation (20.2) as

$$m^* = \frac{M}{(1+r_L)\,L_i}\left(((1+r_D)-(1+r)(1-\gamma))\,D_i - (1+r)\,L_i\right) \qquad (20.19)$$
$$+\alpha\int_0^{m^*} F(m)\,dm.$$

Noting that

$$\frac{\partial}{\partial(1+r_j)}\int_0^{m^*} F(m)\,dm = \frac{\partial m^*}{\partial(1+r_j)}F(m^*), \qquad (20.20)$$
$$\frac{\partial}{\partial(1+r_j)}\int_{m^*}^{M} F(m)\,dm = -\frac{\partial m^*}{\partial(1+r_j)}F(m^*),$$

where $j \in \{L, D\}$ for loan and deposit rates, respectively. We easily obtain that

$$\frac{\partial m^*}{\partial(1+r_D)} = \frac{M}{(1+r_L)\,L_i}\Bigg(((1+r_D) \qquad (20.21)$$
$$-(1+r)(1-\gamma))\frac{\partial D_i}{\partial(1+r_D)} + D_i\Bigg)$$
$$+\alpha\frac{\partial m^*}{\partial(1+r_D)}F(m^*),$$

which can be solved for

$$\frac{\partial m^*}{\partial(1+r_D)} = \frac{1}{1-\alpha F(m^*)}\frac{M}{(1+r_L)\,L_i} \qquad (20.22)$$
$$\times\left((1+r_D)+(1+r)(1-\gamma)\frac{\partial D_i}{\partial(1+r_D)} + D_i\right).$$

Similarly we have

$$\frac{\partial m^*}{\partial(1+r_L)} = -M\frac{\left\{\begin{array}{c}(1+r_L)\,L_i^2\\ -(((1+r_D)-(1+r)(1-\gamma))\,D_i\\ -(1+r)\,L_i)\end{array}\right\}}{(1+r_L)^2\,L_i^2} \qquad (20.23)$$
$$+\alpha\frac{\partial m^*}{\partial(1+r_L)}F(m^*),$$

solving for

$$\frac{\partial m^*}{\partial (1+r_L)} = -\frac{M}{1-\alpha F(m^*)} \times \frac{\left\{ \begin{array}{c} (1+r_L) L_i^2 \\ -(((1+r_D)-(1+r)(1-\gamma)) D_i \\ -(1+r) L_i) \\ \times \left(L_i + (1+r_L) \frac{\partial L_i}{\partial(1+r_L)} \right) \end{array} \right\}}{(1+r_L)^2 L_i^2}. \tag{20.24}$$

Using these results, we get the first order condition from maximizing bank profits over the deposit rate as

$$\begin{aligned} \frac{\partial \Pi_B}{\partial (1+r_D)} &= \left(((1+r_D)-(1+r)(1-\gamma)) \frac{\partial D_i}{\partial (1+r_D)} + D_i \right) \\ &\quad \times \frac{1+F(m^*)}{1-\alpha F(m^*)} \\ &= 0, \end{aligned} \tag{20.25}$$

giving us the optimal deposit rate as

$$1 + r_D = (1+r)(1-\gamma) \frac{N\varepsilon_D}{1+N\varepsilon_D}. \tag{20.26}$$

The deposit rate is solely determined by the deposit market, adjusted for the additional costs of the deposit insurance that is based on the size of deposits, γ. As the deposits are safe due to deposit insurance, depositors have no reason to take into account the loan market and its risks. Only the costs charged on deposits are transferred to depositors by reducing the deposit rate. The cost of deposit insurance based on the actual risk (α) is not borne by depositors as this is an element of the loan market.

The first order condition for the optimal loan rate is given by

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial (1+r_L)} &= \left(L_1 + (1+r_L) \frac{\partial L_i}{\partial (1+r_L)} \right) \\ &\quad \times \left(1 - \frac{1}{M} \int_{m^*}^{M} F(m)\, dm - \frac{\alpha}{M} \int_0^{m^*} F(m)\, dm \right) \\ &\quad + (1-\alpha) F(m^*) (1+r_L) \frac{L_i}{M} \frac{\partial m^*}{\partial (1+r_L)} = 0. \end{aligned} \tag{20.27}$$

Here we see that in general the loan rate will also depend on the deposit rate through its impact on m^*. As high deposit rates make bank failures more likely, given the increased commitment of the bank, a higher loan rate is required to maintain the risk of bank failure and reduce the premium paid for deposit insurance.

If the bank covers the costs of deposit insurance fully, i.e. $\alpha = 1$, we see that the final term in equation (20.27) vanishes, leaving the loan rate to be determined by the first bracketed term. This is not including any terms of the deposit market, but relies solely on the loan market. The result is actually identical to the case of unlimited liability as represented from equation (20.8). The reason is that in this case the bank has fully internalized the costs of bankruptcy through the premium and hence the case is indistinguishable from unlimited liability. If deposit insurance is levied solely on deposit size, this has no effect on the dependence of the loan rate on the deposit rate as the absence of dependency on the risk of the bank, does not allow a full internalization of the bankruptcy costs.

Summary We have seen that if banks have unlimited liability or their bankruptcy costs get internalized through a fully funded deposit insurance, loan and deposit rates are set independently. This is because depositors' outcomes are unaffected by decisions in the loan market and deposit rates have no bearing on the net bankruptcy costs.

If we introduce limited liability without deposit insurance, depositors are affected by the loan rates as higher loan rates make more funds available to cover the deposits, reducing the risk of bank failure. Importantly, we have here assumed that higher loan rates do not affect the likelihood of default by companies. When introducing a partially funded, unfunded or not risk-based funded deposit insurance, the causality reverses. Now the deposit rate is solely based on the deposit market, which arises from the fact that depositors are always fully repaid regardless of the bank's situation, while the loan rate will be affected by the deposit rate. Lower deposit rates reduce the liabilities of the bank in the case of losses and thus reduces the deposit insurance premium and likelihood of bank failure; this will in turn reduce the costs the bank faces and thus reduce loan rates. While in both cases, with partially funded deposit insurance and in the absence of deposit insurance, there will be a positive relationship between loan and deposit rates, the casuality of these relationships is different, though.

Loan and deposit markets are comparable to any other markets, however the price at which these two 'products' are offered by banks will interact in a realistic scenario. The risks taken by banks when providing loans and the interest charged on such loans will affect the ability of banks to repay any deposits, thus affecting the deposit rate that banks are offering. If deposit insurance is charged to banks, the premium will depend on the potential liability of this scheme, which is affected by the deposit rate; while deposits are safe and thus not affected by these considerations, the costs of the deposit insurance will impact the loan rate as the additional costs are charged to borrowers.

Readings Dermine (1986), Klein (1971)

20.2 Balancing cash flows

We might view banks as intermediaries between depositors and borrowers that balance cash inflows resulting from new deposits D_i and cash outflows arising from the provision of new loans L_i. Ignoring other sources of income, in this case the bank could derive value only from any imbalance between the total loans L and total deposits D, i.e. $L - D$. Let us further assume that banks price deposits and loan rates relative to a common benchmark r with a discount γ_D for deposits and a surcharge γ_L for loans such that

$$1 + r = (1 + r_D)(1 + \gamma_D) = \pi (1 + r_L)(1 - \gamma_L), \tag{20.28}$$

where π denotes the repayment rate of loans. The benchmark r is assumed to be stochastic with mean μ and variance σ^2. This benchmark might represent the current risk-free rate or the lending facility of the central bank, which is fluctuating randomly over time. At this interest rate the bank might deposit any excess cash they hold or they are able to obtain funding in case of a cash shortage. Based on these ideas we then have for the profits of banks that

$$\begin{aligned} \Pi_B &= (1 + r)(L - D), \\ E[\Pi_B] &= (1 + \mu)(L - D), \\ Var[\Pi_B] &= \sigma^2 (L - D)^2. \end{aligned} \tag{20.29}$$

The profits of the bank considered here are only those of the interest earned on any cash holdings or the interest paid on any cash shortages. The revenue earned on providing loans and the costs of deposits are ignored here.

With risk aversion z the expected utility of the bank is then

$$U_B = E[\Pi_B] - \frac{1}{2} z Var[\Pi_B]. \tag{20.30}$$

If the bank receives additional deposits D_i, the new deposits become $D + D_i$ and the additional deposits are then kept as cash until a loan request arrives, earning interest at the rate r, or reducing the interest paid. As the deposits the bank has obtained only attract interest r_D, the additional value generated for the bank is the difference to the benchmark interest rate, r. Thus the bank generates value equivalent to $\gamma_D D_i$, the discount on the deposits. The expected utility is then given by

$$\begin{aligned} U_B^D = {} & (1 + \mu)(L - (D + D_i)) + \gamma_D D_i \\ & - \frac{1}{2} z \sigma^2 (L - (D + D_i))^2, \end{aligned} \tag{20.31}$$

when using equation (20.29) in equation (20.30) and taking into account the additional value generated to the bank.

Similarly, a loan request increases L to $L + L_i$ and reduces the amount of cash by L_i, but as the loan is charged r_L, the value generated by the bank increased is reduced by $\gamma_L L_i$ as the bank obtains the surcharge on this additional loan. Thus as before the expected utility of the bank is given by

$$U_B^L = (1+\mu)\left((L+L_i)-D\right)+\gamma_L L_i \\ -\frac{1}{2}z\sigma^2\left((L+L_i)-D\right)^2. \tag{20.32}$$

New deposits and requests for loans are assumed to arrive randomly with arrival rates λ_D and λ_L, respectively, in each time period. We assume that this arrival rate, which we can interpret as the probability of a deposit is received or loan made in a given time period, is given by

$$\lambda_i = \alpha_0 - \alpha_1\gamma_i. \tag{20.33}$$

The higher the surcharge and thus interest rate on loans is or the larger the discount and lower the deposits is, the less likely they are to arrive as loans and deposits become less attractive. We can interpret the value of α_1 as an indication of the level of competition. A higher value for this parameter implies that customers react more sensitively to interest rate changes, for example by moving to other banks. In perfect competition we would observe a very high value of α_1, while for a banking system with less composition this value would be lower.

We firstly get the increase in utility after receiving these additional deposits and loans from equations (20.31) and (20.32) by subtracting equation(20.30) as

$$\begin{aligned} \Delta U_B^D &= U_B^D - U_B \\ &= -(1+\mu)D_i + \gamma_D D_i - \frac{1}{2}z\sigma^2\left(-2(L-D)D_i + D_i^2\right), \\ \Delta U_B^L &= U_B^L - U_B \\ &= (1+\mu)L_i + \gamma_L L_i - \frac{1}{2}z\sigma^2\left(2(L-D)L_i + L_i^2\right). \end{aligned} \tag{20.34}$$

The expected utility increase is given by the likelihood (arrival rate) of new deposits and loans, thus

$$\Delta U_B = \lambda_D \Delta U_B^D + \lambda_L \Delta U_B^L. \tag{20.35}$$

Inserting from equations (20.33) and (20.34), we get the first order conditions for the optimal discount and surcharge as

$$\begin{aligned} \frac{\partial \Delta U_B}{\partial \gamma_D} &= -\alpha_1 \Delta U_B^D + \lambda_D D_i = 0, \\ \frac{\partial \Delta U_B}{\partial \gamma_L} &= -\alpha_1 \Delta U_B^L + \lambda_L L_i = 0. \end{aligned} \tag{20.36}$$

Inserting as necessary, this solves for the optimal discount and surcharge, which become

$$\begin{aligned}\gamma_D &= \frac{\alpha_0}{2\alpha_1} + \frac{\mu + \frac{1}{2}z\sigma^2\left(D_i - 2\left(L - D\right)\right)}{2} \\ &= \frac{\alpha_0 + \alpha_1\mu}{2\alpha_1} + \frac{z\sigma^2\left(D_i - 2\left(L - D\right)\right)}{4}, \\ \gamma_L &= \frac{\alpha_0 - \alpha_1\mu}{2\alpha_1} + \frac{z\sigma^2\left(2\left(L - D\right) + L_i\right)}{4}.\end{aligned} \tag{20.37}$$

We see that in the case that there are more loans than deposits, $L - D > 0$, γ_D (γ_L) decreases (increases), thus increasing both deposit and loan rates when looking at equation (20.28). This increase in interest rates attracts deposits due to higher rates and makes taking loans less attractive, reducing their uptake. This has the combined effect of making it more likely a deposit is obtained rather than a loan, causing $L - D$ to revert back to zero, which here is the optimal level as cash is not attractive to hold given that this implies funding at the risk-free rate r, while funding through deposits is less costly as they are available at a discount. Similarly, if the bank holds more deposits than it has granted loans, $L - D < 0$, the bank seeks to attract more loans as the loan rate will exceed the benchmark interest rate and thus be more attractive. The bank will thus reduce the loan and deposit rates to encourage the uptake of loans and discourage the inflow of additional deposits.

The total difference between loan and deposit rate, the spread as measured by the surcharges γ_D and γ_L is given from equation (20.37) by

$$\gamma_D + \gamma_L = \frac{\alpha_0}{\alpha_1} + \frac{1}{2}z\sigma^2 D_i, \tag{20.38}$$

if we assume $D_i = L_i$ for simplicity. The spread will not depend on the imbalance of loans and deposits, $L - D$, but move in parallel. We also see that with higher levels of competition, α_1, the difference between loan and deposit rates will reduce by increasing deposit rates and reducing loans rates, in line with most economic models.

We have seen that banks will adjust deposit and loan rates to manage the demand for deposits and loans, seeking to attract loans by lowering loan rates if they have more deposits than loans; at the same time they lower the deposit rate to discourage any more deposits. If banks hold more loans than they have deposits, they woudl increase loans rates to discourage the demand for new loans and increase the deposit rate to attract new deposits.

Reading Ho & Saunders (1981)

Conclusions

We have seen that in realistic cases of banks having limited liability and thus can potentially fail, loan and deposit rates are positively related. This relationship arises from the risks depositor exposed to the possibility of banks failing and they will therefore take into account this possibility. A higher loan rate increases the funds available to banks from those loans that are repaid, allowing banks to make larger payments to depositors if a bank fails, thus reducing the required deposit rate. On the other hand, if depositors are protected by deposit insurance, depositors will not be affected by the loans provided. However, the costs of such insurance is increasing as the deposit rate is increasing and these costs will be internalised by banks having to pay for the deposit insurance; the consequence is that loan rates are increasing in the deposit rate. Thus, even though deposit and loan markets address different needs by customers and can be seen as independent from each other, these two markets are connected through the balance sheet of banks and thus affect each other.

Banks are not only directly competing with each other, they are also seeking to manage their balance sheets and will set loan and deposit rates accordingly to ensure they attract the right proportions of deposits and loans. We have seen that loan and deposit rates are moving in parallel to achieve these aims, the differences between these interest rates reducing as competition between banks increases.

Chapter 21

Monopolistic competition

Banks are not only competing with other on the basis of the loan and deposit rates they can offer, but will also attract customers through a range of services they offer. Such services might be a branch network for those preferring personal contact with banking staff, but also the use of online and mobile services for those preferring this format. More specifically it may also include additional benefits of accounts, the flexibility of loan contracts in terms of repayments or taking advantage of previously repaid loans by taking out additional funds if desired. Some banks might limit the amounts that can be withdrawn from deposits within a certain time period, while other banks might not impose such constraints.

Customers will in general have different preferences for such services and account benefits as well as loan and deposit conditions and thus might, everything else being equal, prefer one bank over another for this very reason. We will now take into account such preferences of bank customers and assess how their presence affects the competition between banks. In chapter 21.1 we will consider how other wise identical banks will set their loan and deposit rates as well as their reaction to regulatory constraints on the loan rate they can charge. Chapter 21.2 will then focus on the provision of risky loans, whose risks are covered by the profits of risk-free loans that are given under monopolistic competition between banks of different sizes. We will see how increasing competition between banks can reduce the ability of banks to provide such risky loans, but also how in some conditions, banks might be able to expand risky loans if competition increases.

21.1 Monopolistic competition

Let us assume that banks are not offering a completely homogeneous 'product', but instead differentiate themselves by the services they offer around loans and deposits. Each customer will have their own preferences with respect to the service they ideally would obtain. As banks cannot offer a tailored service to each customer, their customers in general will not obtain their preferred services. We can measure a

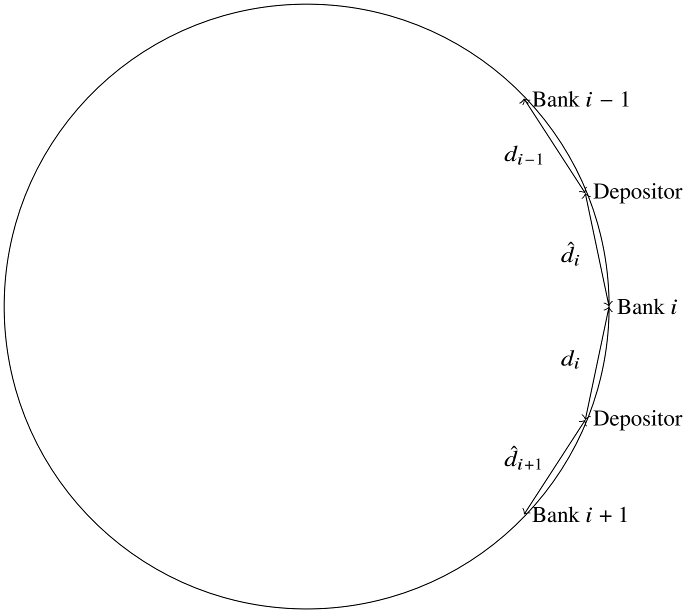


Fig. 21.1: Location of banks offering differentiated services

"distance" between the desired services and the services actually obtained, denoted by d_i for the distance to bank i. We assume that each customer faces costs c per unit of distance for each unit of loan or deposit; then the total costs for a customer with loan L and deposit D are given by cd_iD and cd_iL, respectively.

The N banks in our banking system are assumed to be located on a circle with a circumference normalized to 1 for convenience. Banks will locate themselves equidistant $\frac{1}{N}$ apart on this circle if the assume the M customers are distributed equally around the circle. Figure 21.1 illustrates this setup, where we denote by d_i the distance of a depositor located between banks i and $i+1$, while $\hat{d}_i$ denotes the distance of a depositor located between banks i and $i-1$. Obviously we have that the distance to both nearest banks combined are the distance of the banks together, $d_i + d_{i+1} = \hat{d}_i + \hat{d}_{i+1} = \frac{1}{N}$.

We will use this arrangement of monopolistic competition to assess the optimal determination of loan and deposit rates in cases where the banking market is unregulated and where a regulator imposes a maximum loan rate on banks. We will compare both situations and allow banks to require companies seeking a loan to deposit their funds with the same bank.

Unregulated markets If bank i pays interest r_D^i on deposits D, then when going to bank i, a depositor makes profits

$$\Pi_D^i = \left(1 + r_D^i\right) D - D - cd_iD, \tag{21.1}$$

where the final term accounts for the costs of receiving services which are not exactly as the depositor prefers. We assume here that banks will repay deposits with certainty and hence depositors face no risk.

In order for a depositor to be indifferent between going to bank i or bank $i+1$, we need the profits of using these two banks to be equal, thus $\Pi_D^i = \Pi_D^{i+1}$. Similarly, for indifference between bank i and $i-1$ we need $\Pi_D^i = \Pi_D^{i-1}$. By solving these two equations, we obtain the location of depositors that are indifferent between attending bank i and $i+1$ or i and $i-1$, as

$$\begin{aligned} d_i &= \frac{1}{2N} + \frac{\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right)}{2c} \\ \hat{d}_i &= \frac{1}{2N} + \frac{\left(1+r_D^i\right) - \left(1+r_D^{i-1}\right)}{2c}, \end{aligned} \tag{21.2}$$

where we have used that $d_i + d_{i+1} = \hat{d}_i + \hat{d}_{i+1} = \frac{1}{N}$.

As all depositors located closer to bank i will use this bank, we can interpret $d_i + \hat{d}_i$ as the fraction of depositors using bank i, remembering that the circumference of the circle was normalized to 1. Hence $\left(d_i + \hat{d}_i\right) M$ depositors use bank i, each depositing the amount of D, such that the total deposits of bank i are given by

$$\begin{aligned} D_i &= \left(d_i + \hat{d}_i\right) MD \\ &= \left(\frac{1}{N} + \frac{2\left(1+r_D^i\right) - \left(1+r_D^{1+1}\right) - \left(1+r_D^{i-1}\right)}{2c}\right) MD. \end{aligned} \tag{21.3}$$

In the same way, we can derive the amount of loans given by bank i. The borrowers make profits

$$\Pi_C^i = \pi\left(\left(1+R\right)L - \left(1+r_L^i\right)L\right) - cd_iL, \tag{21.4}$$

where π denotes the probability of success of their investment, R the return on the successful investment and r_L^i the loan rate bank i charges and L the loan amount to the company. Indifference to neighboring banks requires analogously to deposits that $\Pi_C^i = \Pi_C^{i+1}$ and $\Pi_C^i = \Pi_C^{i-1}$, hence

$$\begin{aligned} d_i &= \frac{1}{2N} - \frac{\left(1+r_L^i\right) - \left(1+r_L^{i+1}\right)}{2c} \\ \hat{d}_i &= \frac{1}{2N} - \frac{\left(1+r_L^i\right) - \left(1+r_L^{i-1}\right)}{2c}, \end{aligned} \tag{21.5}$$

such that the total loans supplied by bank i are

$$\begin{aligned} L_i &= \left(d_i + \hat{d}_i\right) ML \\ &= \left(\frac{1}{N} - \frac{2\left(1+r_L^i\right) - \left(1+r_L^{1+1}\right) - \left(1+r_L^{i-1}\right)}{2c}\right) ML. \end{aligned} \tag{21.6}$$

In both cases we observe that if the deposit (loan) rate of bank i is higher (lower), the deposits (loans) the bank has, increases. The extend of this increase is limited, though, by the costs c as changing to a more distant bank reduces their profits, which need to be compensated for by better interest rates, thus higher deposit rates or lower loan rates.

Banks use the deposits D_i to invest into loans L_i and cash reserves R_i, yielding a return r, hence

$$D_i = L_i + R_i. \tag{21.7}$$

With fixed costs F of running the bank, we have the bank profits given by

$$\begin{aligned} \Pi_B^i &= \pi\left(1+r_L^i\right)L_i + (1+r)R_i - \left(1+r_D^i\right)D_i - F \\ &= \left(\pi\left(1+r_L^i\right) - (1+r)\right)L_i - \left(\left(1+r_D^i\right) - (1+r)\right)D_i - F, \end{aligned} \tag{21.8}$$

using equation (21.7). Noting that

$$\begin{aligned} \frac{\partial D_i}{\partial\left(1+r_D^i\right)} &= \frac{MD}{c}, \\ \frac{\partial L_i}{\partial\left(1+r_L^i\right)} &= -\pi\frac{ML}{c} \end{aligned} \tag{21.9}$$

and $D_i = \frac{MD}{N}$, $L_i = \frac{ML}{N}$ in equilibrium as, apart from their location, all banks are identical and would this charge the same interest rates, such that from (21.3) and (21.6) the second term in the brackets will vanish.

The first order conditions for a profit maximum by choosing optimal deposit and loans rates are then given by

$$\begin{aligned} \frac{\partial\Pi_B^i}{\partial\left(1+r_D^i\right)} &= -\left(\left(\left(1+r_D^i\right)-(1+r)\right)\frac{\partial D_i}{\partial\left(1+r_D^i\right)} + D_i\right) \\ &= -\left(\left(\left(1+r_D^i\right)-(1+r)\right)\frac{1}{c} + \frac{1}{N}\right)MD \\ &= 0, \\ \frac{\partial\Pi_B^i}{\partial\left(1+r_L^i\right)} &= \left(\left(\pi\left(1+r_L^i\right)-(1+r)\right)\frac{\partial L_i}{\partial\left(1+r_L^i\right)}\right) + \pi L_i \\ &= -\left(\left(\pi\left(1+r_L^i\right)-(1+r)\right)\frac{1}{c} + \frac{1}{N}\right)ML \\ &= 0, \end{aligned} \tag{21.10}$$

which solves for

$$\begin{aligned} 1+r_D^i &= (1+r) - \frac{c}{N}, \\ 1+r_L^i &= \frac{1+r}{\pi} + \frac{c}{\pi N}. \end{aligned} \tag{21.11}$$

We see that as the number of banks increases, deposit rates will increase and loan rates decrease such that they approach levels in which banks make no profits on deposits or loans as they both approach the interest on cash reserves, in the case of loans adjusted for the risk of non-repayment of the loan. It is apparent that deposit and loan rates are determined independently as there are is term including the loan (deposit) rate in the first order condition for the optimal deposit (loan) rate. As these two 'products' of the bank, deposits and loans, are addressed at different customers, the demand will be determined independently; the funding costs of banks through deposits will not affect the loan rate. Of course, both interest rates are affected by the risk-free rate r and will move in the same direction if this changes.

Inserting equation (21.11) back into equation (21.8), we get the profits of the bank as

$$\Pi_B = \frac{c}{N^2} (L + D) M - F. \tag{21.12}$$

If we had free market entry for banks, competition would erode any profits such that in equilibrium $\Pi_B^{=}0$, or

$$N^* = \sqrt{\frac{c (L + D) M}{F}}. \tag{21.13}$$

Thus, due to the fixed costs of operating a bank, the number of banks in the market will be limited.

Tied contracts in unregulated markets Banks might stipulate that customers can only obtain loans if they have deposits with the same bank, hence they have to use the bank for all their banking activities. As the loan is paid into an account, it instantly becomes a deposit; thus the bank would require the company to use this newly created deposit for their investment and not withdraw this deposit to another bank. Furthermore, companies as well as individuals commonly have both loans and deposits, the latter often as a reserve for unexpected outgoings. If we assume that the preferences by customers in terms of services of the bank, are identical for loans and deposits, their profits from a combined exposure to a bank via tied contracts is given by

$$\Pi_{DC}^i = \Pi_D^i + \Pi_C^i \tag{21.14}$$

with Π_D^i and Π_C^i as given in equations (21.1) and (21.4), respectively. We again determine the location of the customer being indifferent between banks i and $i + 1$, and i and $i - 1$, respectively be setting $\Pi_{DC}^i = \Pi_{DC}^{i+1}$ and $\Pi_{DC}^i = \Pi_{DC}^{i-1}$, which solves for

$$\begin{aligned}
d_i &= \frac{1}{2N} \\
&+ \frac{\left(\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right)\right) D - \pi \left(\left(1+r_L^i\right) - \left(1+r_L^{i+1}\right)\right) L}{2c\left(L+D\right)}, \\
\hat{d}_i &= \frac{1}{2N} \\
&+ \frac{\left(\left(1+r_D^i\right) - \left(1+r_D^{i-1}\right)\right) D - \pi \left(\left(1+r_L^i\right) - \left(1+r_L^{i-1}\right)\right) L}{2c\left(L+D\right)}.
\end{aligned} \tag{21.15}$$

As the customers are tied and thus have deposits and loans with the same bank, the fraction of customers a bank has for deposits and loans are identical, hence

$$\begin{aligned}
D_i &= \left(d_i + \hat{d}_i\right) MD \\
&= \left(\frac{1}{N} + \frac{\begin{array}{c}\left(2\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right) - \left(1+r_D^{i-1}\right)\right) D \\ -\pi\left(2\left(1+r_L^i\right) - \left(1+r_L^{i+1}\right) - \left(1+r_L^{i-1}\right)\right) L\end{array}}{2c\left(L+D\right)}\right) MD \\
L_i &= \left(d_i + \hat{d}_i\right) ML \\
&= \left(\frac{1}{N} + \frac{\begin{array}{c}\left(2\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right) - \left(1+r_D^{i-1}\right)\right) D \\ -\pi\left(2\left(1+r_L^i\right) - \left(1+r_L^{i+1}\right) - \left(1+r_L^{i-1}\right)\right) L\end{array}}{2c\left(L+D\right)}\right) ML
\end{aligned} \tag{21.16}$$

Noting that

$$\begin{aligned}
\frac{\partial D_i}{\partial\left(1+r_D^i\right)} &= \frac{MD^2}{c\left(L+D\right)}, \\
\frac{\partial D_i}{\partial\left(1+r_L^i\right)} &= pi\frac{MLD}{c\left(L+D\right)}, \\
\frac{\partial L_i}{\partial\left(1+r_D^i\right)} &= \frac{MLD}{c\left(L+D\right)}, \\
\frac{\partial L_i}{\partial\left(1+r_L^i\right)} &= -\pi\frac{ML^2}{c\left(L+D\right)},
\end{aligned} \tag{21.17}$$

we get the first order conditions from maximizing equation (21.14) as

$$\begin{aligned}
\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} &= \left(\pi\left(1+r_L^i\right)-(1+r)\right)\frac{\partial L_i}{\partial\left(1+r_D^i\right)} \\
&\quad -\left(\left(\left(1+r_D^i\right)-(1+r)\right)\frac{\partial D_i}{\partial\left(1+r_D^i\right)}+D_i\right) \\
&= \left(\pi\left(1+r_L^i\right)-(1+r)\right)\frac{MLD}{c\left(L+D\right)} \\
&\quad -\left(\left(\left(1+r_D^i\right)-(1+r)\right)\frac{MD^2}{c\left(L+D\right)+\frac{MD}{N}}\right) \\
&= 0 \\
\frac{\partial \Pi_B^i}{\partial\left(1+r_L^i\right)} &= \left(\pi\left(1+r_L^i\right)-(1+r)\right)\frac{\partial L_i}{\partial\left(1+r_L^i\right)} \\
&\quad +\pi L_i-\left(\left(1+r_D^i\right)-(1+r)\right)\frac{\partial D_i}{\partial\left(1+r_L^i\right)} \\
&= -\pi\left(\pi\left(1+r_L^i\right)-(1+r)\right)\frac{ML^2}{c\left(L+D\right)} \\
&\quad +\pi\frac{LM}{N}+\left(\left(1+r_D^i\right)-(1+r)\right)\frac{MDL}{c\left(L+D\right)} \\
&= 0,
\end{aligned} \tag{21.18}$$

using again that in equilibrium all banks are equal such that $D_i=\frac{MD}{N}$ and $L_i=\frac{ML}{N}$. Both equations solve for

$$1+r_D^i=(1+r)\frac{D-L}{D}-\frac{c}{N}\frac{L+D}{D}+\pi\left(1+r_L^i\right)\frac{L}{D}. \tag{21.19}$$

This implies that with tied contracts, interest rates are no longer uniquely determined. As customers have both, deposits and loans, with the same bank, they only consider the relative prices as relevant. A high loan rate is acceptable if it is accompanied by a high deposit rate, or a low deposit rate is acceptable if the loan rate is also low. In contrast to the case of untied contracts, deposit and loan rates are closely related and cannot be set independently. This is the result of deposits and loans being tied, making the two 'products' closely related.

Inserting these results into equation (21.14) we obtain the bank profits as

$$\Pi_B^i=\frac{c}{N^2}\left(L+D\right)M-F. \tag{21.20}$$

Comparing with the bank profits for untied contract in equation (21.13), we see that the profits to banks remain unchanged with tied contracts and banks would be indifferent between individual and tied contracts.

If bank i were to offer untied contracts and sets prices according to equation (21.11), while all other banks set prices according to equation (21.19), we get for bank i from equation (21.15) that $d_i=\hat{d}_i=\frac{1}{2N}$ and inserting into the bank profits of equation (21.8) that $\Pi_B^i = \frac{c}{N^2}(L+D)M - F$. Thus profits remain unchanged and therefore tied and untied contracts can co-exist.

Having established that banks do not benefit from introducing tied contracts, we will now consider the case where a regulator limits the loan rate a bank can charge and explore the implications of such a restriction.

Regulated markets Let us assume now that a regulator stipulates a maximum loan rate $\overline{r}_L$ such that $1 + \overline{r}_L < \frac{1+r}{c} + \frac{c}{\pi N}$; in this case the constraint is binding. With untied contracts, the deposit rate is unaffected as these two markets are independent. Inserting for r_D^i in equation (21.11), we then get the bank profits as

$$\Pi_B^i = (\pi(1+\overline{r}_L) - (1+r))\frac{ML}{N} + \frac{c}{N^2}MD - F, \tag{21.21}$$

where $L_i = \frac{ML}{N}$ as borrowers would choose lenders that are closest to them.

If markets are tied we have the profits of customers given by

$$\Pi_{DC}^i = \pi((1+R) - (1+\overline{r}_L))L - cd_iL + \left(\left(1+r_D^i\right)D - D - cd_iD\right). \tag{21.22}$$

The location of the indifferent depositor from setting $\Pi_{DC}^i = \Pi_{DC}^{i+1}$ and $\Pi_{DC}^i = \Pi_{DC}^{i-1}$, is then determined as

$$\begin{aligned} d_i &= \frac{1}{2N} + \frac{\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right)}{2c}\frac{D}{L+D}, \\ \hat{d}_i &= \frac{1}{2N} + \frac{\left(1+r_D^i\right) - \left(1+r_D^{i-1}\right)}{2c}\frac{D}{L+D}, \end{aligned} \tag{21.23}$$

Hence, as contracts are tied the market shares of deposits and loans are identical, and we have

$$\begin{aligned} D_i &= \left(d_i + \hat{d}_i\right)MD \\ &= \left(\frac{1}{N} + \frac{2\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right) - \left(1+r_D^{i-1}\right)}{2c}\right)MD, \\ L_i &= \left(d_i + \hat{d}_i\right)ML \\ &= \left(\frac{1}{N} + \frac{2\left(1+r_D^i\right) - \left(1+r_D^{i+1}\right) - \left(1+r_D^{i-1}\right)}{2c}\right)ML. \end{aligned} \tag{21.24}$$

Noting that

$$\begin{aligned} \frac{\partial D_i}{\partial\left(1+r_D^i\right)} &= \frac{D}{c}\frac{D}{L+D}M, \\ \frac{\partial L_i}{\partial\left(1+r_D^i\right)} &= \frac{D}{c}\frac{L}{L+D}M, \end{aligned} \tag{21.25}$$

and the first order condition for a profit maximum becomes

$$\frac{\partial \Pi_B^i}{\partial \left(1+r_D^i\right)} = \left(\pi\left(1+\overline{r}_L\right) - (1+r)\right) \frac{D}{c} \frac{ML}{L+D} - \left(\left(1+r_D^i\right) - (1+r)\right) \frac{D}{c} \frac{DM}{L+D} - \frac{DM}{N} = 0, \tag{21.26}$$

which solves for

$$1 + r_D^1 = (1+r) \frac{D-L}{D} - \frac{c}{N} \frac{L+D}{D} + \pi\left(1+\overline{r}_L\right) \frac{L}{D}. \tag{21.27}$$

Firstly we note that again the deposit rate depends on the fixed loan rate. This is because there is no competition in the loan market and the market shares are identical due to the tied markets. Comparing with equation (21.11), we see that with a regulated loan rate the deposit rate is lower as we assumed that $1 + \overline{r}_L < (1+r) + \frac{c}{\pi N}$, which was the loan rate in an unregulated market.

Inserting equation (21.27) into equation (21.8), we obtain that bank profits remain at the unregulated level of $\Pi_B^i = \frac{c}{N^2}(L+D)M - F$. The tied contract allows the bank to fully offset the losses they make from lower loan rates by lowering the deposit rate. As the deposit rate is lower than in the unties contract, while obtaining the same loan rate, the bank profits are higher with tied contracts than with untied contracts.

Assume now that bank i offers an untied contract with the deposit rate set as in equation (21.11) and the other banks use the tied contract with deposit rates as determined in equation (21.27). Inserting this into equation (21.23), we obtain that

$$d_i + \hat{d}_i = \frac{1}{N} - \frac{(1+r) - \frac{c}{N} + \pi\left(1+\overline{r}_L\right) \frac{L}{D}}{c} + \frac{1}{N} - \frac{\left(1+r_D^i\right) + \pi\left(1+\overline{r}_L\right)}{c} < \frac{1}{N}. \tag{21.28}$$

The resulting profits of the bank we can write as

$$\Pi_B = \left(\pi\left(1+\overline{r}_L\right) - (1+r)\right)\left(d_i + \hat{d}_i\right) ML + \frac{c}{N}\left(d_i + \hat{d}_i\right) MD - F. \tag{21.29}$$

We firstly note that although the bank offers an untied contract, it will supply loans to all those it takes deposits from as they cannot obtain loans elsewhere. As $1 + \overline{r}_L < \frac{1=R}{\pi} + \frac{c}{\pi N}$, we have $\pi\left(1 + \overline{r}L\right) - (1+r) < \frac{c}{N}$ and $d_i + \hat{d}_i < \frac{1}{N}$, hence

$$\Pi_B^i < \frac{c}{N^2}(L+D)M - F. \tag{21.30}$$

Therefore, it is not profitable for a bank to deviate from offering tied contracts and in a regulated market all banks will offer tied contracts only to maintain their full profitability.

We have thus established that limiting the loan rate that banks can apply does not affect their profitability if they are able to tie the provision of a loan to the customers also maintaining their deposits at the same bank. In this case, banks are able to reduce deposit rates, allowing them to increase profits where profits from the provision of loans has been reduced due to the regulatory limit on the loan rate.

As the profits of banks are remaining identical with such regulatory restrictions, the number of banks that can profitably operate in the market as determined in equation (21.13) will remain unchanged, too. We can now compare this aspect of the market, the number of banks active in it, with the social optimum.

Social optimum From an overall social point of view the loan and deposit rate merely re-allocate surpluses between depositors, borrowers, and banks, with the aggregate surplus being unaffected. What affects the aggregate surplus, though, are the cost c that are lost due to a mismatch of preferences between customers and the services banks are able to offer. Bank also have to bear fixed costs F that are not recovered directly though interest rates.

In a social optimum, these costs should be minimized. With N banks the distance between them is $\frac{1}{N}$, hence the maximum distance (on each side of the N banks), is $\frac{1}{2N}$ and the total costs are $2N \int_0^{\frac{1}{2N}} d_i dd_i$ for borrowers and lenders combined. Hence we seek to minimize the social costs

$$C = 2Nc\left(D + L\right) M \int_0^{\frac{1}{2N}} D_i dd_i + NF = \frac{c\left(D + L\right)}{4N} M + NF, \tag{21.31}$$

which includes the fixed costs for each bank. The first order condition

$$\frac{\partial C}{\partial N} = F - \frac{c\left(D + L\right)}{4N^2} M = 0 \tag{21.32}$$

solves for

$$N^{**} = \frac{1}{2}\sqrt{\frac{c\left(D + L\right) M}{F}}. \tag{21.33}$$

Comparing this expression with equation (21.13), we see that with free market entry the number of banks active in the market is twice that of the social optimum. We thus see that the high number of feasible banks will reduce the costs to depositors and companies as the average distance to a bank reduces; this will, however, imply high fixed costs for operating such a large number of banks. From the perspective of social welfare it would be more beneficial to limit the number of banks in the market rather than impose restrictions on the loan rates that banks can charge.

Summary We have seen that in unregulated with monopolistic competition banks set loan and deposit rates independently, but that compared to the social optimum too many banks enter the market. Banks may offer tied contracts in which customers can only obtain loans if they place their deposits with the same bank, while other banks might offer untied contracts. This result changes in a regulated market where loans rates are limited; banks will only offer tied contracts. This approach allows banks to maintain their profitability by reducing the deposit rate without losing customers to competitors offering better deposit rates. It is easy to see that imposing minimum deposit rates will have a comparable effect by increasing loan rates through the use of tied contracts.

Monopolistic competition allows banks to maintain some degree of market power from due to the preferences of customers for services provided by banks; this market power can be used to cover fixed costs. If faced with regulatory constraints on the prices for their products, the interest rates on loans and deposits, banks can seek to tie the contracts in that customers can only obtain a loan if they hold their deposits with the same bank. This approach allows banks to offset the lower loan rate (higher deposit rate) imposed by regulation through lower deposit rates (higher loan rates) without affecting the competition between banks and thus their profitability.

Reading Chiappori, Perez-Castrillo, & Verdier (1995)

21.2 Provision of risky loans

Let us assume that we have two types of banks in the market, large banks (L) and small banks (S), in proportion of λ and $1 - \lambda$, respectively. Large banks have cost advantages when monitoring companies' behaviour such that for these costs we have $c_S > c_L$.

Banks provide loans to two types of companies; the first type repays their loan L, including interest to bank i of r_L^i, with certainty due to investing into a safe project, which generates a return R. Banks ensure that the safe investment is chosen through the monitoring of the company. These companies have preferences for a specific bank, for example due to the services they offer; hence banks are in monopolistic competition for these companies. We assume that a company seeking a loan approaches only two banks, who are located at either end of a line of unitary length. The types of banks, larger or small, on either end are randomly selected and companies are uniformly distributed between them. Banks know the location of companies and engage in discriminatory pricing.

The second type of company, which the banks can distinguish ex-ante from the other type of company, are those choosing risky loans. Companies might make this choice because they are difficult to monitor by banks and risky investments are more profitable for them, or do not have any access to risk-free investments. These companies do not always repay their loans $\hat{L}$, causing banks to make a loss of $(1 + r_D)\,\hat{L}$ from having to repay deposits which financed the loans when not obtaining any repayments on the loan, where r_D denotes the interest on the deposits.

We assume that by regulation banks are not allowed to make any losses, even if risky loans fail. Hence they can only give risky loans to the extend that the profits from the safe loans cover any potential losses. Denoting by Π_B the profits of the bank from safe loans, we need $\Pi_B \geq (1 + r_D)\,\hat{L}$, and hence risky loans are limited to the amount of

$$\hat{L} \leq \frac{\Pi_B}{1 + r_D}. \tag{21.34}$$

If we assume that risky loans are profitable to the bank, they would provide the maximum risky loans possible such that equation (21.34) is fulfilled with equality. The focus of this model will be on the amount of lending to risky companies, $\hat{L}$; as the amount of these loans are determined by the profits of lending to safe companies, we will focus on determining the profits from these safe loans.

The lowest loan rate a bank can offer on the safe investment is such that it breaks even, $\Pi_B^i = \left(1 + r_L^i\right) L - (1 + r_D)\, L - c_i L = 0$, where c_i denotes the monitoring costs. Hence the minimum loan rate is given by

$$1 + r_L^i = 1 + r_D + c_i. \tag{21.35}$$

With d_i denoting the 'distance' of a company from bank i and costs of 'travel' c for the entire length, the profits of the safe company are when choosing bank i are given by

$$\Pi_C^i = (1 + R)\, L - \left(1 + r_L^i\right) L - c d_i L. \tag{21.36}$$

The 'distance' represents the preferences of the company for specific bank services and the further the services the bank offers differs from these preferences, the larger the 'distance'. The costs c denote the strength of these preferences and by how much they reduce the utility of a company from not having these preferences met.

We assume that every company can borrow from at least one bank, including the company furthest away from either bank at a distance of $d_i = \frac{1}{2}$. For the company to make a profit in this situation if two small banks with their higher costs c_S are present in the market, thus that $\Pi_C^S = (1 + R)\, L - \left(1 + r_L^S\right) L - \frac{1}{2} c L \geq 0$. Using equation (21.35) for the minimum loan rate, this requirement solves for

$$c_S \leq \frac{1}{2} c - ((1 + R) - (1 + r_D)). \tag{21.37}$$

We assume that this condition is fulfilled throughout.

To determine the market for each bank, we determine the company that is indifferent between banks i and j, hence we require that the profits of the two companies are identical, $\Pi_C^i = \Pi_C^j$. Noting that $d_i + d_j = 1$, we find that the maximum distance a company has from bank i, would be

$$\hat{d}_i = \frac{1}{2} + \frac{c_j - c_i}{2c} \tag{21.38}$$

and any company located closer to bank i would choose this bank, while anyone located further from the bank would choose the other bank. If the same type of banks operate in the market, eithjer both bank being small or both banks being large, then $c_i = c_j$ and from equation (21.38) we see that $\hat{d} = \frac{1}{2}$.

Let us assume the small bank is never squeezed out of the market by the large bank, implying that $\hat{d}_L < 1$. This will require

$$c_S - c_L < c. \tag{21.39}$$

Using these assumptions, we can now evaluate how competition between two such banks will affect the market outcome.

Market competition Let us first assume that companies always can choose between two banks, even if located furthest from a bank, $d_i = 1$, taking a loan from this bank would be profitable. Thus we need $\Pi_C^i = (1+R)L - \left(1 + r_L^1\right)L - cL \geq 0$, or

$$c = \xi_i, \tag{21.40}$$

where we define $\xi_i = (1+R) - (1 + r_D + c_i)$ for convenience. We have inserted the lowest possible loan rate here as that would be the loan rate used to attract the marginal company. Similarly, the bank will charge each company a loan rate such that the company does not switch to the other bank. Hence it will set its loan rate such that

$$\begin{aligned}\Pi_C^i &= (1+R)L - \left(1 + r_L^i\right)L - cd_iL = (1+R_L)L - \left(1 + r_D + c_j\right)L - cd_jL \\ &= \Pi_C^j.\end{aligned}$$

Using $d_i + d_j = 1$, this solves for

$$1 + r_L^i = \left(1 + r_D + c_j\right) - c\left(2d_i - 1\right). \tag{21.41}$$

As we assumed that banks can engage in discriminatory pricing, banks can set loans rates dependent on the characteristics of the company, which is here captures by its distance. The loan rates will then set such that the company will not prefer to switch to the other bank.

Using the market size as defined in equation (21.38), we get the expected profits of bank i as

$$\Pi_B^i = \int_0^{\hat{d}_i} \left(\left(1 + r_L^i\right)L - \left(1 + r_D + c_i\right)L\right) dd_i = \frac{\left(c + \left(c_j - c_i\right)\right)^2}{4c} L. \tag{21.42}$$

with discriminatory prices as determined by equation (21.41). Banks gain all loans of companies from their location at 0 to $\hat{d}$, with their competitor providing loans to those companies located further than $\hat{d}$ away.

If $c > \xi_i$, then the bank cannot offer a company at distance $d_i = 1$ a loan they would accept as they would make a loss even at the loan rate of $1 + r_L^i = 1 + r_D + c_i$. This means the other bank can expect full monopoly rents from the company by setting the loan rate such that

$$\Pi_C^i = (1+R)L - \left(1 + \hat{r}_L^i\right)L - cd_iL = 0,$$

which solves for the loan rate to be

$$1 + \hat{r}_L^i = (1 + R) - cd_i. \tag{21.43}$$

We see that the loan rate reduces at half the rate compared to the presence of competition in equation (21.41. Such a monopoly lasts until the other bank can make a more attractive loan offer when using their marginal loan rate $1 + r_L^j = 1 + r_D + c_j$. With the other bank located at a distance $1 - d_j$, we need

$$\Pi_C^j = (1 + R)\, L - \left(1 + r_L^j\right) L - c\left(1 - d_j\right) = 0$$

for the maximum distance of such a monopoly. Here the other bank offers a loan rate that gives the company the same profits as its current bank, making the company indifferent between them. This equation solves for a distance to the current bank of

$$\widetilde{d_i} = 1 - \frac{(1 + R) - \left(1 + r_D + c_j\right)}{c} = \frac{\xi_j}{c}. \tag{21.44}$$

Hence bank profits are given by the monopoly rent from its position to $\widetilde{d_i}$ and then the competitive rate from $\widetilde{d_i}$ to $\hat{d}_i$. Using equations (21.38), (21.41), (21.43), and (21.44), we obtain the bank profits as

$$\begin{aligned} \hat{\Pi}_B^i = & \int_0^{\widetilde{d_i}} \left(\left(1 + \hat{r}_L^i\right) L - \left(1 + r_D + c_i\right) L\right) dd_i \qquad (21.45)\\ & + \int_{\widetilde{d_i}}^{\hat{d}_i} \left(\left(1 + r_L^i\right) L - \left(1 + r_D + c_i\right) L\right) dd_i \\ = & \left(\xi_j - \frac{\xi_j^2}{2c} - \frac{c}{4} + \frac{c_j - c_i}{2} + \frac{\left(c_j - c_i\right)^2}{4c}\right) L. \end{aligned}$$

Figure 21.2 illustrate the price setting behaviour of two banks, one being a large bank and the other being a small bank. Until $\widetilde{d_i}$ bank i enjoys a monopoly and charges loan rates accordingly. From then on, the bank faces competition by the other bank and the loan rate will fall faster as companies become more distant. When the loan rate crosses the marginal loan rate of the other bank, $1 + r_D + c_S$, the loan rate continues to fall and the bank undercuts its competitor, as the lower loan rate allows to attract more companies, while still being profitable. This continues until at $\hat{d}_i$, it would no longer be profitable for the bank to reduce the loan rate further. At this point, companies further afield will use the other bank. We assume here that the other bank does not enjoy an area of monopoly power and thus has to price competitively throughout. The consequence is that the other bank cannot fully exploit their market power and for banks of the same distance to it than to the other bank, has to charge a lower loan rate.

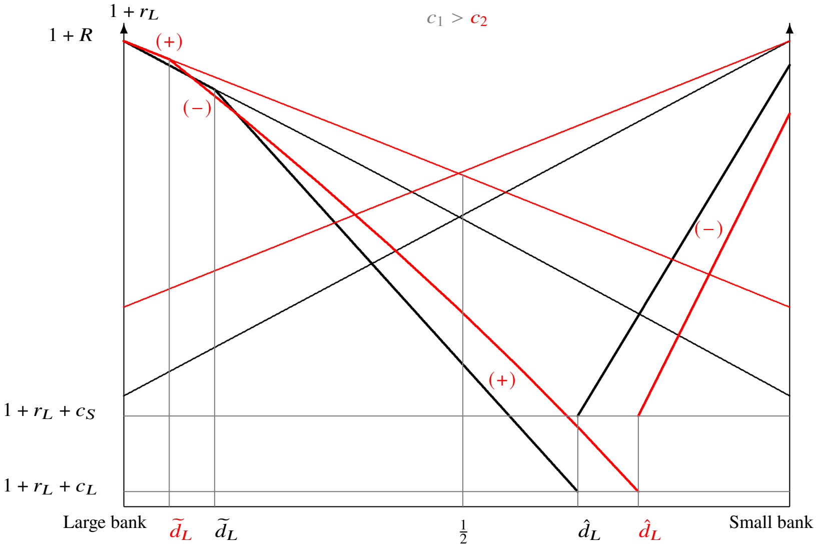


Fig. 21.2: Price setting by competing banks of different sizes

Expected profits Let us first (case A) assume that small banks are able to cover the whole market, this requires the 'travel' costs of companies to be sufficiently low such that $c < \xi_S$. This implies that no bank has monopoly over companies located close to them. Regardless of the types of banks competing, loan rates are always set competitively.

If we have a market where two identical banks are competing, either two large or two small banks, we get from equation (21.42) with Π^{ij} denoting the profits of bank i if competing with bank j, that

$$\Pi^{SS} = \Pi^{LL} = \frac{c}{4}. \tag{21.46}$$

If one bank is large while the other bank is small, we get the profits of the small and large banks, respectively, as

$$\begin{aligned} \Pi^{SL} &= \frac{(c + (c_L - c_S))^2}{4c}, \\ \Pi^{LS} &= \frac{(c + (c_S - c_L))^2}{4c}. \end{aligned} \tag{21.47}$$

With λ denoting the fraction of large banks and banks being paired randomly, the total profits of banks in this case are then given by

$$\begin{aligned}\Pi^A &= \lambda^2 \Pi^{LL} + (1-\lambda)^2 \Pi^{SS} + \lambda(1-\lambda)\Pi^{SL} + \lambda(1-\lambda)\Pi^{LS} \\ &= \frac{c}{4} + \lambda(1-\lambda)\frac{(c_S - c_L)^2}{2c},\end{aligned} \tag{21.48}$$

when inserting from equations (21.46) and (21.47).

The second case (case B) is the situation where a small bank cannot cover the entire market, thus $c > \xi_S$, and the other bank has monopoly power over companies close to them. In addition, we assume that a large bank will cover the entire market, hence $c < \xi_L$, and therefore any bank matched with a large bank will face competition for all companies they can reach. In summary we assume that $\xi_S < c < \xi_L$.

With two small banks, they would both enjoy monopoly powers and from equation (21.45) we have the profits of these banks given by

$$\Pi^{SS} = \xi_S - \frac{\xi_S^2}{2c} - \frac{c}{4}, \tag{21.49}$$

using that $c_i = c_j = c_S$. Similarly, two large banks do not enjoy any monopoly power and we have from equation (21.42) and $c_i = c_j = c_L$ that bank profits are given by

$$\Pi^{LL} = \frac{c}{4}. \tag{21.50}$$

If one bank is large while the other bank is small, the large bank enjoys a monopoly over companies located close to them, while the small bank does not enjoy any monopoly power. Hence

$$\begin{aligned}\Pi^{LS} &= \xi_S - \frac{\xi_S^2}{2c} - \frac{c}{4} + \frac{c_S - c_L}{2} + \frac{(c_S - c_L)^2}{4c}, \\ \Pi^{SL} &= \frac{(c + (c_L - c_S))^2}{4c} = \frac{c}{4} + \frac{(c_S - c_L)^2}{4c} - \frac{c_S - C_L}{2}.\end{aligned} \tag{21.51}$$

Thus, similar to equation (21.48), the total profits in this case are given by

$$\Pi^B = (2\lambda - 1)\frac{c}{4} + (1-\lambda)\left(\xi_S - \frac{\xi_S^2}{2c}\right) + \lambda(1-\lambda)\frac{(c_S - c_L)^2}{2c}. \tag{21.52}$$

The final case (case C) is that small and large banks enjoy monopoly power as neither can cover the entire market, thus we consider the case that $c > \xi_L$. In this case we get the bank profits as

$$\begin{aligned}
\Pi^{LL} &= \xi_S - \frac{\xi_S^2}{2c} - \frac{c}{4}, \\
\Pi^{SS} &= \xi_L - \frac{\xi_L^2}{2c} - \frac{c}{4}, \\
\Pi^{SL} &= \xi_L - \frac{\xi_L^2}{2c} - \frac{c}{4} + \frac{c_L - c_S}{2} + \frac{(c_L - c_S)^2}{4c}, \\
\Pi^{LS} &= \xi_S - \frac{\xi_S^2}{2c} - \frac{c}{4} + \frac{c_S - c_L}{2} + \frac{(c_S - c_L)^2}{4c}
\end{aligned} \tag{21.53}$$

and the total profits are given by

$$\begin{aligned}
\Pi^C = \lambda \left(\xi_S - \frac{\xi_S^2}{2c} \right) + (1 - \lambda) \left(\xi_L - \frac{xi_L^2}{2c} \right) \\
-\frac{c}{4} + \lambda (1 - \lambda) \frac{(c_S - c_L)^2}{2c}.
\end{aligned} \tag{21.54}$$

Noting that $\hat{L} = \frac{\Pi_B}{1+r_D}$ from above, we can now analyze the properties of these profits with the aim to evaluate the implication of competition on the provision of risky loans. If the profits of the bank from granting safe loans are increased, banks can increase the supply of risky loans. It is therefore that banks profits as determined here and the provision of risky loans are equivalent.

The impact of competition To start, we asses a market in which small and large banks face equal costs, i.e. $c_L = c_S$, and hence large banks have no comparative advantage over small banks. In this situation, case B requiring $\xi_S < c < \xi_L$ cannot occur as $\xi_S = \xi_L$ from the definition of ξ_i, thus case B does not need to be evaluated.

An increase in competition we interpret here as a reduction in c, the cost of companies to take loans from banks, depending on their location. This is justified as a reduction of c will reduce the monopoly power, if it exists, as we see from equation (21.44). Reduced costs to companies also implies that it is less costly for them to switch banks and any differences in loan rates become more and more important until, at $c = 0$, the loan rate is the only determinant for the choice of bank.

Looking at equations (21.48) and (21.54), we obtain

$$\begin{aligned}
\frac{\partial \Pi^A}{\partial c} &= \frac{1}{4} > 0, \\
\frac{\partial \Pi^C}{\partial c} &= \frac{\lambda \xi_S^2 + (1 - \lambda) \xi_L^2}{2c^2} - \frac{1}{4},
\end{aligned} \tag{21.55}$$

where the latter is positive if

$$c < \sqrt{2 \left(\lambda \xi_S^2 + (1 - \lambda) \xi_L^2 \right)}. \tag{21.56}$$

Thus for sufficiently small costs $c < \xi_s$, an increase in c increases bank profits and hence the provision of risky loans. This implies that if competition is sufficiently high (low costs c), then increasing competition even more (lowering c), reduces the provision of risky loans. In less competitive markets, increasing competition would increase the provision of risky loans.

Looking at figure 21.2 allows us to understand this result. Lower costs c, thus increased competition, reduces the monopoly power of the bank as seen there. What remains of the monopoly can be exploited more given the higher surplus of companies that can be extracted, indicated by a plus sign in the figure. The smaller monopoly market, though, reduces profits, indicated by a minus sign. Finally, the lower costs extends the market of the larger bank, increasing profits, indicated by a plus sign in the lower area. This in turn reduces the profits of the smaller bank, as shown by a minus sign on the right. As we can see in the figure here, the profits will on balance reduce, leading to less risky loans being given.

If the costs are higher, though, such that both banks have a monopoly over some companies, we can see by focussing on large bank that the profit changes are much more balanced and as costs decrease, this will lead to an increase in the market share of this bank, outweighing the reduction of profits in the monopoly area.

For the more general case that banks have different costs, we assume that $c_S - c_L > \xi_S\sqrt{2}$, such that the cost advantage of large banks in monitoring are sufficiently large. In this case we rule out case A which required that $c < \xi_S < \frac{c_S - c_L}{\sqrt{2}} < \frac{c}{\sqrt{2}}$ using equation (21.39). Similarly, case C can be ruled out as we would need $c > \xi_L = \xi_S + (c_S - c_L) > \left(1 + \sqrt{2}\right)\xi_S$, where the relationship of ξ_S and ξ_L emerge from their definitions. If costs c were this large, small banks would not be able to cover the full market as $1 + \sqrt{2} > 2$, contradicting our assumption that all companies can obtain a loan. This leaves us with only case B to evaluate.

We find that the influence of the costs on bank profits are given by

$$\frac{\partial \Pi^B}{\partial c} = \frac{2\lambda - 1}{4} - (1 - \lambda)\left(\xi_S^2 - \lambda\left(c_S - c_L\right)^2\right). \tag{21.57}$$

If $\frac{\partial \Pi^B}{\partial c} > 0$ then again an increase in competition would result in fewer risky loans. We can solve equation (21.57) and obtain that this is the case if

$$\left|\lambda + \frac{c^2 - (c_S - c_L)^2 - \xi_S^2}{2\left(c_S - c_L\right)^2}\right| > \sqrt{\left(\frac{c^2 - (c_S - c_L)^2 - \xi_S^2}{2\left(c_S - c_L\right)^2}\right)^2 - \frac{2\xi_S^2 - c^2}{2\left(c_S - c_L\right)^2}},$$

of which only the upper root can be shown to yield a solution within the relevant interval of $\lambda \in [0; 1]$. We can also show that the solution is fulfilled if $\lambda > \frac{1}{2}$.

Thus if many large banks are present, increasing competition between them reduces their profits as neither have monopoly powers over some companies to fall back on. Only small banks allow large banks to create monopolies they can exploit. Hence if they are sufficiently active in the market, profits can increase with increasing competition (lower c) due to the better exploitation of the home market.

This is consistent with the impact of more large banks on bank profits and hence the provision of risky loans,

$$\frac{\partial \Pi^B}{\partial \lambda} = \frac{c}{2} - \left(\xi_S - \frac{\xi_S^2}{2c}\right) + (1-2\lambda)\frac{(c_S - c_L)^2}{2c}, \tag{21.58}$$

being positive. This is the case if

$$\lambda < \frac{1}{2} + \frac{(c-\xi_X)^2}{2(c_S - c_L)^2} < 1, \tag{21.59}$$

with the last inequality arising from $c - \xi_S < c_S - c_L$, which solves for $c < \xi_L$, our assumption here. Hence only if there are sufficiently many, but not too many large banks in the market, does increasing their presence increase the provision of risky loans. The reduction of the monopoly for small banks, increases profits to companies that can be extracted by banks, increasing their profits, which can then be used to provide risky loans. Once there are many large banks, the competition between them, as they are frequently competing with each other, reduces profits, though.

Summary We have seen that competition between banks can actually reduce the provision of risky loans. This is the case if competition between banks is low and many large banks dominate the market. Hence increasing competition does not always lead to better access to risky loans. We might not interpret λ strictly as the number of large banks, but as the likelihood that a large bank is considered by a company as one of the two potential lenders. In this sense it might be interpreted as the market share of a bank, here a bank that has a cost advantage over other banks.

Banks subsidize risky loans through the provision of risk-free loans for which they compete in monopolistic competition; the profits generated from this monopolistic are then used to cover the potential losses from the provision of risky loans. In many cases, increased competition for such risk-free loans will reduce the profits they generate and thus limit the provision of risky loans.

Reading Heddergott & Laitenberger (2017)

Conclusions

We have seen that in monopolistic markets loan and deposit rates are set independently as they reflect different markets. If banks require customers to use their services for borrowing and depositing, the profits of banks remain unchanged and even regulatory constraints on one interest rate, for example the loan rate, will not affect these profits, but instead banks will reduce the deposit rate, or if the deposit rate is restricted, increased the loan rate, to maintain their profitability. The profits that banks generate from their local monopoly power due to the preferences of some individuals for their specific services allows them to generate profits in otherwise homogeneous markets. However, if the competition between banks is increasing due the preferences of customers for specific services becoming less important to customers, these profits might decrease; but a heterogeneity in the costs banks face when providing loans, such as competitive advantages for monitoring loans by large banks, might in some instances increase the overall profits in the banking system as in particular large banks might face an even stronger competitive advantage. This would allow, for example banks to use such profits to generate more risky loans, where their profits allow them to cover potential losses from these, making it viable to grant such risky loans. Thus increased competition might have an adverse effect on the ability of banks overall to provide risky loans.

Chapter 22

Competition through loan size

It is common to assume that banks compete for customers through their loan and deposit rates, or other fees for account services. This is not the only way banks can compete with each other, they might also compete for companies seeking loans by offering them larger loans than competitors, making them less reliant on other funding sources, such as equity or the bond market. We will here explore the competition between banks seeking to attract borrowers by offering them larger loans than their competitors and explore the properties of the resulting equilibrium.

Let us assume that a company can choose between two investment projects. The first investment project succeeds with probability π_H and generates a return of R_H on the initial investment I. The alternative investment project succeeds with probability $\pi_L < \pi_H$ and hence is more risky but returns $R_L > R_H$. We furthermore assume that $\pi_H (1 + R_H) > 1 + r > \pi_L (1 + R_L)$, where r denotes the risk-free rate. This assumption implies that the low-risk investment is socially desirable as it exceeds the costs of funding from a risk-free fund, while the high risk investment is socially desirable, given that its return is below this threshold. As the bank financing the investment cannot profitably charge a loan rate below the risk-free rate, it would want the company to choose the low risk investment. It might, however, be more profitable for the company to choose the high risk investment, thus a moral hazard problem emerges.

Banks counter this moral hazard problem by monitoring the company such that it ensures the safe investment is chosen. Banks face monitoring costs C, which will depend on the characteristics of the company, such as the industry they operate in, the complexity of the business model, or similar aspects. We model these characteristics by assuming that companies have 'distance' to the bank, denoted d, that captures this ability of the bank to monitor companies. We assume that for the total monitoring costs we have $C = cd$, where c is the monitoring cost of a company one unit away. As a normalization we assume $0 \leq d \leq 1$. Thus a company with a larger 'distance' from the bank is more difficult to monitor and the bank faces higher costs of doing so.

Loan demand Companies finance their investments I using bank loans L and they have access to the bond market, borrowing $\hat{L}$. The bond market has no ability to monitor the company. The total investment financed by debt, a bank loan or a bond, and the company has no own funds, hence

$$L + \hat{L} = I. \tag{22.1}$$

As we will want to make sure that companies only choose the socially desirable safe investment, the return to an investor in the bond market, yielding a loan rate $\hat{r}_L$, must at least give the risk-free return r to be attractive, thus we require

$$\pi_H (1 + \hat{r}_L) \hat{L} \geq (1 + r) \hat{L}. \tag{22.2}$$

Similarly, banks pay depositors a rate of r_D and have to recover their monitoring costs as well, with loan rate r_L on the bank loan we thus require

$$\pi_H (1 + r_L) L - C \geq (1 + r_D) L, \tag{22.3}$$

assuming that banks finance their loans fully by deposits.

In order to monitor the company, banks need to make higher returns if companies choose the safe investment over the risky investment, hence

$$\pi_H (1 + r_L) L - C \geq \pi_L (1 + r_L) L. \tag{22.4}$$

The revenue generated by the company is distributed to the bank and bond market to repay the loans with interest, with the company retaining profits Π_C. Thus in the case the investment is successful, the company makes profits of

$$\Pi_C = (1 + R_H) I - (1 + r_L) L - (1 + \hat{r}_L) \hat{L}. \tag{22.5}$$

Finally, the company needs to be better off choosing the safe investments over the risky investment, thus we require

$$\begin{aligned} &\pi_H \left((1 + R_H) I - (1 + r_L) L - (1 + \hat{r}_L) \hat{L}\right) \\ &\quad \geq \pi_L \left((1 + R_L) I - (1 + r_L) L - (1 + \hat{r}) \hat{L}\right). \end{aligned} \tag{22.6}$$

Using these constraints, we can get from condition (22.4) that

$$1 + r_L \geq \frac{C}{(\pi_H - \pi_L) L}, \tag{22.7}$$

which inserted into condition (22.3) solves for

$$L \leq \frac{\pi_L}{\pi_H - \pi_L} \frac{C}{1 + r_D}. \tag{22.8}$$

and hence combining with condition (22.7) we obtain the minimum loan rate banks require as

$$1 + r_L \geq \frac{1 + r_D}{\pi_L}. \tag{22.9}$$

Finally, when combining condition (22.8) with condition (22.2), we obtain the minimum loan rate in bonds markets as

$$1 + \hat{r}_L \geq \frac{1 + r_D}{\pi_H}, \tag{22.10}$$

which implies that the minimum loan rate required by the bank is higher than that of the bond market as $\pi_L < \pi_H$. The reason for this requirement is that the bank needs to recover their costs of monitoring.

With equation (22.1) we obtain using condition (22.8) that the lowest possible amount of bonds used to finance the investment of the company is given by

$$\hat{L} \geq I - \frac{\pi_L}{\pi_H - \pi_L} \frac{C}{1 + r_D}. \tag{22.11}$$

Inserting all this information into equation (22.5), we get the maximum profits of the company as

$$\Pi_C \leq (1 + R_H) I - \frac{1 + r}{\pi_H} I - \frac{C}{\pi_H - \pi_L} \left(1 - \frac{1 + r}{1 + r_D} \frac{\pi_L}{\pi_H} \right). \tag{22.12}$$

As our focus will not be on the competition for companies on the basis of loan rates, but instead we focus on the monitoring of the company by banks, let us assume that the minimum loan rates are charged and the bank provides the maximum amount of loans possible. Thus we assume that the loan market is perfectly competitive in terms of the loan rates that are charged; this implies that relationships (22.8)-(22.12) hold with equality.

In order for condition (22.6) that companies choose the low-risk investment to be fulfilled, we obtain, after inserting for all variables that we require, that the monitoring costs have to be sufficiently small, namely

$$C \leq \frac{\pi_H (1 + R_H) - \pi_L (1 + R_L) - \frac{\pi_H - \pi_L}{\pi_H} (1 + r)}{1 - \frac{1+r}{1+r_D} \frac{\pi_L}{\pi_H}} I. \tag{22.13}$$

If monitoring costs are sufficiently small, companies are able to obtain a loan and find it profitable to do so; note here that the bond market free-rides on the monitoring effort of the bank as they make no contribution to the monitoring costs of the bank but benefit from the company choosing the low-risk investment.

We assume that the 'distance' d of companies from the bank, thus the difficulties in monitoring companies, are uniformly distributed in the interval $[0; 1]$. This allows us to calculate the total demand for bank loans, using $C = cd$. The total demand for loans will be given equal to the amount of deposits and hence we require

$$D = \int_0^{\overline{d}} L dd = \frac{\pi_L}{\pi_H - \pi_L} \frac{c}{1 + r_D} \int_0^{\overline{d}} d dd = \frac{1}{2} \frac{\pi_L}{\pi_H - \pi_L} \frac{c\overline{d}^2}{1 + r_D}. \tag{22.14}$$

The total demand for loans is given for all companies as determined in equation (22.8). Given that the amount of deposits is exogenously given, the bank may not be able to serve the entire market, but only up to a distance of $\overline{d}$; banks would choose those companies that are closest to them as they require the lowest monitoring costs to banks and thus provide them with the highest profits.

We can solve equation (22.14) for the deposit rate, which becomes

$$1 + r_D = \frac{1}{2} \frac{\pi_L}{\pi_H - \pi_L} \frac{c\overline{d}^2}{D}, \tag{22.15}$$

which is decreasing in the amount of deposits the bank can attract; thus smaller banks would be able to pay higher deposit rates than larger banks. This is because smaller banks only give fewer loans that are located closer to them, which reduces monitoring costs and thus allows them to pay higher deposit rates without reducing their profits. Alternatively smaller banks could provide lower loan rates, but we assumed these to be fixed in a competitive market. We are here not concerned about how banks can maintain different deposit rates, such as through some form of market segmentation, but instead focus on the monitoring by banks.

Inserting $C = c\overline{d}$ and equation (22.15) into equation (22.13), we obtain after some transformations that

$$\begin{aligned} c\overline{d}^2 - &\left(\pi_H (1 + R_H) - \pi_L (1 + R_L) - \frac{\pi_H - \pi_L}{\pi_H} (1 + r)\right) \overline{d} \\ &- 2\frac{\pi_H - \pi_L}{\pi_H} D (1 + r) = 0, \end{aligned}$$

which solves for

$$\overline{d} = \frac{\xi + \sqrt{\xi^2 + 8\frac{\pi_H - \pi_L}{\pi_H} cD (1 + r)}}{2c}, \tag{22.16}$$

with $\xi = \pi_H (1 + R_H) - \pi_L (1 + r_L) - \frac{\pi_H - \pi_L}{\pi_H} (1 + r)$. We assumed here that equation (22.13) holds with equality due to the competitive loan market which we assumed. We thus see that a bank might not be able to provide loans to all companies and those that are too 'distant' would impose too high monitoring costs for banks to be able to complete with the bond market, despite the knowledge of the bond market that companies would pursue the high-risk investment.

The lowest possible deposit rate would be $1 + r_D = 1 + r$ as otherwise the bank would demand infinite deposits and invest them into the risk-free asset. If equation (22.15) would imply that $1 + r_D < 1 + r$, the bank has excessive deposits that cannot all be lent out. This is the case if

$$D \geq \overline{\overline{D}} = \frac{1}{2} \frac{\pi_L}{\pi_H - \pi_L} \frac{c\overline{\overline{d}}^2}{1 + r} \tag{22.17}$$

as we obtain by inserting this relationship into equation (22.15). We can derive the maximum 'distance' of a company from the bank that ensures that the deposit rate the bank can pay does exceed the risk-free rate, $\overline{\overline{d}}$, by inserting $1 + r_D = 1 + r$ and $C = c\overline{\overline{d}}$ into equation (22.13). We then obtain

$$\overline{\overline{d}} = \frac{\xi I}{c\frac{\pi_H - \pi_L}{\pi_H}}. \tag{22.18}$$

Throughout we will assume that the deposits banks attract fulfill $D \leq \overline{\overline{D}}$ and thus banks do not have excess deposits they are not able to use for loans and that they will cover the entire market allowing every company to obtain a loan, thus $\overline{d} \geq 1$ if $D = \overline{\overline{D}}$. Hence from equation (22.18) with $\overline{\overline{d}} > 1$, we need that

$$c < \frac{\xi I}{1 - \frac{\pi_L}{\pi_H}}. \tag{22.19}$$

Was this condition not fulfilled, banks could specialise and locate themselves such that for one company the distance to one bank is 1, while a distance to another bank is 0; banks could then act as monopolists to companies located at $d = 0$, which corresponds to $d = 1$ for the other bank, as the other bank would not offer a loan to this company. By requiring that each bank is able to offer a loan to each company we ensure that banks compete for all companies.

Competition between banks Let us now consider two such banks competing to provide a loan to a company. This company is located at a distance d from bank 1 and with our assumption that the maximal distance between companies is 1, a distance of $1 - d$ to bank 2. The size of bank loans the two companies could obtain are given from equation (22.8) as

$$\begin{aligned} L_1 &= \frac{\pi_L}{\pi_H - \pi_L} \frac{cd}{1 + r_D^1}, \\ L_2 &= \frac{\pi_L}{\pi_H - \pi_L} \frac{c(1-d)}{1 + r_D^2}, \end{aligned} \tag{22.20}$$

where we used again that $C = cd$ ($C = c(1-d)$) for obtaining the bank loan from bank 1 (2), which offers a deposit rate of $1 + r_D^1$ ($1 + r_D^2$). The profits of companies are given by

$$\Pi_C^i = \pi_H \left((1 + R_H) I - (1 + r_L) L_i - (1 + \hat{r}_L) \hat{L}_i\right) \tag{22.21}$$

if the company borrows from bank i. Inserting from the above we can now determine the location of the company being indifferent between choosing between bank 1 and bank 2 when solving the equation $\Pi_C^1 = \Pi_C^2$ for the distance d and denoting the solution by $\hat{d}$.

With bank 1 providing loans to companies located between 0 and $\hat{d}$ and bank 2 to those located between $\hat{d}$ and 1, we get the loan demands for each bank and can equate this to the deposits the banks have obtained. Thus we find

$$\begin{aligned} D_1 &= \int_0^{\hat{d}} L_1 dd = \frac{1}{2}\frac{\pi_L}{\pi_H - \pi_L}\frac{c\hat{d}^2}{1+r_D^1}, \\ D_2 &= \int_{\hat{d}}^1 L_2 dd = \frac{1}{2}\frac{\pi_L}{\pi_H - \pi_L}\frac{c\left(1-\hat{d}\right)^2}{1+r_D^2}, \end{aligned} \tag{22.22}$$

which solve for the deposit rates offered to be

$$\begin{aligned} 1 + r_D^1 &= \frac{1}{2}\frac{\pi_L}{\pi_H - \pi_L}\frac{c\hat{d}^2}{D_1} \\ 1 + r_D^2 &= \frac{1}{2}\frac{\pi_L}{\pi_H - \pi_L}\frac{c\left(1-\hat{d}\right)^2}{D_2} \end{aligned} \tag{22.23}$$

Inserting these deposit rates back into equation (22.22) and then into equation (22.21), we can write the condition that banks make the same profits, $\Pi_C^1 = \Pi_C^2$, as

$$\mathcal{L} = \frac{D_1}{\hat{d}^2} - \frac{D_2}{\left(1-\hat{d}\right)^2} - \frac{1}{2}\frac{c^2}{1+r}\frac{\pi_H}{\left(\pi_H - \pi_L\right)^2}\left(2\hat{d} - 1\right) = 0. \tag{22.24}$$

We require the profits of the two banks to be equal as otherwise the 'positions' of the banks would not be an equilibrium and the bank making smaller profits would prefer to change their 'position', or their monitoring expertise, as to increase their profits at the expense of the bank making higher profits. In the case that banks are of the same size, $D_1 = D_2$, the solution to this equation is $\hat{d} = \frac{1}{2}$ and each bank provides loans to half the companies in the market.

Banks of different sizes We can now investigate how the market share of each bank changes as banks become different in size. Let us assume that bank 2 is the larger bank and set $D_2 = kD_1$, with $k \geq 1$. We then easily obtain

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \hat{d}} &= -\frac{2D_1}{\hat{d}^3} - \frac{2kD_1}{\left(1-\hat{d}\right)^3} - \frac{c^2\pi_H}{(1+r)\left(\pi_H - \pi_L\right)^2} < 0, \\ \frac{\partial \mathcal{L}}{\partial k} &= -\frac{D_1}{\left(1-\hat{d}\right)^2} - \frac{kD_1}{\left(1-\hat{d}\right)^3}\frac{\partial \hat{d}}{\partial k}. \end{aligned} \tag{22.25}$$

Totally differentiating the condition for banks having equal profits from equation (22.24) we obtain

$$d\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \hat{d}} d\hat{d} + \frac{\partial \mathcal{L}}{\partial k} dk = 0$$

and hence

$$\frac{d\hat{d}}{dk} = -\frac{\frac{\partial \mathcal{L}}{\partial k}}{\frac{\partial \mathcal{L}}{\partial \hat{d}}}. \tag{22.26}$$

Inserting from equation (22.25), we obtain

$$\frac{d\hat{d}}{dk} = \frac{\frac{D_1}{\left(1-\hat{d}\right)^2}}{-\frac{2D_1}{\hat{d}^2} - \frac{3kD_1}{\left(1-\hat{d}\right)^3} - \frac{c^2 \pi_H}{(1+r)(\pi_H - \pi_L)}} < 0. \tag{22.27}$$

Hence, if bank 2 increases its size relative to bank 1, an increase in k, it will increase its market share, $1 - \hat{d}$ and attract additional companies from the smaller bank. The larger amount of deposits makes it desirable for the bank to expand their market through the provision of larger loans to companies. These larger bank loans are attractive to companies and hence they prefer a loan from a larger bank as that allows them to reduce their reliance on the bond market.

We also have

$$\frac{\partial \mathcal{L}}{\partial D_1} = \frac{1}{\hat{d}^2} - \frac{k}{\left(1-\hat{d}\right)^2} = (1-k)\,\frac{\left(\hat{d} - \frac{1}{1-k}\right)^2 + \frac{k}{(1-k)^2}}{\hat{d}^2\left(1-\hat{d}\right)^2} < 0, \tag{22.28}$$

where we note that $k \geq 1$. Similarly, after totally differentiating equation (22.24) and noting the signs in equations(22.25) and (22.28), we find that

$$\frac{d\hat{d}}{dD_1} = -\frac{\frac{\partial \mathcal{L}}{\partial D_1}}{\frac{\partial \mathcal{L}}{\partial \hat{d}}} < 0. \tag{22.29}$$

This result implies that even if the size of both banks increases through an increase of deposits they obtain, the larger bank will expand its market share. As the large bank has a larger increase in size as we keep the ratio of the sizes, k, constant, it will more aggressively increase the loan sizes and capture additional companies who find the larger size of banks loans attractive.

The maximum deposit rates the banks can pay are also affected by a change in the size of deposits. If the larger bank increases its size further, the smaller bank is able to support a higher deposit rate, given its lower monitoring costs due to loans being provided at a smaller distance than the larger bank. This difference is increasing as the larger bank attracts more and more customers, who are ever more distant, and thus incurring higher monitoring costs. Noting that the ratio of deposit rates is given by

$$\mathcal{L} = \frac{1+r_D^1}{1+r_D^2} = k\frac{\hat{d}^2}{\left(1-\hat{d}\right)^2},$$

we have

$$\frac{\partial \mathcal{L}}{\partial k} = \frac{\hat{d}}{\left(1-\hat{d}\right)^3}\left(2k\left(1-2\hat{d}\right)\frac{\partial \hat{d}}{\partial k} + \hat{d}\left(1-\hat{d}\right)\right) > 0, \tag{22.30}$$

as $\frac{\partial \hat{d}}{\partial k} < 0$ from equation (22.27) and $\hat{d} \geq \frac{1}{2}$. Hence the positive sign indicates that as the large bank becomes larger, the smaller bank offers a relatively better deposit rate than the larger bank.

If we increase the size of both banks, we obtain

$$\frac{\partial \mathcal{L}}{\partial D_1} = 2k\frac{\hat{d}}{\left(1-\hat{d}\right)^3}\frac{\partial \hat{d}}{\partial D_1} < 0, \tag{22.31}$$

and the differences in the deposit rates are reducing. As both banks become larger, they becomes less constrained on lending and will increase the size of their loans, but this also means that they will become more alike and hence slowly both deposit rates converge towards the risk free rate r.

Competition between banks manifests itself here through the provision of loans and an incursion into the area of expertise of the other bank to monitor companies; loan rates are taken as given. If we assume that deposit rates cannot vary, because depositors are shifting to the most profitable bank, then the common deposit rate would be below the deposit rate of the larger bank, $1 + r_D^2$, as both banks need to be profitable to operate. If the smaller bank offers a higher deposit rate than the larger bank, it would attract all deposits and hence grow, making this high deposit rate unsustainable as the deposits need to be invested into loans, which in turn would increase the monitoring costs of that banks, necessitating the deposit rate to reduce.

An implication of this interpretation is that small banks are more profitable than large banks. It has to be noted, though, that the sizes of the banks are exogenously determined in this model and not an equilibrium; if banks compete for deposits their deposit rates would have to be equal and this would determine their relative size. In the current model, small banks will specialize in a small area of expertise and large banks cover a wider range of the market.

Summary We have seen how banks compete for loans not by offering lower loans rates but by offering companies larger loans that reduce the reliance on bond markets. The monitoring of banks ensures that companies seek the low-risk investment and this in turn allows banks to charge lower loan rate, making this investment attractive to them. Offering loans at better conditions than the bond market could, who has no ability to monitor companies and will thus charge a higher loan rate. The attraction for a company to switch from one bank to another is that this bank would be able to offer a larger loan than the competitor and thus allow the company to reduce their reliance on the bond market.

Reading Almazan (2002)

Chapter 23

Shadow banking

Simplified, banks collect funds from the general public (deposits), that can be withdrawn at any time. These funds (deposits) are then used to provide long-term loans to borrowers. Some money market funds and hedge funds provide similar arrangements indirectly. They collect funds from wealthy individual investors as well as institutional investors, which often can be withdrawn at short notice, and use these funds to purchase assets; often such assets are securitized loans. These securitized loans have originated in a bank and have then been sold to a special purpose vehicle, which in turn issues the securities money market funds buy. Thus indirectly these money market funds are providing loans to borrowers as the banks can use the proceeds of the loan sale to grant more additional loans. These money market funds are not subject to the same regulations as banks, especially with respect to capital and liquidity requirements. As they fulfill similar roles as banks they are commonly referred to as ’shadow banks’.

An important aspect of shadow banking is that their investors seek to invest in safe assets, i.e. shadow banks will typically demand securitized loans that have no or little risks. This then allows them to provide the safety to their investors, similar to deposits. Figure 23.1 illustrates how these shadow banks typically operate when investing into securitized loans.

While in principle such shadow banks and actual banks look similar from the perspective of depositors, there are distinct differences. Firstly, money market funds (shadow banks) are not or much less regulated than banks. For example, there are no strict capital or liquidity requirements imposed on shadow banks, and the types of assets shadow bank can invest in is generally not restricted. Thus the safeguards for depositors with shadow banks are much less developed and depend on the contractual arrangements of the depositor with the shadow bank itself. In addition, a safeguard to the equivalent of deposit insurance for bank deposits are not available to shadow banks.

In order for a shadow banking system to be able to exist, banks must be willing to provide risk-free securitized loans to shadow banks; we will explore the conditions that have to be met in chapter 23.1. The existence of shadow banks is often

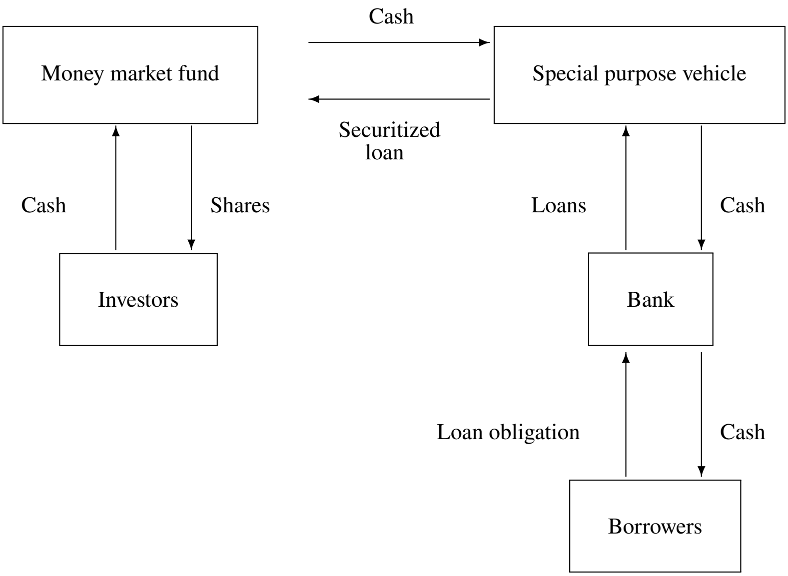


Fig. 23.1: Shadow banking and securitisation

attributed to the lighter regulation they are facing compared to banks and that they take advantage of this. For this reason, chapter 23.2 will look at regulatory arbitrage and the benefits of banks setting up divisions engaged in shadow banking to benefit from these less strict regulatory requirements.

23.1 Provision of risk-free securities by banks

Let us assume that investors seek to invest into risk-free assets only; if the funds to be invested exceed the amount of readily available risk-free assets, investors can turn to risky assets that have been securitized such that investors do not face risks, but any risks are retained by banks generating such risky assets, for example loans. It often are these securitized loans that shadow banks are purchasing with the funds they obtain from their investors and hence banks need to have incentives to provide such risk-free securitized loans.

Banks can give two types of loans, safe and risky loans. Safe loans attract an interest rate r_L^S and are repaid with certainty. There is, however, a limited supply of these loans such that a bank can give no more than $\overline{L}$ of safe loans, i.e. the amount lent, L_S, fulfills $L_S \leq \overline{L}$. The risky loans get repaid in full, including interest, with probability π_i; otherwise the bank does not obtain any payments. The bank does not know the probability π_i of the loan being repaid, only that it is either π_H or π_L, where $\pi_H > \pi_L$. It is known that a fraction p_i of loans will exhibit probability of success π_i. We denote by $\pi = \pi_H p_H + \pi_L p_L$ the average success probability. Risky loans are supposed to be available without constraints and they attract a loan rate of r_L^R, where the amount of such loans given is L_R. We assume that $1 + r_L^S > \pi \left(1 + r_L^R\right)$, making the return on the safe loan higher than the expected return on the risky loan.

Banks can securitize their safe (risky) loans to the amount of S_S (S_R) and obtain a price of P_S (P_R) for their loans. They can also buy securitized loans using the safe (risky) loan to the amount of T_S (T_R) at the same price. As we assume that securitized loans bought from different banks, the resulting diversification eliminates the idiosynchratic default risk in loans as it is assumed defaults are independent across banks. Hence securitized loans have no default risk, even for risky loans, but only certain losses from a fixed default rate.

The net amount of safe and risky loans a bank is exposed to is given by $L_i + T_i - S_i$, i.e. the loans given and securitized loans bought, less those securitized loans sold. Banks also obtain funds from selling their loans for securitization and lose funds when buying securitized loans, $P_i (S_i - T_i)$ and face costs by paying interest r_D on the funds by investors, D. In addition they may retain any unused funds, $L_S + L_R - D$. This gives us the expected profits of the bank as

$$\begin{aligned}\Pi_B = {} & (1 + r_S)(L_S + T_S - S_S) + P_S (S_S - T_S) \\ & + \pi \left(1 + R - L^R\right)(L_R + T_R - S_R) \\ & + P_R (S_R - T_R) - (1 + r_D) D - (L_S + L_R - S_S - S_R).\end{aligned} \tag{23.1}$$

Banks can only lend out and use to purchase securitized loans to the amount they obtained from investors and selling loans for securitization. Thus

$$D - (L_S + L_R + P_S (T_S - S_S) + P_R (T_R - S_R)) \geq 0. \tag{23.2}$$

Investors seek safe investments, hence we cannot allow bank to fail and thus need to ensure that even in the worst case scenario of a bank losing their risky loans and the probability of success being π_L, no loses occur. Hence we need

$$\left(1 + r_L^S\right)(L_S + T_S - S_S) + \pi_L \left(1 + r_L^R\right) T_R - (1 + r_D) D \geq 0, \tag{23.3}$$

where securitized loans held are risk-free in that we know the fraction being repaid.

Further constraints are that banks cannot securitize more than the loans they have given and the limited supply of safe loans :

$$\begin{aligned} L_S - S_S &\geq 0 \\ L_R - S_R &\geq 0 \\ \overline{L} - L_S &\geq 0 \end{aligned} \tag{23.4}$$

Banks will maximize their profits from equation (23.1) subject to these constraints from equations (23.2), (23.3), and (23.4). If we let ξ_i denote the Lagrange multipliers for each constraint, the first order conditions are given by

$$\frac{\partial \Pi_B}{\partial L_S} = \left(1 + r_L^S\right)(1 + \xi_2) - \xi_1 + \xi_3 + \xi_5 - 1 = 0, \tag{23.5}$$

$$\frac{\partial \Pi_B}{\partial S_S} = -\left(1 + r_L^S\right)(1 + \xi_2) + p_S(1 + \xi_1) - \xi_3 = 0, \tag{23.6}$$

$$\frac{\partial \Pi_B}{\partial T_S} = \left(1 + r_L^S\right)(1 + \xi_2) - p_S(1 + \xi_1) = 0 \tag{23.7}$$

$$\frac{\partial \Pi_B}{\partial D} = -(1 + r_D)(1 + \xi_2) + (1 + \xi_1) = 0, \tag{23.8}$$

$$\frac{\partial \Pi_B}{\partial L_R} = \pi\left(1 + r_L^R\right) - \xi_1 + l\xi_4 - 1 = 0, \tag{23.9}$$

$$\frac{\partial \Pi_B}{\partial S_R} = -\pi\left(1 + r_L^R\right) + p_R(1 + \xi_1) - \xi_4 = 0, \tag{23.10}$$

$$\frac{\partial \Pi_B}{\partial T_R} = \left(1 + r_L^R\right)(\pi + \pi_L\xi_2) - p_R(1 + \xi_1) = 0. \tag{23.11}$$

As securitized loans of safe loans must be safe, the return on them and safe deposits must be identical, thus

$$1 + r_D = \frac{1 + r_L^S}{P_S}. \tag{23.12}$$

Using equations (23.12) and (23.8), we get from equation (23.6) that $\xi_3 = 0$ and therefore the constraint that $L_S \geq S_S$ cannot be binding. In this case we see that due to equation (23.12) the marginal benefits in equations (23.6) and (23.8) are identical. Thus there is no benefit from raising funds through securitization compared to raising these directly from investors as deposits. Therefore there is no incentive to securitize safe loans and we set $S_S = 0$. If no loans are securitized, then no securitized loans can be bought, thus $T_S = 0$.We thus see that safe loans are never securitized.

For risky loans, investors will prefer securitized risky loans if their returns exceed those of the deposits, i.e. if $\frac{\pi\left(1+r_L^R\right)}{P_R} > 1 + r_D$. After inserting equstion (23.8), we obtain from equation (23.11) that

$$\begin{aligned}\frac{\partial \Pi_B}{\partial T_R} &= \left(1 + r)L^R\right)(\pi + \pi_L\xi_2) - P_R(1 + r_D)(1 + \xi_2) \\ &\geq \left(\pi\left(1 + r_L^R\right) - P_R(1 + r_D)\right)(1 + \xi_2) \\ &> 0,\end{aligned} \tag{23.13}$$

where the first inequality uses that $\pi \geq \pi_L$ and the last inequality the higher returns on securitized risky loans. With investors demanding securitized loans directly and the condition in equation (23.13), demand for risky loans would be infinite. Hence we need $\frac{\pi\left(1+r_L^R\right)}{P_R} \leq 1 + r_D$ to eliminate the demand of investors and any securitized loans would be held by banks only, i.e. $T_R = S_R$.

Let us assume that $S_R = T_R = 0$ such that the constraint $L_R \geq S_R$ is not binding, implying $\xi_4 = 0$, hence from equation (23.10) we have

$$\frac{\partial \Pi_B}{\partial S_R} = -\pi\left(1 + r_L^R\right) + P_R\left(1 + \xi_1\right) = 0, \tag{23.14}$$

which we can insert into equation (23.11) to obtain that

$$\frac{\partial \Pi_B}{\partial T_R} = \xi_2 \pi_L \left(1 + r_L^R\right). \tag{23.15}$$

Hence, if constraint (23.3) is binding such that $\xi_2 > 0$, then $S_R = T_R = 0$ cannot be optimal. If on the other hand $\xi_4 > 0$, then the constraint is binding and $S_R = T_R = L_R$, such that all risky loans are securitized.

If, on the other hand, constraint (23.3) is not binding and $\xi_2 = 0$, then when inserting equation (23.11) into equation (23.10), we see that $\xi_4 = 0$, implying that $S_R < L_R$. If there are no benefits to maximize securitization, then no securitization would be optimal, thus $S_R = T_R = 0$.

These results imply that safe loans are never securitized and risky loans are fully securitized if constraint (23.3) is binding and the securitized loans are held in the banks themselves. If constraint (23.3) is not binding, no securitization should occur.

We can now use these results to determine the securitization strategy of the bank. Let us firstly assume that $D < \overline{L}$ and hence as only D can be invested into loans, $L_S < \overline{L}$, implying that $\xi_5 = 0$. We know that safe loans are not securitized, thus $\xi_3 = 0$. The marginal benefits of investing into safe loans are higher than those from investing into risky loans, if from comparing equations (23.5) and (23.9), we have

$$\left(1 + r_L^S\right) - \pi\left(1 + r_L^R\right) \geq \xi_4 - \xi_2\left(1 + r_L^S\right). \tag{23.16}$$

The left hand side is positive by the assumption above and $\xi_2 \geq 0$, hence for $\xi_4 = 0$, risky loans are not being securitized, safe loans are preferred as their marginal benefits are higher. If $\xi_4 > 0$ we know that $S_R = T_R$ and hence given that only net positions of securitized loans are relevant, we need $\frac{\partial \Pi_B}{\partial S_R} = -\frac{\partial \Pi_B}{\partial T_R}$. Inserting from equations (23.10) and (23.11), we obtain that

$$\xi_4 = \xi_2 \pi_L \left(1 + r_L^R\right). \tag{23.17}$$

Inserting this expression back into equation (23.16), we see that as $\pi > \pi_L$ and $1 + r_L^S > \pi\left(1 + r_L^R\right)$, the expression is always fulfilled. Thus banks will invest into safe loans first.

Hence for $D < \overline{L}$ all loans are safe and constraint (23.3) cannot be binding, such that $\xi_2 = 0$. Using equation (23.8) and that $\xi_3 = 0$ we can rewrite equation (23.5) as

$$\frac{\partial \Pi_B}{\partial L_S} = \left(1 + r_L^S\right) - (1 + r_D) = 0. \tag{23.18}$$

hence

$$1 + r_D = 1 + r_L^S. \tag{23.19}$$

If $D > \overline{L}$, then banks invest into risky loans as long as it is profitable to do so, i.e. $\pi\left(1 + r_L^R\right) > 1$, and banks avoid expected losses. Let us assume again that equation (23.3) is not binding, i.e. $\xi_2 = 0$. As any securitized risky loans would be safe, we will have $P_R = 1$. The marginal benefits of granting risky loans from equation (23.9) will exceed those of buying securitized loans from equation (23.11) if $\xi_4 \geq 0$ as can easily be verified. Given that banks prefer granting new loans to buying securitized loans, there will be no demand for securitization and hence no loans can be securitized, such that $S_R = T_R = 0$. To ensure constraint (23.3) is fulfilled with $L_S = \overline{L}$, given the safe loans are taken up first and $T_R = 0$, we need

$$D \leq D^* = \frac{1 + r_L^S}{\pi\left(1 + r_L^R\right)}\overline{L}. \tag{23.20}$$

For $\overline{L} < D \leq D^*$ banks invest $\overline{L}$ into safe loans and $D - \overline{L}$ into risky loans. No securization is needed as the return on the safe loans are sufficient to cover any losses arising from the small amount of risky loans.

Increasing deposits beyond D^* requires then to conduct securitization to supplement the safe loans in ensuring risk-free deposits. We know from above that with securitization, all risky loans are securitized, hence $T_R = L_R = D - \overline{L}$ and $L_S = \overline{L}$. Thus (23.3) requires that

$$D \leq D^{**} = \frac{1 + r_L^S - \pi\left(1 + r_L^R\right)}{1 + r_D - \pi_L\left(1 + r_L^R\right)}\overline{L}. \tag{23.21}$$

In this case constraint (23.3) is not binding and we maintain $\xi_2 = 0$. Furthermore $\xi_4 = 0$ as this constraint is also not binding. Inserting this and equation (23.8) into equation (23.9), we easily get

$$1 + r_D = \pi\left(1 + r_L^R\right), \tag{23.22}$$

which inserted into equation (23.21) gives

$$D^{**} = \frac{1 + r_L^S - \pi\left(1 + r_L^R\right)}{(\pi - \pi_L)\left(1 + r_L^R\right)}\overline{L}. \tag{23.23}$$

Thus for $D^* \leq D \leq D^{**}$ risky loans are securitized but the return of the safe loans allows to maintain risk-free deposits. We can interpret the safe loans as collateral for the securitized risky loans.

If D increases further, constraint (23.3) becomes binding and the only way to maintain this constraint is by reducing the interest paid on deposits such that

$$1 + r_D = \pi_L \left(1 + r_L^R\right) - \left(\pi_L \left(1 + r_L^R\right) - \left(1 + r_L^S\right)\right) \frac{\overline{L}}{D}, \tag{23.24}$$

which reduces as D increases. The reduction is not the result of the increased supply of deposits directly, but is required to maintain a safe return.

Investors are not willing to make losses as they could retain these deposits as cash. Thus we require $r_D \geq 0$, or another benchmark at which funds can be invested elsewhere. From equation (23.24) this implies that

$$D \leq \overline{D} = \frac{\left(1 + r_L^S\right) - \pi_L \left(1 + r_L^R\right)}{1 - \pi_L \left(1 + r_L^R\right)} \overline{L}. \tag{23.25}$$

Thus deposits larger than $\overline{D}$ are not sustainable as the bank cannot guarantee them to be safe and yield a positive return.

Risky loans are securitized to diversity the default risk and the safe loans can be interpreted as additional collateral, for example in form of government securities, to ensure the safety of the securitized risky loans. These thus securitized loans can be off-loaded to a special purpose vehicle and sold to a shadow bank. Therefore this model shows that banks have incentives to provide such shadow banks with suitable securitized loans.

Reading Gennaioli, Shleifer, & Vishny (2013)

23.2 Regulatory arbitrage with shadow banks

Banks can avoid the costs of regulatory constraints, in particular capital requirements, by securitizing their loans and selling them to a third party. thereby banks remove the loans from their balance sheets and can use to proceeds obtained from the securitized loans to grant additional loans. It are often shadow banks that purchase such securitized loans and we will explore here how capital requirements for banks, while no such requirements exist for shadow banks, can induce banks to securitize loans and support a shadow banking system as a way of regulatory arbitrage, that is the exploitation of different sets of regulations for banks and shadow banks.

We assume that banks finance their loans L using equity E and deposits D such that $L = E + D$. By regulation, banks are required to hold a fraction κ of their assets as equity, hence

$$\begin{aligned} E &= \kappa L, \\ D &= (1 - \kappa) L. \end{aligned} \tag{23.26}$$

On the loan portfolio of the bank, two possible outcomes are possible. The first outcome, which happens with a probability p, consists of a high repayment rate π_H on their loans; the other outcome, occurring with probability $1 - p$, has higher defaults such that its repayment rate is $\pi_L \geq \pi_H$. The bank does not know in advance which state occurs. With r_D denoting the interest on deposits and r_L on loans, we assume that

$$\pi_H (1 + r_L) > 1 + r_D > \pi_L (1 + r_L) \,. \tag{23.27}$$

If the default rate is high, π_L, the bank cannot repay its depositors and we assume that a regulator seizes the bank and covers any shortfall such that deposits are risk-free. This equivalent to an implied deposit insurance by the government and the bank will fail, with the bank owner making a full loss of their equity.

In order to attract equity for the bank, its expected return must exceed that of deposits. If the bank does not fail, hence the low default rate is realised, the return on equity is r_E and this expected return needs to exceed the the deposit rate as otherwise investors would not provide equity but deposits only. Thus we need $p(1 + r_E) > 1 + r_D$. Setting an equity risk premium of τ, we can determine the return on equity as

$$1 + r_E = \frac{1 + r_D}{p} + \tau. \tag{23.28}$$

Incentives to securitize loans If loans are securitized and sold to shadow banks, the yield on these securitized is set as r_S if the low default rate π_H prevails, with the bank profiting from the difference to $\pi_H (1 + r_L)$, and if the outcome is less favourable with a high default rate π_L, they obtain the return generated, $\pi_L (1 + r_L)$. If the high default rate π_L is realised, the projected yield on the securitized loans is not achieved and the purchasers of the loans are given the full proceeds the bank obtains on the loans; consequently banks do not make any profits in this case.

Investors into securitized loans are assumed to be indifferent between this investment and providing deposits such that $p(1 + r_S) + (1 - p)\pi_L (1 + r_L) = 1 + r_D$. This easily solves for a yield on the securitized loans of

$$1 + r_S = \frac{(1 + r_D) - (1 - p)(\pi_L (1 + r_L))}{p}. \tag{23.29}$$

The highest feasible capital requirements are such that even with high default rates, the bank can repay its deposits. From $\pi_L (1 + r_L) L > (1 + r_D) D$, we get with equation (23.26) that

$$\kappa \leq \overline{\kappa} = \frac{(1 + r_D)\pi_L (1 + r_L)}{1 + r_D}. \tag{23.30}$$

In general, the bank will retain a fraction λ of the loans on its books and sell a fraction $1 - \lambda$ of loans to shadow banks through securitization, promising to pay $1 + r_S$, but not having to make up the shortfall if a return of $\pi_L (1 + r_L)$ is realized. Hence the bank will make a profit of $(\pi_H (1 + r_L)) - (1 + r_S)$ if the high state is realized and none otherwise. We thus have the profits of the bank as

$$\begin{aligned} \Pi_B = {} & p(\pi_H (1 + r_L) - (1 - \kappa)(1 + r_D) - \kappa(1 + r_E))\lambda L \\ & + p(\pi_H (1 + r_L) - (1 + r_S))(1 - \lambda) L. \end{aligned} \tag{23.31}$$

These profits consist of the loans retained yielding a high repayment rate and which are financed by a combination of deposits and equity, and the profits from the loans sold, which do not require any financing. Note that in case the high default rate is realised, the bank will fail and the bank profits will be zero due to limited liability.

Any capital requirements imposed will be binding as we can show that

$$\frac{\partial \Pi_B}{\partial \kappa} = p\left(\left(1+r_D\right)-\left(1+r_E\right)\right)\lambda L = -p\left(\left(1+r_D\right)\frac{1-p}{p}+\tau\right)\lambda L < 0, \quad (23.32)$$

where the second equality emerges from inserting equation (23.28). Hence banks prefer lower capital requirements as this would increase their profits.

As the bank profits are linear in λ, the bank will choose to either retain all loans on the book or sell all loans to shadow banks. If the profits from shadow banking, the last term in equation (23.31), exceed those of loans staying on the book, the first term in equation (23.31), then shadow banking will occur and traditional banking disappear. We thus need

$$\pi_H\left(1+r_L\right)-\left(1+r_S\right) \geq \pi_H\left(1+r_L\right)-\left(1-\kappa\right)\left(1+r_D\right)-\kappa\left(1+r_E\right),$$

which after inserting from equations (23.28) and (23.29) becomes

$$\kappa \geq \hat{\kappa} = \frac{\left(1-p\right)\left(\left(1+r_D\right)-\left(\pi_L\left(1+r_L\right)\right)\right)}{\left(1-p\right)\left(1+r_D\right)+p\tau}. \quad (23.33)$$

We can easily show that for $\tau > 0$, thus the presence of an equity premium, we have $\hat{\kappa} < \overline{\kappa}$ and for sufficiently high capital requirements shadow banking emerges, in line with the idea of regulatory arbitrage. Banks would securitize loans and sell them to shadow banks as the absence of capital requirements reduces their funding costs, allowing banks to offer a yield on these loans that is attractive to shadow banks, but less than their funding costs that include the equity premium.

Bank guarantees for securitized loans It is common for banks to provide guarantees or collateral for their securitized loans. Let us assume therefore that banks would transfer a fraction γ of their retained assets, λL, to the shadow bank if the high default rate is realized. The return to shadow banks with such guarantees are denoted by $1+\hat{r}_S$. If the high default rate is realized, the shortfall the shadow bank faces is $\left(\left(1+\hat{r}_S\right)-\pi_L\left(1+r_L\right)\right)\left(1-\lambda\right)L$. In order to guarantee the full return $1+\hat{r}_S$ to the shadow bank, the transfer $\gamma\lambda L$ needs to be sufficient to cover this shortfall. Hence we need

$$\gamma\lambda L > \left(\left(1+\hat{r}_S\right)-\pi_L\left(1+r_L\right)\right)\left(1-\lambda\right)L.$$

This can be solved for the maximum return that can be supported,

$$1+\hat{r}_S \leq \pi_L\left(1+r_L\right)+\frac{\lambda}{1-\lambda}\gamma. \quad (23.34)$$

The returns the shadow bank obtains has again to be equivalent to that of deposits, such that they are attracted to it. Thus we require

$$p\left(1+\hat{r}_S\right)+(1-p)\min\left\{1+\hat{r}_S,\left(\pi_L\left(1+r_L\right)\right)+\frac{\lambda}{1-\lambda}\gamma\right\}=1+r_D. \qquad (23.35)$$

If the default rate is low, the shadow bank obtains the agreed yield on the securitized loans, but of the default rate is high, it will obtain the lower return on the loans, adjusted by the guarantee provided by the bank, up to the agreed yield. This condition solves for the yield on securitized loans to be

$$1+\hat{r}_S=\begin{cases}\frac{(1+r_D)-(1-p)\left(\pi_L(1+r_L)+\frac{\lambda}{1-\lambda}\gamma\right)}{p} & \text{if } 1+\hat{r}_S\geq\pi_L\left(1+r_L\right)+\frac{\lambda}{1-\lambda}\gamma\\ 1+r_D & \text{if } 1+\hat{r}_S<\pi_L\left(1+r_L\right)+\frac{\lambda}{1-\lambda}\gamma\end{cases}. \qquad (23.36)$$

Inserting the first case into equation (23.34), we can solve this equation for

$$\lambda\leq\frac{\left(1+r_D\right)-\pi_L\left(1+r_L\right)}{\left(1+r_D\right)-\pi_L\left(1+r_L\right)+\gamma}. \qquad (23.37)$$

Thus we can rewrite equation (23.36) as

$$1+\hat{r}_S=\begin{cases}\frac{(1+r_D)-(1-p)\left(\pi_L(1+r_L)+\frac{\lambda}{1-\lambda}\gamma\right)}{p} & \text{if } \lambda\leq\frac{(1+r_D)-\pi_L(1+r_L)}{(1+r_D)-\pi_L(1+r_L)+\gamma}\\ 1+r_D & \text{if } \lambda>\frac{(1+r_D)-(\pi_L(1+r_L))}{(1+r_D)-\pi_L(1+r_L)+\gamma}\end{cases}. \qquad (23.38)$$

We easily see that the yield bank promise on securitized loans is lower with the bank guarantee, $1+\hat{r}_S<1+r_S$, due to the additional return the guarantee provides if the high default rate is realised.

Banks will only provide such guarantees if it is profitable to do so. If the capital requirements are high, $\kappa>\hat{\kappa}$, we have seen that even without guarantees all loans are sold and only shadow banks exist. Hence for guarantees to be provided in this case, we require that the profits when providing guarantees and offering a yield of $\hat{r}_S$ on securitized loans, equation (23.31), exceed the profits when not providing the guarantee and offering a yield of r_S; as loans are securitized, banks have no funding costs.

$$\begin{aligned}&p\left(\left(\pi_H\left(1+r_L\right)\right)-(1-\kappa)\left(1+r_D\right)-\kappa\left(1+r_E\right)\right)\lambda L\\&\quad+p\left(\left(\pi_H\left(1+r_L\right)\right)-\left(1+\hat{r}_S\right)\right)(1-\lambda)L\\&\quad\geq p\left(\left(\pi_H\left(1+r_L\right)\right)-\left(1+r_S\right)\right)L.\end{aligned}$$

We note that the payment of the guarantee does not affect bank profitability as the guarantee is only paid if the high default rate is realised and in this case the bank generates no profits in any case as it fails. This condition can be transformed into

$$\begin{aligned}&\lambda\left(\kappa\left(\left(1+r_E\right)-\left(1+r_D\right)\right)-\left(\left(1+\hat{r}_S\right)\left(1+r_D\right)\right)\right)\\&\quad\leq\left(1+r_E\right)-\left(1+\hat{r}_S\right).\end{aligned} \qquad (23.39)$$

We can now insert from equations (23.28), (23.29), and (23.38) to obtain the condition for guarantees to be provided. For the case in equation (23.38) that $\lambda \leq \frac{(1+r_D)-\pi_L(1+r_L)}{(1+r_D)-\pi_L(1+r_L)+\gamma}$, this can be transformed into

$$\kappa \leq \kappa^* = \hat{\kappa} + \frac{(1-p)\,\gamma}{(1-p)\,(1+r_D)+p\tau}. \tag{23.40}$$

For larger levels of securitization to shadow banks, λ, we can show that the condition in equation (23.39)is always fulfilled.

In the case that $\kappa < \hat{\kappa}$, only banks would exist without guarantees. Thus for guarantees to be given by banks we need

$$\begin{aligned} &p\,((\pi_H\,(1+r_L)) - (1-\kappa)\,(1+r_D) - \kappa\,(1+r_E))\,\lambda L \\ &\quad + p\,((\pi_H\,(1+r_L)) - (1+\hat{r}_S))\,(1-\lambda)\,L \\ &\qquad \geq p\,((\pi_H\,(1+r_L)) - (1-\kappa)\,(1+r_D) - \kappa\,(1+r_E))\,L. \end{aligned}$$

Hence for guarantees to be provided in this case, we require that the profits when providing guarantees and offering a yield of $\hat{r}_S$ on securitized loans, equation (23.31), exceed the profits when not providing the guarantee and retaining the loans, incurring funding costs. This condition simplifies to

$$(1+\hat{r}_S) - (1+r_D) \leq \kappa\,((1+r_E) - (1+r_D)). \tag{23.41}$$

Using equations (23.28) and (23.38) we can easily show that for $\lambda > \frac{(1+r_D)-\pi_L(1+r_L)}{(1+r_D)-\pi_L(1+r_L)+\gamma}$, equation (23.41) is always fulfilled and guarantees are always provided. For smaller levels of securitization of loans we require that

$$\kappa \geq \kappa^{**} = \hat{\kappa} - \frac{\lambda}{1-\lambda}\frac{(1-p)\,\gamma}{(1-p)\,(1+r_D)+p\gamma}. \tag{23.42}$$

Therefore if $\kappa^{**} \leq \kappa \leq \kappa^*$ banks would provide guarantees to their shadow banks. The optimal amount λ to transfer can be obtained when maximizing bank profits

$$\begin{aligned} \hat{\Pi}_B = &\; p\,((1+r_H) - (1-\kappa)\,(1+r_D) - \kappa\,(1+r_E))\,\lambda L \\ &+p\,((\pi_H\,(1+r_L)) - (1+\hat{r}_S))\,(1-\lambda)\,L, \end{aligned} \tag{23.43}$$

which is identical to the profits in equation (23.31), except for the yield on securitized loans taking into account the existence of a guarantee.

If $\lambda > \frac{(1+r_D)-\pi_L(1+r_L)}{(1+r_D)-\pi_L(1+r_L)+\gamma}$, we have from equation (23.38) that $1+\hat{r}_S = 1+r_D$ and it is easy to show that we obtain $\frac{\partial \hat{\Pi}_B}{\partial \lambda} < 0$ and banks would choose to securitize as little as possible, making this area of high transfer unsustainable as the condition requires a high level of securitization.

In the case that $\lambda \leq \frac{(1+r_D)-\pi_L(1+r_L)}{(1+r_D)-\pi_L(1+r_L)+\gamma}$ the first order condition $\frac{\partial \hat{P}B}{\partial \lambda} = 0$ solves for the optimal amount to transfer and similarly we can obtain the optimal guarantee by setting $\frac{\partial \hat{P}B}{\partial \gamma} = 0$.

We thus see that for $\kappa^{**} \leq \kappa \leq \kappa^*$ banks will provide guarantees to shadow banks and therefore banks and shadow banks will co-exist. In the case that $\kappa < \kappa^{**}$ only banks will exist and for $\kappa > \kappa^*$ only shadow banks will exist. The provision of guarantees to shadow banks extends the viability of shadow banks for lower capital requirements to banks and the viability of banks is extended to higher capital requirements. Regulatory arbitrage works in that banks reduce the burden of capital requirements, and thus the higher funding costs from holding equity, by off-loading same loans to shadow banks and guaranteeing their loans, where we assumed that such guarantees do not attract capital requirements. However, as banks' assets are seized in the low state, exactly when banks need to honour their guarantee, they do not incur any losses directly.

Figure 23.2 shows the results graphically. It can be shown that the optimal degree of securitization, λ is increasing in the capital requirements, κ. If no securitization occurs, $\lambda = 0$, we can interpret this as only banks being present in the market, while for $\lambda = 1$, that is full securitization of loans, only shadow banks are present in the market; intermediate levels of securitization, $0 < \lambda < 1$, correspond to the co-existence of banks and shadow banks.

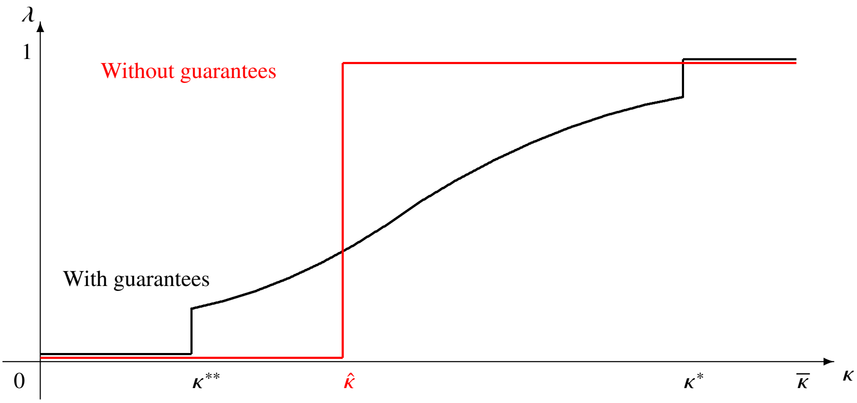


Fig. 23.2: Optimal level of securitization with regulatory arbitrage

The guarantee γ is usually held in the form of liquid securities that are pledged to the shadow bank. They can thus form part of the liquidity reserves a bank is required to hold, imposing no further costs on the bank in holding liquid assets that cannot be invested into loans. Furthermore, rather than focussing on capital requirements, the same results would emerge if liquidity requirements are imposed on banks but not on shadow banks. In this case, the lower return on liquidity reserves compared to loans will impose a cost on banks that shadow banks do not face, resulting in a comparable incentive for banks to securitize loans and by reducing their balance sheet, reducing deposits and thus liquidity requirements.

Summary We have seen that if the costs of banks are higher than the costs of shadow banks, for example due to capital requirements, it can be beneficial for banks to securitize loans and sell these to shadow banks. As shadow banks face lower costs when holding these loans, they will be requiring a lower return than banks and if the cost-differential is sufficiently high, banks will forego the full profits from retaining the loans and sell them to shadow banks. Banks can provide guarantees against a too low return on the securitized loans, making them more attractive to shadow banks, but also less costly to banks as the guarantee only becomes relevant if the banks fails. In this case, we see a gradual increase of the level of securitization as the costs of banks, the capital requirements, are increasing.

The emergence of shadow banks is here driven by the cost advantage of shadow banks due to the lack of regulation of their activities. We can therefore interpret the securitization of banks as being the result of regulatory arbitrage; banks could set up shadow banks, such as money market funds, and transfer loans to these entities with the aim to reduce their cost of the regulatory requirements.

Reading Górnicka (2016)

Conclusions

We have seen that shadow banking emerges as a response to regulatory arbitrage. Banks will seek to off-load some or all loans from their balance sheet to other entities, shadow banks, that are facing lower regulatory costs on holding such loans, such as reduced or no capital requirements. Despite their reduced return from granting such loans in the first place, the lower regulatory costs make this process more profitable to banks than retaining the loans on their balance sheet. We have also seen that banks are not only willing to sell loans, but that they will find it optimal to generate securitized loans in such a way that they are risk-free to the purchaser, the shadow bank.

Chapter 24

Alternative lending arrangements

Banks are not only competing with each other to obtain deposits and grant loans, but they also face competition from other financial institutions that are not recognised as banks in the traditional sense and also not subject to comparable regulatory constraints. From the perspective of companies taking out loans, it is not important whether the loan is provided by a bank or another non-bank organisation as they are only concerned about the conditions of any such loans. In this chapter we will look at a number of such alternative sources of loans and explore how their presence might affect the banks.

Direct lending from a 'depositor' to a borrower using a suitable matching mechanism is commonly referred to as peer-to-peer lending and seen as a viable alternative for borrowers and for depositors who do not rely on access to their funds. Chapter 24.1 will therefore explore the implication the presence of such lending mechanism has on banks. Banks finance their loans through deposits and have no direct interest in how the company spends the proceeds from the loan. In contrast to that finance companies affiliated with retailers might have an inherent interest in granting a loan that is then spend on the products they sell. Such consumer finance and how it affects the ability of borrowers obtaining loans is considered in chapter 24.2.

Both, peer-to-peer lending and consumer finance provide traditional loans to borrowers where an amount is advanced to the borrower and this is then repaid by that borrower at a later point with interest. Due to religious beliefs, some societies do not allow interest to be charged and banking has to adjust to this requirement. In chapter 24.3 we will explore the impact the presence of Islamic banks has on conventional banks and also how conventional banks affect Islamic banks. Another alternative form of borrowing funds is through group lending; here each borrower is liable for the repayment not only of his own loan but also of the loans of all other members of the group, it thus a joint liability loan. Such loans are common in many developing countries and typically comprise small loans. The impact of such joint liability will be explored in more detail in chapter 24.4.

24.1 Peer-to-peer lending

Depositors can provide loans directly to borrowers, which is often referred to as peer-to-peer lending. Such lending is usually conducted with the help of an intermediary, commonly through an online platform, that facilitates this lending by assessing the risks of borrowers and managing the disbursement of funds to the borrower as well as any payment to the lender. This intermediary does not take any active role in the lending process and is only an administrator with the risks arising from the ability of the borrower to repay the loan solely borne by the 'depositor'.

Such peer-to-peer lending platforms will compete with banks for depositors and borrowers. Compared to peer-to-peer lending, banks are assumed to have an advantage in managing the risks of companies borrowing funds; the probability of a loan being repaid to banks, π_i, will be higher than the probability of a loan being repaid in peer-to-peer lending, $\hat{\pi}_i$, hence $\pi_i > \hat{\pi}_i$. This difference might arise from a better ability of banks to monitor companies after a loan has been granted to ensure funds are used appropriately and the additional support and advice banks might provide to companies in danger of not being able to repay their loans.

We assume that companies are either having a high probability of repaying the loan, π_H and $\hat{\pi}_H$, respectively, ot a low probability of repaying the loan, π_L and $\hat{\pi}_L$, respectively, with $\pi_H > \pi_L$ and $\hat{\pi}_H > \hat{\pi}_L$. We assume that the relative disadvantage of the peer-to peer lender is the same for both types of companies. Thus we assume

$$\frac{\pi_H}{\hat{\pi}_H} = \frac{\pi_L}{\hat{\pi}_L}. \tag{24.1}$$

In addition, peer-to-peer lenders face costs of c to gain customers from banks. Peer-to-peer lenders face no funding costs and hence their profits are

$$\Pi_P = \hat{\pi}_i \left(1 + \hat{r}_L^i\right) L - L - c, \tag{24.2}$$

where L is the loan amount and $\hat{r}_L^i$ the loan rate for type i, which we assume to be known to both the bank and peer-to-peer lender.

Banks have to hold equity E and finance their loans L with deposits D such that $L = D + E$. Furthermore, the regulatory scrutiny that banks are subjected to, costs them the amount of C.

We propose that bank lending itself needs to be profitable as else the bank could invest into safe assets. With perfect competition, the profits from lending itself, ignoring the requirements of holding capital and the regulatory costs, will be eroded such that

$$\pi_L \left(\left(1 + r_L^i\right) L - \left(1 + r_D^i\right) D\right) = 0,$$

which solves for

$$\left(1 + r_D^i\right) D = \left(1 + r_L^i\right) L. \tag{24.3}$$

We have here assumed that banks have limited liability and will not be able to repay their deposits if the loans are not repaid.

For depositors to provide funds to the bank, we need them to break even, hence

$$\pi_i \left(1 + r_D^i\right) D = D,$$

or when using equation (24.3), we have

$$D = \pi_i \left(1 + r_L^i\right) L. \tag{24.4}$$

As companies have no costs of switching their borrowing from banks to peer-to-peer lending, we need the loan rates of these two lending forms to be identical, $1 + r_L^i = 1 + \hat{r}_L^i$. With perfect competition, peer-to-peer lenders generate no profits, $\Pi_P = 0$, and we get from equation (24.2) that

$$\left(1 + r_L^i\right) L = \frac{L + c}{\hat{\pi}_i}. \tag{24.5}$$

Using equations (24.3), (24.4), and (24.5), we get for the profits of banks as

$$\begin{aligned} \Pi_B^i &= \pi_i \left(\left(1 + r_L^i\right) L - \left(1 + r_D^i\right) D\right) - E - C \qquad (24.6)\\ &= \pi_i \left(\left(1 + r_L^i\right) L - \left(1 + r_D^i\right) D\right) - (L - D) - C \\ &= \pi_i \left(1 + r_L^i\right) L - \pi_i \left(1 + r_D^i\right) D - L + D - C \\ &= -L + \pi_i \left(1 + r_L^i\right) L - C \\ &= -L + \frac{\pi_i}{\hat{\pi}_i} (L + c) - C \\ &= \frac{\pi_i}{\hat{\pi}_i} c + \frac{\pi_i - \hat{\pi}_i}{\hat{\pi}_i} L - C. \end{aligned}$$

If $\Pi_B^L < 0$, then the bank would not offer companies of type L a loan. This condition is fulfilled if

$$c < \overline{c} = \frac{\hat{\pi}_L}{\pi_L} C - \frac{\pi_L - \hat{\pi}_L}{\pi_L} L. \tag{24.7}$$

Thus, if the costs of peer-to-peer lenders to attract depositors from banks are sufficiently low, the competition between banks and peer-to-peer lenders will not allow banks to lend profitably to companies with high risks, that is those with a low probability to repay loans. The reason is that the additional regulatory costs that bank face makes it unprofitable for them to offer loans to high-risk companies; peer-to-peer lending with its lower regulatory costs can provide these loans at lower costs, provided the costs of attracting depositors to them is not high.

To continue offering loans to companies of type H, we need $\Pi_B^H \geq 0$. This requires

$$c \geq \underline{c} = \frac{\hat{\pi}_H}{\pi_H} C - \frac{\pi_H - \hat{\pi}_H}{\pi_H} L. \tag{24.8}$$

Competition between banks and peer-to-peer lenders will be sufficiently low to retain low risk companies, those with high probabilities of repaying the loan, if the costs of attracting depositors from banks to peer-to-peer lending is sufficiently high. The additional regulatory costs give banks an inherent disadvantage compared to peer-to-peer lending and they will only be able to retain low-risk loans if the costs of attracting deposits to peer-to-peer lending are sufficiently high.

Using our assumption that $\frac{\pi_H}{\hat{\pi}_H} = \frac{\pi_L}{\hat{\pi}_L}$, we can easily show that $\underline{c} < \overline{c}$. Hence if $\underline{c} < c < \overline{c}$, high-risk companies, L, will migrate to peer-to-peer lenders and low risk companies, H, will remain with banks. If $c > \overline{c}$, peer-to-peer lenders face too high costs to gain deposits, or the regulatory costs C to banks are too low for them to stop providing high-risk loans. For small costs $c < \underline{c}$, all firms switch to peer-to-peer lending as the regulatory costs for banks are too high.

We have thus established that the presence of peer-to-peer lending competing with traditional banks can lead to a situation where high-risk companies are provided with loans directly through peer-to-peer lending while low-risk companies remain with traditional banks. This result is drive by the advantage that peer-to-peer lending has due to a lower regulatory burden; the costs of attracting deposits from banks do not fully outweigh these cost advantages and this benefits foremost high-risk companies as they are less profitable to banks.

Reading de Roure, Pelizzon, & Thakor (2021)

24.2 Consumer finance

Banks face competition in lending to consumer from finance companies associated with the seller of the goods. The seller has an interest in increasing his sales as this would generate a profit and if consumers are not able to make purchases without obtaining a loan, these profits might affect the willingness to provide loans. Sellers offering loans to their customers, commonly through associated finance companies, is a common practice in the purchase of high-value goods such as cars, household appliances, and electronic goods.

Let us assume a consumer seeks to purchase a good at price P for which he requires a loan to the same amount. It is common knowledge that the consumer repays the loan with probability π and there is no collateral available for this loan. The loan rate r_L is also given and identical for both the bank and the consumer finance such that consumers are indifferent between the originator of their loan.

A bank will finance a loan to the amount of P fully through deposits on which interest r_D is payable and deposits are always repaid; thus we do not allow the bank to default. Banks will provide a loan to the consumer of it is profitable to do so, thus

$$\Pi_B^i = \pi \left(1 + r_L\right) P - \left(1 + r_D\right) P \geq 0 \tag{24.9}$$

or

$$\pi \geq \pi^* = \frac{1 + r_D}{1 + r_L}. \tag{24.10}$$

Provided the risk the consumer poses is not too high, the bank will grant the loan.

The company selling the product might now seek to expand its market by offering to provide loans to customers themselves. We assume that for such loans the funding costs are higher than those of banks, $\hat{r}_D > r_D$. Funding costs for consumer funding will in general be higher than for banks due the higher risks involved,a s we will establish below, and the lack of regulatory oversight of such companies. In addition, companies make profits by selling the good at price P but having costs of C, thus making a profit of $P - C$ from the same of the good. These profits from selling the good, which would not be possible if not granting the loan, will add to the profits generated by the loan itself. The company will grant the loan as long as doing so is profitable, thus as long as

$$\hat{\Pi}_C^i = (P - C) + (\pi (1 + r_L) P - (1 + \hat{r}_D) P) \geq 0, \tag{24.11}$$

where the first term denotes the profits from selling the product and the second term the profits from financing the sale. Hence a consumer loan will be provided to customers if

$$\pi \geq \hat{\pi}^* = \frac{1 + \hat{r}_D}{1 + r_L} - \frac{P - C}{P} \frac{1}{1 + r_L}. \tag{24.12}$$

The financing of the purchase through a loan would cause a loss for the company if $\pi < \frac{1+\hat{r}_D}{1+r_L}$, but as long as $\pi \geq \hat{\pi}^*$ this loss is compensated for by the profits from the same of the good. We easily see that the conditions under which loans are granted are extended to more risky consumers, requiring $\hat{\pi}^* < \pi^*$, if

$$1 + \hat{r}_D < 1 + r_D + \frac{P - C}{P} \tag{24.13}$$

and hence the financing costs of the company are not too large compared to those of banks.

Consumer finance will be extended to consumers with higher risks than with bank lending, provided the funding costs in consumer finance is not too high. If we were to assume that banks could offer a loan at a slightly lower loan rate to consumers than companies could, as is often observed in markets, then low-risk consumers, those with $\pi > \pi^*$ would seek bank loans, while consumers with higher risks , those with $\hat{\pi}^* < \pi < \pi^*$, will have to seek a loan from the finance company. Thus finance companies are more likely to provide more risky loans, justifying their higher funding costs.

Reading Barron, Chong, & Staten (2008)

24.3 Islamic banking

In Islamic banking the conventional forms of interest on loans and deposit are not allowed. For lending, the most common form of 'murabaha' involves the bank buying the asset the borrower wishes to acquire and then sells this asset on to the borrower at a markup and allowing the borrower to pay in installments. Hence the bank is exposed to the risk the asset might be destroyed before the borrower takes possession and thus the bank makes a loss. Depositors to not get a fixed interest rate but in the form known as 'mudarabah' they obtain a fraction of the profits made by the bank. Thus depositors do not obtain a fixed interest payment but participate in the profits the bank is able to generate.

Let us denote by π_D the probability that the asset the 'borrower' seeks to acquire is destroyed or damaged while in the possession of the bank. The profits of the bank are then given by

$$\begin{aligned} \Pi_B = & \left(1-\pi_D\right)\left(\pi\left(\left(1+r_L\right)L-L\right)-\left(1-\pi\right)L\right) \\ & -\pi_D L-\left(1-\pi_D\right)\left(\pi\left(\left(1+\beta r_L\right)D-D\right)-\left(1-\pi\right)D\right) \\ & +\pi_D D. \end{aligned} \tag{24.14}$$

Here the first term denotes the investment outcome if the asset is not destroyed and the first part is then the return $r_L L$ required by the bank if the investment succeeds, which it does with probability π. The second part shows the loss to the bank if the investment does not succeed and it does not receive the installment payments. The 'interest' $r_L L$ denotes the markup the bank applies when re-selling the asset to the 'borrower'. The second term denotes the loss to the bank if the asset is destroyed as then the borrower does not have to make any payments. The next two terms similarly denote the payments made to 'depositors' that are only due if the asset is not destroyed and the investment is successful. In this case the 'depositors' are paid a fraction β of the return to the bank. If the asset is destroyed or the installment payments not made, the 'despositors' are not repaid as they participate in the losses of the bank.

'Depositors' need to break even in order to participate, hence the payments the depositors receive from the Islamic bank have to cover the initial outlay. With perfect competition between 'depositors' we require that

$$\left(1-\pi_D\right)\left(\pi\left(\left(1+\beta r_L\right)D-D\right)-\left(1-\pi\right)D\right)-\pi_D D=0. \tag{24.15}$$

We require $D = L$ such that the 'deposits' are fully lent out and no funds left unused. Inserting this and equation (24.15) into equation (24.14), we obtain

$$\Pi_B = \left(1-\pi_D\right)\pi\left(1+r_L\right)L-L. \tag{24.16}$$

We assume that the demand for loans is decreasing in the markup rate (the equivalent to the loan rate), thus $\frac{\partial L}{\partial(1+r_L)} < 0$. The first order condition for a profit maximum of the bank for the optimal markup becomes

$$\begin{aligned}\frac{\partial \Pi_B}{\partial\left(1+r_L\right)} &= \left(1-\pi_D\right)\pi L+\left(1-\pi_D\right)\pi\left(1+r_L\right)\frac{\partial L}{\partial\left(1+r_L\right)} \\ &\quad -\frac{\partial L}{\partial\left(1+r_L\right)} \\ &= 0,\end{aligned} \tag{24.17}$$

which solves for

$$1+r_L=\frac{1}{\pi\left(1-\pi_D\right)}+\frac{L}{\frac{\partial L}{\partial(1+r_L)}}. \tag{24.18}$$

Solving from equation (24.15) we obtain

$$\beta r_L=\frac{\pi_D}{\pi\left(1-\pi_D\right)}+\frac{1-\pi}{\pi}. \tag{24.19}$$

We can compare this 'deposit rate' with the results of a commercial bank. Similarly to equation (24.14), we obtain the profits of a conventional bank as

$$\begin{aligned}\hat{\Pi}_B &= \pi\left(\left(1+\hat{r}_L\right)L-L\right) \\ &= \left(1-\pi\right)L-\pi\left(1+\left(1+r_D\right)D-D\right)+\left(1-\pi\right)D.\end{aligned} \tag{24.20}$$

Here the first term denotes the profits if the company repays, the second the loss if it fails. Similarly the final terms denote the payments to depositors. Note that banks whose loans default cannot repay deposits and hence they are kept as indicated in the final term. The loan rate from conventional banks is denoted by $\hat{r}_L$.

Depositors will only provide deposits if it is profitable to do so and with perfect competition between them this implies that their profits are zero. Hence we require that

$$\pi\left(\left(1+r_D\right)D-D\right)-\left(1-\pi\right)D=0. \tag{24.21}$$

As before we require that loans are fully funded by deposits, $D=L$, and inserting this as well as equation (24.21) into equation (24.20) gives us the bank profits as

$$\hat{\Pi}_B=\pi\left(1+\hat{r}_L\right)L-L. \tag{24.22}$$

The first order condition when maximizing the bank profits for the optimal loan rate, we obtain that

$$\frac{\partial\hat{\Pi}_B}{\partial\left(1+\hat{r}_L\right)}=\pi L-\pi\left(1+\hat{r}_L\right)\frac{\partial L}{\partial\left(1+\hat{r}_L\right)}-\frac{\partial L}{\partial\left(1+\hat{r}_L\right)}=0, \tag{24.23}$$

which gives us

$$1+\hat{r}_L=\frac{1}{\pi}-\frac{L}{\frac{\partial L}{\partial(1+\hat{r}_L)}}. \tag{24.24}$$

From equation (24.21) we easily obtain that

$$1 + r_D = \frac{1}{\pi}. \tag{24.25}$$

Comparing equations (24.18) and (24.24) we see that the total charge of the Islamic bank, $1 + r_L$, exceeds that of the conventional bank, $1 + \hat{r}_L$. Thus borrowing from an Islamic bank is more expensive. A depositor in an Islamic bank obtains a return of βr_L as defined in equation (24.19), this is the equivalent of the deposit rate in equation (24.25). Comparing these two expressions we see that depositors in Islamic banks obtain a higher return. Thus loan and deposit rates in conventional banks are lower than their equivalent in Islamic banks.

If conventional and Islamic banks are in direct competition, they would be attractive for depositors, but not for borrowers. In our model, the key difference driving this result is the additional risk of the asset being destroyed while in the possession of the bank, π_D. We can reasonably claim that the shorter the time period is that the bank owns the asset the lower this probability is. It is straight forward to show that for such shorter holding periods we have $\lim_{\pi_D \to 0} (1 + r_L) = 1 + \hat{r}_L$ and $\lim_{\pi_D \to 0} \beta r_L = r_D$.

Reducing the period the Islamic bank holds the asset before handing it to the 'borrower' to zero is, however, against the idea of Islamic banking in which banks and borrowers share the risks of an investment. Without holding the asset, no risk sharing occurs and the resulting markup could be interpreted as conventional interest, making such a practice not viable.

Thus conventional and Islamic banks can only co-exist if 'borrowers' have a clear preference for Islamic banking on moral or religious grounds. In this case, Islamic banks can sustain the higher 'interest rate' on 'loans' and those preferring conventional banks will accept lower deposit rates than they would be able to obtain in Islamic banks.

Reading Azmat, Azad, Bhatti, & Ghaffar (2020)

24.4 Group lending

It is common that the person obtaining a loan is solely responsible for its repayment, including any guarantees as a collateral that have been agreed. An alternative to such borrowing arrangements could be that a number of borrowers are made jointly responsible for a pool of loans, thus borrowers have joint liability and are required to cover the default of other borrowers. This arrangement is often used in the provision of small loans in developing countries, in what is often referred to as 'microfinance'.

We will investigate the properties of such lending arrangements by firstly looking into the impact such joint liability has on the ability of obtaining loans in chapter 24.4.1. As the key property of such loans is the joint liability for the loans of all member of a selected group, we will explore in chapter 24.4.2 how such groups are selected and chapter 24.4.3 will explore the optimal size of such groups. Joint liability for loans will also affect how risk taking is limited within such groups as we will discuss in chapter 24.4.4, and chapter 24.4.5 will discuss how group lending affects the asymmetric information between borrowers and the lenders. While the entire discussion of group lending and joint liability has assumed that such a contract is offered and borrowers accept them, chapter 24.4.6 will explore conditions under which such joint liability contracts are optimal and prevent strategic default.

24.4.1 Joint liability

The key property in group lending is that all members of the group are jointly responsible for the repayment of all loans that have been granted. This joint liability for the debt of other group members increases the potential costs of borrowers and thus it will have to be balanced against other benefits borrowers obtain from such an arrangement.

Assume that banks are facing a moral hazard problems in that borrowers have the choice between two investments. The return on one investment, R_L, is higher than on the return on the other investment, R_H, hence $R_L > R_H$. On the other hand, the investment with the higher return is also more risky as the probability of success pi_i fulfills $\pi_L < \pi_H$. We assume that $\pi_H (1 + R_H) > 1 + r_D > \pi_L (1 + R_L)$, where r_D denotes the deposit rate; if we assume that loans are fully funded by deposits, this assumption implies that the low-risk investment, pi_H, can generate profits for the bank, while high-risk investment π_L is not able to generate profits. This is because a loan rate cannot be set such that the costs of the banks are covered and the company itself makes a profit from the investment.

In addition to the costs from loans, borrowers face fixed costs $\overline{L}_i$, such that $\overline{L}_H < \overline{L}_L$, and only obtain returns on investments that exceed this amount, thus returns are only generated on an investment of $\hat{L}_i = L - \overline{L}_i$, where L denotes the size of the loan and thus the investment. The high-risk-investment π_L is facing high fixed costs but might generate high returns, while the low-risk investment π_H faces low fixed costs but will also only be able to generate low returns.

We assume that companies are free to choose between these two investments and banks are not able direct the type of investment. As the high-risk investment will not generate profits to the bank, it will seek to provide incentives such that only the low-risk investment is chosen. This reflects the moral hazard problem in the choice of investments by companies.

Furthermore, in addition to its own loan, the joint liability for loans within the group requires borrowers to repay a fraction q of the loan of other borrowers, if those borrower default and the borrower himself succeeds. For simplicity we assume here that only two borrowers are forming a group and by allowing that $q < 0$ we do not impose full joint liability but only that a fraction of the loan of the other borrower is covered. If both borrowers choose the low-risk investment, the profits of the bank are given by

$$\Pi_B = \pi_H (1 + r_L) L + \pi_H (1 - \pi_H) (1 + r_L) qL - (1 + r_D) L = 0 \tag{24.26}$$

for a competitive banking sector in which banks make no profits and where r_L denotes the loan rate the bank applies. Thus we can obtain the loan rate as

$$1 + r_L = \frac{1}{1 + (1 - \pi_H)\, q} \frac{1 + r_D}{\pi_H}. \tag{24.27}$$

The conventional loan in which a borrower is only liable for his own loan is captured by the case that $q = 0$.

Similarly the profits of the borrower is given by

$$\begin{aligned} \Pi_C^i &= \pi_i^2 \left((1 + R_i)\, \hat{L}_i - (1 + r_L)\, L\right) \\ &\quad + \pi_i (1 - \pi_i) \left((1 + R_i)\, \hat{L}_i - (1 + r_L)\, L - (1 + r_L)\, qL\right) \\ &= \pi_i \left((1 + R_i)\, \hat{L}_i - (1 + r_L)\, L\right) - \pi_i (1 - \pi_i)(1 + r_L)\, qL. \end{aligned} \tag{24.28}$$

The first term denotes the case in which both borrowers succeed with their investment and each repays their own loan. The second term denotes the case where a borrower succeeds, but the other borrower fails and in addition to his own loan has to repay a fraction q of the loan of the other borrower. In case the borrower itself does not succeed, no repayments are required.

Inserting from equation (24.27) we thus obtain for companies investing into the low-risk and high-risk investments, respectively that

$$\begin{aligned} \Pi_C^H &= \pi_H (1 + R_H)\, \hat{L}_H - (1 + r_D)\, L, \\ \Pi_C^L &= \pi_L (1 + R_L)\, \hat{L}_L - \frac{\pi_L}{\pi_H} \frac{1 + (1 - \pi_L)\, q}{1 + (1 - \pi_H)\, q} (1 + r_D)\, L. \end{aligned} \tag{24.29}$$

We see that for companies choosing the low-risk investment the higher costs from covering another borrower's default is exactly offset by the lower loan rate charged by the bank, making his profits indistinguishable from conventional loans. This is not the case if the company chooses the high-risk loan.

We now easily obtain from the above that

$$\begin{aligned} \frac{\partial \Pi_C^H}{\partial q} &= 0, \\ \frac{\partial \Pi_C^H}{\partial L} &= \pi_H (1 + R_H) - (1 + r_D) > 0, \\ \frac{\partial \Pi_C^L}{\partial q} &= -\frac{\pi_L}{\pi_H} \frac{\pi_H - \pi_L}{(1 + (1 - \pi_H)\, q)^2} < 0 \\ \frac{\partial \Pi_C^L}{\partial L} &= \pi_L (1 + R_L) - \frac{\pi_L}{\pi_H} \frac{1 + (1 - \pi_L)\, q}{1 + (1 - \pi_H)\, q} (1 + r_D)\, . \end{aligned} \tag{24.30}$$

We now assume that $0 < \frac{\partial \Pi_C^H}{\partial L} < \frac{\partial \Pi_C^L}{\partial L}$ such that the marginal benefits to the borrower of the high-risk investment are higher than of the low-risk investment, if taking into account the financing costs.

As $\frac{\partial \Pi_C^H}{\partial L} > 0$, we see that companies would prefer the highest possible amount of loans. Higher loan amounts provide an incentive to switch to the more risky investment as can be seen from the assumption $\frac{\partial \Pi_C^L}{\partial L} > \frac{\partial \Pi_C^H}{\partial L}$ and hence banks need to limit the amount of loans accordingly such low-risk loans are chosen. We thus need that the profits of companies choosing the low-risk loan is higher than choosing the high-risk loan, $\Pi_C^H \geq \Pi_C^L$ and the highest possible loan amount is where $\Pi_C^H = \Pi_C^L$.

Looking at the difference of profits generated by the two different investments, we get that in order to maintain a constant difference in the profits between the two investments it is rfequired that

$$d\left(\Pi_C^H - \Pi_C^L\right) = \left(\frac{\partial \Pi_C^H}{\partial L} - \frac{\partial \Pi_C^L}{\partial L}\right) dL + \left(\frac{\partial \Pi_C^H}{\partial q} - \frac{\partial \Pi_C^L}{\partial q}\right) dq = 0. \tag{24.31}$$

From this we obtain the result that

$$\frac{dL}{dq} = -\frac{\frac{\partial \Pi_C^H}{\partial q} - \frac{\partial \Pi_C^L}{\partial q}}{\frac{\partial \Pi_C^H}{\partial L} - \frac{\partial \Pi_C^L}{\partial L}} > 0, \tag{24.32}$$

where we used equation (24.30) and our assumption that $\frac{\partial \Pi_C^L}{\partial L} > \frac{\partial \Pi_C^H}{\partial L}$. As companies seek the highest possible loan amount, companies want the highest possible value for q that is feasible; the lower loan rate associated with a higher joint liability will compensate for the costs of repaying the loan of the other company.

The highest feasible q is determined such that the resources of the company, after repaying their own loan, are sufficient to cover the additional liability. Thus we require that

$$qL \leq (1 + R_H)\,\hat{L}_H - (1 + r_L)\,L \tag{24.33}$$

which after inserting from equation (24.27) and solving the resulting equation becomes

$$q \leq \overline{q} = -\frac{L - (1-\pi_H)(1+R_H)\,\hat{L}_H}{2(1-\pi_H)\,L} + \sqrt{\left(\frac{L - (1-\pi_H)(1+R_H)\,\hat{L}_H}{2(1-\pi_H)\,L}\right)^2 + \frac{\Pi_C^H}{\pi_H(1-\pi_H)\,L}}. \tag{24.34}$$

From the requirement that $\Pi_C^H \leq \Pi_C^L$ and equation (24.29), we obtain the optimal loan size at the equality of these profits as

$$L = \frac{\pi_H(1+R_H)\,\overline{L}_H - \pi_L(1+R_L)\,\overline{L}_L}{\pi_H(1+R_H) - \pi_L(1+R_L) - \left(1 - \frac{\pi_L}{\pi_H}\frac{1+(1-\pi_L)q}{1-(1-\pi_H)q}\right)(1+r_D)}, \tag{24.35}$$

where we need $\pi_G(1+R_G)\,\overline{L}_G > \pi_B(1+R_B)\,\overline{L}_B$ for a positive loan amount.

Finally, companies will only participate if they make profits, thus we require $\Pi_C^H > 0$, or

$$L \geq \frac{\pi_H (1 + R_H) \overline{L}_H}{\pi_H (1 + R_H) - (1 + r_D)}. \tag{24.36}$$

Assuming that these two conditions are fulfilled, the amount of loans that companies can obtained are given by equation (24.35) and this loan amount is higher than would be possible without joint liability as indicated by equation (24.32). Without joint liability companies would not be able to secure loans of such size as companies would have an incentive to conduct the high-risk investment. The lower loan rate in case of joint liability provides incentives to choose the low-risk investment for larger loans.

We have thus established that joint liability will allow banks to provide larger loans to companies than would be feasible without joint liability. It might therefore be that joint liability is attractive to companies that face severe moral hazard constraints but have investment opportunities that require a larger loan than a bank would be willing to provide as a conventional loan.

Reading Stiglitz (1990)

24.4.2 Peer selection

Companies with a wide range of characteristics can be attracted to group lending and banks need to establish criteria how these companies are grouped together to establish joint liability. We assume here that the key characteristic of companies is the risks of their investments, defined by the probability of success an investment has, π_i. Let us assume for simplicity that there are only two types, low-risk companies with π_H and high-risk companies with π_L, where $\pi_L < \pi_H$. While companies know their characteristics and the characteristics of other companies, banks do not know the risk these companies are exposed to. The assumption that companies know each other's characteristics can be justified by the knowledge companies have of the market they operate in or through social networks.

As banks cannot distinguish between companies with different risk profiles, they will only be able to offer a single loan rate r_L to all companies. The joint liability is established through a company with a successful investment having to repay a fraction q of the loan of the other companies within their group. As banks cannot distinguish between companies, this parameter will be identical for all companies, regardless of their risk characteristics. For simplicity we assume that banks form groups of two companies, such that the way companies grouped comes down the either matching a high-risk (low-risk) company with another high-risk (low-risk) company or matching a high-risk and a low-risk company.

Taking into account that companies with different risk profiles might be grouped, the profits of company i being grouped with company j will, in line with equation (24.28), become

$$\begin{aligned}\Pi_C^{ij} &= \pi_i \pi_j \left((1 + R) L - (1 + r_L) L\right) \\ &\quad - \pi_i \left(1 - \pi_j\right) (1 + r_L) q L.\end{aligned} \tag{24.37}$$

For simplicity we here assume that the return on the investment, R, is identical across companies, regardless of the risks companies take.

A borrower of type i can choose to combine with a borrower of type H or L. We find that the difference in profits when grouping with these two types of companies to be

$$\begin{aligned}\Pi_C^{iH} - \Pi_C^{iL} &= \pi_i (\pi_H - \pi_L) ((1+R) L - (1+r_L)(1+q) L) \\ &> 0.\end{aligned} \tag{24.38}$$

Using the same assumption as in equation (24.33), the last term is positive and as $\pi_H > \pi_L$, the entire expression is positive. Hence companies of both types prefer to join with a low-risk borrower.

The benefits of a high-risk company being grouped with a low-risk company rather than a high-risk company are from equation (24.38) given by $\pi^L \left(\pi_H - \pi^L\right) (1+R) L - (1+r_L)(1+q) L$. This benefit is the maximum inducement the high-risk could pay to be grouped with a low-risk company. The benefits for a low-risk company being grouped with a low-risk company rather than a high-risk company are then similarly given by $\pi_H (\pi_H - \pi_L) (1+R) L - (1+r_L)(1+q) L$. As $\pi_H > \pi_L$, these benefits are higher than any payments the high-risk company could make.

Therefore we will find that low-risk borrowers join each other and then, similarly, high-risk borrowers join each other. This is because low-risk borrowers could provide additional payments to the bank for being grouped with another company they have identified as low-risk, and a high-risk company could not match the size of this payment; consequently the bank will allocate the low-risk company to another low-risk company. This then leaves high-risk companies to be grouped together as they have no other partner.

We thus find that companies of similar risks tend to group together, which will also simplify the risk assessment for banks as they can pool the information they hold on companies in a group and knowing they all will have similar risk characteristics are more easily able to make a correct judgement on the risks of companies.

Reading Ghatak (1999)

24.4.3 Optimal group size

When borrowing with joint liability, the size of the group can be important. A larger group size will allow any responsibilities from a group member not repaying their loan to be spread wider, reducing the costs; on the other hand, it also will increase the number of group members that will not be able to repay their loans, providing a trade-off of these two effects.

Let us assume that the loan rate r_L is fixed but lenders instead compete on the size of the group they lend to. Borrower have identical risks and investments succeed with probability π, yielding a return of R in that case. Thus a successful borrower can repay at most $(1+R) L$ to cover the default of other borrowers. If there are M borrowers failing, and hence $N-m$ borrowers with successful investments, then the most a bank can obtain is $(N - M)(1 + R) L$, where N is the number of borrowers in the group and L the loan amount. Of course, the bank is only entitled to $N(1 + r_L) L$ of payments at most. The probability of M borrowers defaulting is given by $\binom{N}{M}\pi^{N-M}(1-\pi)^M$ and hence the bank profits are given by

$$\begin{aligned}\Pi_B = &\sum_{M=0}^{N}\binom{N}{M}\pi^{N-M}(1-\pi)^M \\ &\times\min\{N(1+r_L)L;(N-M)(1+R)\}L \\ &-(1+r_D)L,\end{aligned} \tag{24.39}$$

with r_D denoting the deposit rate that applies to the deposits fully financing the loans.

Borrowers will only repay their loan if they are successful, which happens with probability π, and then obtain a profit from the investment of $(R - r_L) L$, after repaying the loan. From this amount we have to deduct their share of the failing repayments they need to cover, up to their surplus, $R - r_L$. The share to be covered is M failures with $N - M$ borrowers providing this coverage. We thus have

$$\begin{aligned}\Pi_C = \pi\Bigg((R-r_L) - &\sum_{M=0}^{N}\binom{N}{M}\pi^{N-M}(1-\pi)^M \\ &\times\min\left\{\frac{M}{N-M}(1+r_L);R-r_L\right\}\Bigg)L.\end{aligned} \tag{24.40}$$

Borrowers taking out individual loans would get a profit of

$$\hat{\Pi}_C = \pi(R - \hat{r}_L) L, \tag{24.41}$$

at a different interest rate $\hat{r}_L$. Banks and group lenders competing would ensure that $\Pi_C = \hat{\Pi}_C$. A bank lending individually would get profits of

$$\begin{aligned}\hat{\Pi}_B &= \pi(1+\hat{r}_L)L - (1+r_D)L \\ &= \pi(1+R)L - (1+r_D)L - \hat{\Pi}_C,\end{aligned} \tag{24.42}$$

where we solved equation (24.41) for $\hat{r}_L$ and inserted this into equation (24.42). Using equations (24.39), (24.40), and (24.42), we can now show that

$$\frac{\Pi_B}{N} = \hat{\Pi}_B. \tag{24.43}$$

Thus banks make the same profits per loan in group lending as in lending individually. Therefore, both lending forms can co-exist in the market.

Adding an additional group member causes the banks additional losses from default of $(1 - \pi)(1 + r_L) L$ but also increases benefits by $\pi(R - r_L) L$ from the increased cover of losses by this additional borrower. If $\pi > \frac{1+r_L}{1+R}$, then these benefits exceed the costs and hence $\frac{\partial \Pi_B}{\partial N} > 0$. This condition is assumed to be fulfilled if we assume that lending is socially desirable as it implies that the expected return of the investment exceeds the financing costs in form of the loan rate. Using equation (24.43) and inserting inserting (24.42), it is obvious that in this case $\frac{\partial \Pi_C}{\partial N} < 0$. Thus banks prefer larger groups and borrowers prefer smaller groups, leading to a conflict of interest in the optimal group size.

If banks are competitive, we have $\Pi_B = N\hat{P}i_B = 0$ and hence from equation (24.42) we easily obtain the loan rate for individual loans as

$$1 + \hat{r}_L = \frac{1 + r_D}{\pi}. \tag{24.44}$$

As borrowers will only participate if $\Pi_C = \hat{\Pi}_C \geq 0$, we get from equation (24.41), after inserting equation (24.44), that we require

$$\pi \geq \pi^* = \frac{1 + r_D}{1 + R}. \tag{24.45}$$

Not surprisingly, loans are only available to borrowers that are not too risky.

Using equation (24.43) it is now easy to see that

$$\frac{\partial \Pi_B}{\partial \pi} = \frac{\partial N \hat{\Pi}_B}{\partial \pi} = N\,(1 + R)\,L > 0. \tag{24.46}$$

Using that with perfect competition $\Pi_B = 0$ and the implicit function theorem we get with our above result that $\frac{\partial \Pi_B}{\partial N} > 0$ that

$$\frac{\partial N}{\partial \pi} < 0. \tag{24.47}$$

We thus see that the optimal group size is larger, the more risky the loans are. For more risky loans, the chances of a borrower being required to repay the loans of other group members is low as the probability of a successful investment is low, making the additional costs from a larger group size small and thus supporting larger groups.

Reading Krause (2022b)

24.4.4 Reducing moral hazard

Companies can often control the level of risks they are taking, although reducing risks (increasing the probability of success of an investment) will normally be costly. Such costs might include additional risk management, the employment of better managers demanding a higher salary, changing working practices, and many other measures companies can take. In group lending companies are not only exposed to the risks they themselves take, but also the risks the other companies take. This will affect the incentives to reduce risks and can affect the moral hazard from such lending arrangements.

Let us assume that companies can control the risks their investment takes. They make investments that are successful with probability π_i when obtaining a loan L, which yields a return of R if successful; the bank charges interest r_L and the company can control the probability of success with costs c. If the bank finances the loan fully with deposits that attract interest r_D, the social welfare is then given by

$$\Pi_W = \pi_i (1 + R) L - (1 + r_D) L - \frac{1}{2} c \pi_i^2 L, \tag{24.48}$$

such that the socially optimal level of risk is given from the first-order condition

$$\frac{\partial \Pi_W}{\partial \pi_i} = (1 + R) L - c \pi_i L = 0 \tag{24.49}$$

as

$$\pi^* = \frac{1 + R}{c}. \tag{24.50}$$

In traditional lending, the company profits are given by

$$\Pi_C = \pi_i ((1 + R) L - (1 + r_L) L) - \frac{1}{2} c \pi_i L, \tag{24.51}$$

with the assumption that the loan is either repaid with probability π_i, or no payment is made. The first order condition for the optimal risk level is given from the first-order condition

$$\frac{\partial \Pi_C}{\partial \pi_i} = ((1 + R) L - (1 + r_L) L) - c \pi_i L = 0. \tag{24.52}$$

The bank profits are given by

$$\Pi_B = \pi_i (1 + r_L) L - (1 + r_D) L, \tag{24.53}$$

where we assume that deposits are always repaid. In competitive markets banks make no profits, thus $\Pi_B = 0$. Solving equation (24.52) for $1 + r_L$ and inserting for this into equation (24.53), we easily obtain

$$c \pi_i^2 - \pi_i (1 + R) + (1 + r_D) = 0 \tag{24.54}$$

which solves for the optimal level of risk companies choose:

$$\pi_i^{**} = \frac{(1 + R) + \sqrt{(1 + R)^2 - 4c (1 + r_D)}}{2c} < \pi_i^*. \tag{24.55}$$

We only consider the larger root of equation (24.54) as this gives the lower risk. The success rate that companies choose is below the social optimum and hence companies choose investments that are too risky; hence in conventional lending the moral hazard of companies having to exert effort to reduce risks causes the risk to be too high.

We now introduce joint liability for loans by making loans the responsibility of two borrowers that form a group. If the other borrower defaults, a borrower has to repay a fraction λ of the original loan. This can only happen if the borrower does not itself default. The other borrower's risk is denoted π_j, such that the company's profits are given by

$$\Pi_C = \pi_i \left((1+R) L - (1+r_L) L \right) - \pi_i \left(1 - \pi_j \right) \lambda L - \frac{1}{2} c \pi_i^2 L. \tag{24.56}$$

The first term denotes the repayment of the loan the company as obtained and the second term the repayment of the loan of the other company, provided that company defaults and the company itself has been successful; the final term denotes the costs of companies increasing the success rate of their own investments.

The first order condition for the optimal choice of the risk is

$$\frac{\partial \Pi_C}{\partial \pi_i} = (1+R) L - (1+r_L) L - \lambda \left(1 - \pi_j \right) L - c \pi_i L = 0. \tag{24.57}$$

The bank profits are then given by

$$\Pi_B^i = \pi_i (1+r_L) L + \pi_i \left(1 - \pi_j \right) \lambda L - (1+r_D) L, \tag{24.58}$$

where the first term denotes the repayment of the loan by the company that has received it and the second term captures the repayment of the loan by the other group member.

As both borrowers are ex-ante identical, we need in equilibrium that the risks they are choosing are the same, $\pi_i = \pi_j$. Using this requirement in equation (24.57) and solving for the loan rate, $1 + r_L$, to insert into equation (24.58), using here again that $\pi_i = \pi_j$ and as before that $\Pi_B = 0$, we get

$$c \pi_i^2 - \pi (1+R) + (1+r_D) = 0. \tag{24.59}$$

This expression is identical to equation (24.54) and hence joint liability does not affect risk-taking.

Companies have an interest in the risk the other company in a joint liability group takes. Let us therefore now assume that borrowers can monitor each other's risk-taking behaviour such that we can ensure $\pi_i = \pi_j$ from the start and not only as an equilibrium condition. In this case, the profits of the company are given by

$$\Pi_C = \pi_i \left((1+R) L - (1+r_L) L \right) - \pi_i (1 - \pi_i) \lambda L - \frac{1}{2} c \pi_i^2 L \tag{24.60}$$

and the first order condition for the optimal level of risk becomes

$$\frac{\partial \Pi_C}{\partial \pi_i} = (1+R) L - (1+r_L) L - \lambda (1 - 2\pi_i) L - c \pi_i L = 0. \tag{24.61}$$

The bank profits are then easily obtained as

$$\Pi_B = \pi_i \left(1 + r_L\right) L + \pi_i \left(1 - \pi_i\right) \lambda L - \left(1 + r_D\right) L. \tag{24.62}$$

Solving again equation (24.61) for the loan rate $1 + r_L$ and inserting this into equation (24.62), we obtain after using $\Pi_B = 0$ that

$$(c - \lambda) \pi_i^2 - \pi_i \left(1 + R\right) + \left(1 + r_D\right) = 0, \tag{24.63}$$

which solves for

$$\pi_i^{***} = \frac{(1 + R) + \sqrt{(1 + R)^2 - 4 (c - \lambda) (1 + r_D)}}{2 (c - \lambda)} \geq \pi_i^{**} \tag{24.64}$$

The indicated inequality applies if we assume that $c > \lambda$ as then the denominator in equation (24.64) is smaller than in equation (24.55), making the expression larger. Furthermore the numerator increases as $c - \lambda < c$ and hence a smaller amount is deducted.

It thus that with joint liability the risks the company takes will be smaller as long as the company can ensure through monitoring that the other company in the group takes the same level of risk as themselves. The requirement for monitoring the other company and ensuring they take the same risks arises because this will eliminate an externality. A company taking a higher risk, exposes the other company in their group to a higher risk of having to cover their default; this increased cost of success to the company will induce them to take a higher risk than they would otherwise do. If monitoring each other can ensure that the risks taken are identical for both companies, this externality gets fully internalized as the company know that if it will increase risks and impose costs on the other company, the other company will do so likewise.

We can derive that $\pi_i^{***} < \pi_i^{*}$ if $(1 + R)^2 < \frac{c^2}{\lambda} (1 + r_D)$. Hence if the returns on the project are not too big, the social optimum is not achieved, but the risk taken is smaller than with individual loans and approaches the social optimum more closely. For very profitable projects not enough risk would be taken, $(1 + R)^2 > \frac{c^2}{\lambda} (1 + r_D)$, the risk taken would be less than is socially optimal.

We thus see that joint lending can reduce the moral hazard of borrowers, but only if they can monitor each other to ensure that the other company does not increase the risks of their investments. Without the ability to monitor each other, the risks taken are not different between individual and group lending.

Reading Ghatak & Guinnane (1999)

24.4.5 Eliminating asymmetric information

In lending, banks are often less well informed than the companies themselves, who might well have much knowledge than banks of the risks their investments involve. Similarly, companies might have similarly good information about other companies, in particular those in similar businesses. We will see how in group lending, banks can extract information about the riskiness of companies from their decision to prefer group lending over a traditional loan.

Let us assume that banks do not know the riskiness of their borrowers, but borrowers are able to evaluate each other perfectly. We assume there are two types of borrowers, one with a high repayment rate π_H and one with a low repayment rate for loans, $\pi_L < \pi_H$. It is known that a fraction p of borrowers have a high probability repaying the loan, while a fraction $1 - p$ have a low probability of making such repayments. With a loan rate r_L, deposit rate r_D and a loan size of L, fully funded by deposits, we get the bank profits for a traditional loan as

$$\Pi_B = (p\pi_H + (1-p)\,\pi_L)\,(1+r_L)\,L - (1+r_D)\,L. \tag{24.65}$$

Because the bank does not know the type of borrower, it will only be able to use the average probability of the loan being repaid, $p\pi_H + (1-p)\,\pi_L$. In a competitive market profits are eliminated such that $\Pi_B = 0$ and hence

$$1 + r_L = \frac{1+r_D}{p\pi_H + (1-p)\,\pi_L}. \tag{24.66}$$

Depending on thetype of borrower, he will make profits of

$$\Pi_C^i = \pi_i\,((1+R)\,L - (1+r_L)\,L)\,, \tag{24.67}$$

where R denotes the return on the investment and we assume that if the investment is not successful no loan repayments are required due to limited liability.

Let us now assume there is a contract that allows for joint liability of two borrowers. If one borrower defaults, the other borrower has to repay a fraction λ of the original loan, in addition to the repayment of its own loan. Thus, for two borrowers of types i and j to form such a group, the profits to borrower i are given by

$$\begin{aligned} \Pi_C^{ij} &= \pi_i\pi_j\,((1+R)\,L - (1+\hat{r}_L)\,L) \\ &\quad + \pi_i\left(1-\pi_j\right)((1+R)\,L - (1+\hat{r}_L)\,L - \lambda L)\,, \end{aligned} \tag{24.68}$$

where $\hat{r}_L$ denotes the loan rate in this case. Here the first term covers the case that both borrowers can repay the loan and the second term that borrower j is not able to do so and thus borrower i repays a fraction of the original loan obtained by the other borrower.

Assume now we determine the loan rate in group lending from

$$1 + \hat{r}_L = 1 + r_L - (1-\pi_L)\,\lambda + \gamma\lambda, \tag{24.69}$$

where γ is some parameter. Inserting equation (24.69) into equation (24.68) and noting equation (24.67), we get the profits of the different combinations of borrowers as

$$\begin{aligned}\Pi_C^{HH} &= \Pi_C^H + \pi_H(\pi_H - \pi_L - \gamma)\lambda L, \\ \Pi_C^{LL} &= \Pi_C^L - \pi_L\gamma\lambda L, \\ \Pi_C^{HL} &= \pi_C^H - \pi_H\gamma\lambda L.\end{aligned} \tag{24.70}$$

We now can easily obtain that

$$\begin{aligned}\frac{\partial \Pi_C^{HH}}{\partial \lambda} &= \pi_H(\pi_H - \pi_L - \gamma)L, \\ \frac{\partial \Pi_C^{LL}}{\partial \lambda} &= -\pi_H\gamma L < 0, \\ \frac{\partial \Pi_C^{HL}}{\partial \lambda} &= -\pi_H\gamma L < 0.\end{aligned} \tag{24.71}$$

We see immediately that for $\gamma > 0$ the combination of two high-risk borrowers, L, will prefer no joint liability as the optimal fraction of repayment from the other borrowers loan will be $\lambda = 0$, given the marginal profits are always negative. The same applies if a low-risk borrower combines with a high-risk borrower. If $\gamma < \pi_H - \pi_L$, the two low-risk borrowers H would want the highest possible joint liability as the marginal profits are strictly positive, thus $\lambda = 1$.

We also see immediately that $\Pi_C^{HH} > \Pi_C^{HL}$ and hence low-risk borrowers will not accept high-risk borrowers into their group, leading to a complete separation into groups of high-risk and low-risk borrowers. Furthermore, as $\Pi_C^{HH} > \Pi_C^H$ for $\gamma \leq \pi_H - \pi_L$, we see that low-risk borrowers will want to engage in joint-liability lending, while $\Pi_C^{LL} < \Pi_C^L$ and high-risk borrowers would prefer individual loans. Hence by setting $0 < \gamma < \pi_H - \pi_L$ the bank can separate the two types of borrowers; a low-risk borrower will seek a group loan with other low-risk borrowers and high-risk borrowers will seek individual loans.

We can determine the bank profits from lending to a group of two low-risk borrowers and obtain

$$\begin{aligned}\Pi_B^{HH} &= 2\pi_H^2(1 + \hat{r}_L)L + 2\pi_H(1 - \pi_H)((1 + \hat{r}_L)L + \lambda L) \\ &\quad -2(1 + r_D) \\ &= 2(\pi_H(1 + r_L)L - (1 + r_D)L) \\ &\quad -2\pi_H(\pi_H - \pi_L - \gamma)\lambda L,\end{aligned} \tag{24.72}$$

where the second equality has been obtained by inserting from equation (24.69). The first term denotes the repayment of both loans as both borrowers' investments have been successful, and the second term denotes the repayments if only one of the borrowers' investments was successful and this borrower also covers some of the repayments of the loan of the unsuccessful borrower.

For banks to offer this joint liability contract, we need these profits to exceed that of an individual loan, $\Pi_B^{HH} > 2\Pi_B^H = 0$, which solves for

$$\gamma > \gamma^* = \frac{\pi_H (\pi_H - \pi_L) \lambda - (\pi_H (1 + r_L) - (1 + r_D))}{\pi_H \lambda}. \tag{24.73}$$

With the definition of $1 + r_L$ in equation (24.69), we can easily confirm that $\gamma^* < \pi_H - \pi_L$. Competition between banks resulting in $\Pi_B^{HH} = 0$, will lead to γ^* being chosen accordingly.

Inserting equation (24.73) into equation (24.72), we see that the banks profits in joint lending, Π_B^{HH}, does not depend on the degree of joint liability λ; any combination of γ and λ fulfilling equation (24.73) can be used.

The single liability contract for high-risk borrowers would be structured such that the profits to these borrowers are identical to the profits from group lending by choosing an appropriate loan rate $\hat{\hat{r}}_L$, hence $\Pi_B^{LL} = \pi_L \left(\left(1 + \hat{\hat{r}}_L \right) - (1 + r_D) L \right)$. From this we obtain that $1 + \hat{\hat{r}}_L = \frac{1+r_D}{\pi_L}$. Even though high-risk borrowers are getting lower profits due to the higher loan rate, they cannot join to joint liability lending as they would not be accepted by low-risk borrowers.

We have seen that banks could use group lending contracts to identify high-risk and low-risk borrowers, with low-risk borrowers preferring group lending contracts and high-risk borrowers taking out individual loans. This separation of low-risk and high-risk borrowers can be compared to a similar separation using collateral as discussed in chapter 9.2.1. Here the potential loss from repaying the loan of the other borrower has the same impact as the loss of collateral in case of default; the loan rate is also lower here due to this additional payment the bank receives from the other borrower in case of default. Thus the commitment to cover the loss of the other borrower can be interpreted as a collateral and only low-risk borrowers are willing to provide this.

Reading Tassel (1999)

24.4.6 Avoiding strategic default

Companies may have incentives to default strategically is discussed in chapter 7. A common argument to prevent strategic default is that companies forego future profits by defaulting and if these future profits outweigh the benefits from defaulting, companies will not do so voluntarily. Alternatively, banks might be monitoring their lenders and thereby prevent strategic defaults by identifying them as such and charging a fine, in addition to the repayment of the loan. We will here investigate how in group lending incentives exist for companies to monitor each other and through such monitoring prevent strategic default at no additional cost to the bank.

Let us assume that a company not defaulting benefits from future lending opportunities to the value of V; such future benefits arise as the company can continue to obtain loans and conduct investments, which after a default is no longer possible. Companies could strategically default on their loan L, on which interest r_L is payable and the investment returns R, if the profits generated after repaying the loan and obtaining future benefits are less than retaining the proceeds of the investment and not repaying the loan as well as forgoing any future benefits from further loans.

Thus for a strategic default we require

$$(1+R)\,L-(1+r_L)\,L+V \leq (1+R)\,L.$$

Banks will charge the highest possible loan rate such that this relationship is fulfilled with equality, just preventing strategic default. Thus we have the loan rate given as

$$(1+\hat{r}_L)\,L = V. \tag{24.74}$$

As companies are successful with probability π, the expected profits of the company and bank, respectively, are then given by

$$\begin{aligned}\hat{\Pi}_C &= \pi\left((1+R)-(1+\hat{r}_L)\,L+V\right)=\pi\,(1+R)\,L, \qquad (24.75)\\ \hat{\Pi}_B &= \pi\,(1+\hat{r}_L)\,L-(1+r_D)\,L=\pi V-(1+r_D)\,L,\end{aligned}$$

where r_D denotes the deposit rate on deposits, which finance the loan fully.

If two companies are now jointly liable for a loan of L to each of them, they might monitor each other and if monitoring happens, strategic default cannot occur; assume such monitoring happens with probability p. We can also interpret this probability as the likelihood with which startegic default can be identified as such by the monitoring company. If a borrower seeks to strategically default but this is detected through monitoring, then there is a private cost W in doing so; such costs might include a loss of reputation in or companies being more reluctant to join a group with this company in future group lending. The joint liability states that if a borrower is successful in its project, it has to repay a fraction λ of the other borrower's loan if that borrower defaults. The cost of monitoring for strategic default is c; as in joint liability the originator of a default cannot clearly be identified to outsiders, we assume that banks exclude borrowers involved in a group that experiences default only with probability β and thus deprive them of their future benefits V.

If a borrowers is successful with his project, it will not strategically default if

$$\begin{aligned}&(1+R)\,L-\pi\,(1+r_L)\,L-(1-\pi)\,(1+\lambda)\,(1+r_L)\,L+V\\ &\quad\geq (1+R)\,L+\pi V+(1-\pi)\,(1-\beta)\,V-pW.\end{aligned}$$

Here the left-hand side shows the payments if the borrower does not default. He will receive the return on the project, repay only its own loan if the other borrower is also successful (π), or repay his and a fraction λ of the other borrower's loan $(1-\pi)$ if he fails, but retains the future value V, The pay-off from strategic default on the right-hand side show the retained project value, the retention of future benefits if the other borrower succeeds and pays off his loan and if he does not succeed, the possible loss of this value, $(1-\pi)\,(1-\beta)$, less the private costs if monitored. This expression can easily be solved for

$$p \geq p^* = \frac{(1+\lambda\,(1-\pi))\,(1+r_L)\,L-(1-\pi)\,\beta V}{W}. \tag{24.76}$$

For $p \geq p^*$ no strategic default happens and as monitoring is costly we set $p = p^*$.

This relationship gives us the benefits of being monitored, but we also need to ensure monitoring at level p^* happens. The benefits of monitoring need to exceed those of not monitoring for strategic default, thus we require that

$$\begin{aligned}&\pi^2\lambda\left(1+r_L\right)L+\left(1-\left(1-\pi\right)^2\right)V+\left(1-\pi\right)^2\left(1-\beta\right)V-cp^*\\&\quad\geq \pi V+\left(1-\pi\right)\left(1-\beta\right)V.\end{aligned}$$

If monitoring occurs and both borrowers succeed (π^2), the borrower saves his fraction λ of the repayment as no strategic default happens. If both borrowers do not fail and therefore repay their loans, $1-(1-\pi)^2$, the future benefits are retained. It is also retained if both fail, but only with probability $1-\beta$. The costs of monitoring c reduce these benefits. If no monitoring happens, the value of future borrowing is retained if either the borrower repays the loan himself of if he is not successful, but retains it anyway. We can rewrite this condition as

$$\pi^2\lambda\left(1+r_L\right)L+\pi\left(1-\pi\right)\beta V-cp^*\geq 0. \tag{24.77}$$

This restriction is least strong if $\beta = 1$ as then monitoring will happen more easily and also p^* is lower as we see from equation (24.76); hence banks would set $\beta = 1$ and if a joint loan fails, both borrowers lose future benefits for sure.

Bank profits are given as

$$\Pi_B=\left(2\pi^2+2\pi\left(1-\pi\right)\left(1+\lambda\right)\right)\left(1+r_L\right)L-2\left(1+r_d\right)L. \tag{24.78}$$

The bank is repaid if either both repay their loans or one borrower repays the loan for both borrowers. In the former case the bank obtains two repayments and in the latter $1 + \lambda$ repayments. Maximizing the bank profits in equation (24.78) subject to equation (24.77) as an equality, with ξ denoting the Lagrange multiplier, we get the first order condition as

$$\begin{aligned}\frac{\partial\Pi_B}{\partial\lambda}&=2\pi\left(1-\pi\right)\left(1+r_L\right)L\\&\quad+\xi\left(\pi^2\left(1+r_L\right)L-\frac{c}{W}\left(1-\pi\right)\left(1+r_L\right)L\right)\\&=0,\end{aligned} \tag{24.79}$$

where we inserted for p^* from equation (24.76). We can rewrite equation (24.79) as

$$2\pi\left(1-\pi\right)+\xi\left(\pi^2-\frac{c}{W}\left(1-\pi\right)\right)=0.$$

If we assume

$$\frac{c}{W}\leq\frac{\pi^2}{1-\pi}, \tag{24.80}$$

then both terms are positive and as $\xi \geq 0$, the only solution is to set $\lambda = 1$, such that a successful borrower will fully repay the loan of the unsuccessful borrower. This then implies that equation (24.77) is binding and using $\lambda = \beta = 1$, we can solve this expression for the loan rate, which is determined as

$$(1 + r_L) L = (1 - \pi) \frac{\frac{c}{W} + \pi}{\frac{c}{W}(2 - \pi) - \pi^2} V. \tag{24.81}$$

Inserting these results, we get the bank profits as

$$\Pi_B = 2(2 - \pi)(1 - \pi) \pi \frac{\frac{c}{W} + \pi}{\frac{c}{W}(2 - \pi) - \pi^2} V - 2(1 + r_D) L. \tag{24.82}$$

Banks would offer group lending contracts if it is more profitable than individual contracts, thus we require $\Pi_B \geq 2\hat{\Pi}_B$, which from equation (24.75) becomes

$$\frac{c}{W} \leq \frac{1 + (1 - \pi)^2}{2 - \pi}. \tag{24.83}$$

The constraint in equation (24.80) on the optimality of full joint liability for the group loan is less strict than this constraint, thus if group lending is optimal for banks, it will encompass full liability of the borrowers or each other's loans.

Similarly, we get the company profits from group lending as

$$\begin{aligned} \Pi_C = & \pi^2 ((1 + R) L - (1 + r_L) L) \\ & + \pi (1 - \pi) ((1 + R) L - 2(1 + r_L) L) \\ & + \left(1 - (1 - \pi)^2\right) V - cp^*. \end{aligned} \tag{24.84}$$

The first term denotes the repayment of the loan if both borrowers are successful, the second term the fact that if the other borrower is not successful, both loans need to be repaid by this borrower. Unless both borrowers defaults, the future value is retained as the loan is repaid. In addition, the monitoring costs need to be covered.

The company seeks out joint loans if the profits are higher than taking out an individual loan, $\Pi_C \geq \hat{\Pi}_C$, which when inserting from equations (24.75), (24.76), and (24.81), becomes

$$\frac{c}{W} \geq 2 - \pi. \tag{24.85}$$

The condition in equation (24.80) that makes full liability for each other's loans optimal is consistent with the constraint of companies seeking group lending if

$$\pi \geq \frac{2}{3}. \tag{24.86}$$

The conditions for banks preferring group lending, equation (24.83), and the condition for companies to prefer bank lending, equation (24.85), can never be fulfilled together. Hence here companies would not be able to agree with banks on group lending if simple contracts are on offer; if banks prefer group lending, companies would prefer individual loans and if companies were to prefer group lending, banks would prefer individual loans.

Thus, borrowers monitoring each other in a group lending can effectively prevent strategic defaults by borrowers as it is in the interest of each borrower that the other does not strategically default, causing the other group member to become responsible for the repayment of both loans. While such an arrangement might be optimal for companies, it would not be optimal for banks and if it is optimal for banks, it would not be optimal for companies. Thus the ability of companies to engage in group lending if this is beneficial to them would depend on banks being required to offer such loans.

Reading de Aghion (1999)

Résumé

In group lending a number of borrowers are jointly responsible for the repayment of all loans granted to the members of that group. We have seen that such arrangements can have benefits to banks and borrowers alike. It allows borrowers to obtain larger loans than they would be able to secure when taking out individual loans due to the reduced moral hazard in group lending. But not only does group lending to allow moral hazard, it also allows banks to identify the risks of borrowers more easily if borrowers know each other well. Borrowers taking similar risks will form groups and it is even that group lending acts similarly to collateral in that the choice between individual and group lending can reveal the risk type of the borrowers; low-risk borrowers will prefer group loans and high-risk borrowers will prefer individual loans. If borrowers are able to monitor each other, it will even be possible to avoid strategic default without any costs to the bank.

Group lending is a prominent lending form only in developing countries for small loans to individual entrepreneurs, often also known as ’microfinance’. This limitation to developing markets might be the result of severe moral hazard and asymmetric information between banks and borrowers; the lack of experience in monitoring borrowers and assessing risks combined with strong incentives to use loan proceeds to better their immediate living standards rather than investing into their business, makes the provision of bank loans difficult. Group lending allows the mutual monitoring within groups, who often have strong social connections, and despite the absence of collateral allows borrowers to signal to banks their low-risk investments by deciding to take up a group loan rather than seeking an individual loan.

Conclusions

Banks are not only competing against each other, but also against other financial institutions offering loans. We have discussed some of such alternative forms of loan provision and seen that such alternative lenders might be attractive to some borrowers. In particular, peer-to-peer lending, a form of loans that originates directly from ’depositors’ without the use of banks, was attractive to high-risk borrowers, while low-risk borrowers found traditional bank lending more attractive. Similarly, borrowers with higher risks were able to secure loans from the sellers of goods they seek to purchase with the loan in what is commonly referred to as ’consumer finance’. Group lending, often known as ’microfinance’, where a number of borrowers are jointly responsible for the loans to all group members, can be attractive to low-risk borrowers, and can be found particularly in situations where lending is subject to significant moral hazard and asymmetric information. Finally, in Islamic banking conventional interest cannot be charged and the total return afforded to borrowers and depositors are such that borrowing would be more costly but deposits would attract a higher return.

Such alternative forms of providing loans are on the one hand a competition to banks as they will react to the possibility of borrowers seeking loans elsewhere, but on the other hand the existence of such alternatives may also allow banks distinguish between borrowers they would want to provide loans to, such as low-risk borrowers, and those they do not want to provide loans to, high-risk borrowers; without alternative lenders banks might not be able to distinguish between these types of borrowers and thus be exposed to adverse selection. Being able to distinguish borrower types, or finding alternative mechanisms to reduce moral hazard, will improve the social welfare of the economy as all borrowers will obtain loans at conditions that are suitable to them.

Chapter 25

Risk-taking behaviour

Economic analysis of the impact of competition commonly focusses on the price of the goods, in the context of banking this would mainly be the loan and deposit rates, as well as the quantities of goods supplied and demanded, the amount of loans and deposits. We will here consider whether competition has an influence on the risk-taking by banks. With increasing competition, the profit margins of banks will reduce and this might induce them to provide loans to companies that pursue more risky investments as this would allow them to charge higher loan rates and hence enhance their profitability. In chapter 25.1 we will explore this effect of competition on the risk of individual banks loans that are provided by banks. With banks not only providing a single loan but a portfolio of such loans, the correlations between defaults will become important as we will investigate in chapter 25.2.

It is important to assess this impact of competition between banks ensure any assessment of competition policy does not only consider the impact on loan and deposit rates, and thus how borrowers and depositors are directly affected, but also the implications on the risk of banks. This aspect is central to assess any potential social costs of competition policy, such as an increase, or a reduction, of the risks banks take. But it will also be important to note the risks that banks are willing to take with each loan as this might give an indication of the type of investments that banks are willing to finance. Such loan provision can have significant impact on the future growth prospects on an economy and fir this reason should be included in any full assessment of competition policy.

25.1 The effect on loan risks

The main focus when analysing the effect of competition between banks is the influence on loan and deposit rates. However, banks have another decision variable whose optimality might change as competition changes, namely the risks of the loans they provide. Using the common assumption that higher risks result in higher returns, it follows that banks might compensate for lower loan rates by taking higher risks that increase their profitability in the face of increased competition. In this section we will investigate the effect such competition has on the risk-taking of banks.

We will specifically look in chapter 25.1.1 on the consequences of increased competition from ever more banks entering the market if banks are purchasing deposit insurance. In contrast to that, chapter 25.1.2 assesses the effect on the risk-taking of banks if competition requires them to reduce their loan rates in a situation where both adverse selection and moral hazard are present. While the risks are taken when providing loans, competition in deposit markets can nevertheless affect the risks optimally taken in the loan market as chapter 25.1.3 will show.

25.1.1 Risk-taking with deposit insurance

In most cases, banks have limited liability and thus generate value to their owners only if they do no fail. We will explore here how a bank seeking to maximize its value in light of ever diminishing market shares due to competition from more banks chooses the optimal level of risk for the loans they provide.

A bank gives a loan L at interest r_L, that the borrower repays with probability π. This probability of repayment will be reducing the higher the loan rate is, $\frac{\partial \pi}{\partial(1+r_L)} < 0$, as higher loan rates require companies to make more risky investments in order to generate the returns needed to repay their loan. If we assume that the bank has no equity, it will fail if the loan is not repaid. We furthermore assume that loans are fully financed through deposits D that attract interest r_D; hence we have $L = D$. In addition, depositors are protected by a deposit insurance that pays any shortfalls on their claim against the bank if the bank fails. To finance these payments, banks pay a premium P for the deposit insurance. Thus the bank profits are given by

$$\begin{aligned} \Pi_B &= \pi \left((1 + r_L) L - (1 + r_D) D - P \right) \qquad (25.1) \\ &= \pi \left(((1 + r_L) - (1 + r_D)) D - P \right). \end{aligned}$$

Pre-determined deposit insurance premium The first order conditions for a profit maximum by choosing the loan and deposit rates optimally, with a given deposit insurance premium P, is obtained as

$$\begin{aligned} \frac{\partial \Pi_B}{\partial (1 + r_L)} &= \frac{\partial \pi}{\partial (1 + r_L)} \left(((1 + r_L) - (1 + r_D)) D - P \right) + \pi D \qquad (25.2) \\ &= 0, \\ \frac{\partial \Pi_B}{\partial (1 + r_D)} &= \pi \left(((1 + r_L) - (1 + r_D)) \frac{\partial D}{\partial (1 + r_D)} \right. \\ &\qquad \left. -D - \frac{\partial P}{\partial (1 + r_D)} \right) \\ &= 0. \end{aligned}$$

The deposit insurance premium is set fairly such that the expected shortfall in funds if the loan defaults, $\pi \left((1 + r_D) D + P \right)$ equals the expected payout the deposit insurance needs to make, $(1+r_D)\,D$. Hence

$$\pi\left((1+r_D)\,D+P\right)=(1+r_D)\,D,$$

which solves for the deposit insurance premium to be set at

$$P=\frac{1-\pi}{\pi}\,(1+r_D)\,D. \tag{25.3}$$

We assume now that the deposit insurer announces the premium $\frac{P}{D}$ before the bank makes lending decisions, anticipating the choice of risk, π, and the deposit rate, $1+r_D$, by the bank. We can thus treat the deposit insurance premium as fixed. In this case we have $\frac{\partial P}{\partial D}=\frac{1-\pi}{\pi}\,(1+r_D)$. Inserting these relationship, we can rewrite (25.2) as

$$\begin{aligned}\frac{\partial\pi}{\partial\,(1+r_L)}\left((1+r_L)-\frac{1+r_D}{\pi}\right)+\pi&=0 \\ \left((1+r_L)-\frac{1+r_D}{\pi}\right)\frac{\partial D}{\partial\,(1+r_D)}-D&=0\end{aligned} \tag{25.4}$$

where we used $\frac{\partial P}{\partial(1+r_D)}=\frac{\partial P}{\partial D}\frac{\partial D}{\partial(1+r_D)}$.

Let us now assume that the amount of deposits in an economy is given and the degree of competition between banks can be assessed through the number of banks active in the market. With a given amount of deposits, and hence loans, being available in the market, the deposits D that each banks obtains reduces as the number of banks increases. If we approach perfect competition with an infinite number of banks, the size of deposits in each banks will approach zero. From the second line in equation (25.4), we see that this requires $1+r_L-\frac{1+r_D}{\pi}=0$. Inserting this result into the first line, we obtain that $\pi=0$. Hence as competition increases, modelled by increasing the number of banks, the probability of loans being repaid, π, reduces, implying riskier loans are given with increasing competition.

Deposit insurance price schedule Rather than setting the deposit insurance premium in advance, let us now assume that the deposit insurer publishes a pricing schedule for their premium, depending in the risk the bank takes and the deposit rate they promise. The pricing schedule will be given by the fair pricing in equation (25.3) and banks will consider the impact of the default on loans and deposit interest on their profits.

Inserting equation (25.3) into equation (25.1), we get the bank profits as

$$\Pi_B=\pi\left((1+r_L)-\frac{1+r_D}{\pi}\right)D=\left(\pi\,(1+r_L)-(1+r_D)\right)D. \tag{25.5}$$

The first order conditions for a profit maximum are then given by

$$\frac{\partial \Pi_B}{\partial (1+r_L)} = \left(\frac{\partial \pi}{\partial (1+r_L)} (1+r_L) + \pi \right) D = 0, \tag{25.6}$$
$$\frac{\partial \Pi_B}{\partial (1+r_D)} = (\pi (1+r_L) - (1+r_D)) \frac{\partial D}{\partial (1+r_D)} - D = 0.$$

We now see that from the first line of these first order conditions that the default rate does not depend on the level of competition; hence competition does not affect risk-taking. The first line will equal zero of the term in brackets is zero and this does only include terms of the risk, π, and the loan rate, r_L, but the size of deposits, and hence the level of competition between banks, is not relevant for the solution of this condition.

What we observe here is the internalization of the bankruptcy costs covered by the deposit insurance through the deposit insurance premium, which makes increasing risks with increasing competition unattractive due to an increasing deposit insurance premium.

The same result would be valid in the absence of deposit insurance. In this case depositors face the risk of not being repaid and thus need adequate compensation through the deposit rate such that their payment from providing the bank with deposits that are only repaid with probability π gives them the same value as retaining the funds without the prospect of earning interest, $\pi (1+r_D) D = D$. Using this result and taking into account that without deposit insurance no deposit insurance premium is payable, $P = 0$, the bank profits in equation (25.1) become

$$\Pi_B = (\pi (1+r_L) - 1) D, \tag{25.7}$$

which is similar to the bank profits in equation (25.5) if we were to set $r_D = 0$. As the deposit rate did not affect the risk-taking by the bank, the results obtained will be similar and in the absence of deposit insurance banks would not alter the level of risks in the loans they provide as competition between banks increases. This is because the costs of this risk, the possibility of default, are internalised through higher deposits rates. In the same way as with deposit insurance before it was internalised through a higher deposit rate.

Summary We have thus seen that if deposit insurance is provided such that the deposit insurance premium is determined in advance of the bank making the decision on the riskiness of the loans based on the inference of the bank's risk, higher competition will increase the risks banks are willing to take. In contrast, with deposit insurance priced to fully take into account the risks actually taken, then no such incentive exists; the same result holds if no deposit insurance is provided. In the latter two cases, the risks of the loans provided are fully internalised by the bank through higher deposit insurance premia and higher deposit rates, respectively, while the pre-determined deposit insurance premium does allow banks to increase risks without increasing their costs. As the loan rates on high-risk loans are higher, the banks would make more profits as there is no correspondent increase in costs. This effect increases as competition becomes more intense because the size of deposits, and thus loans, reduce, making profits ever smaller; consequently banks seek to compensate for this risk by taking on higher risks in order to increase profits in the case the loan is repaid.

Readings Allen & Gale (2001), Grochulski & Kareken (2004)

25.1.2 Moral hazard and adverse selection

When making decisions to provide a loan, banks face the problems of moral hazard and adverse selection. The moral hazard arises where a company has the choice between investment opportunities and some of them might not generate profits for the banks, for example due to the high risks they are exposed to, but might be more profitable to the company. The bank might also not be aware if the company's investments will allow them to repay the loan given the investment opportunities they have; this gives rise to adverse selection if the company is aware of their ability to repay a loan but the bank does not have thin information. We will evaluate how these two effects influence the level of risks banks are willing to take as the competition between banks changes.

Let us assume that banks are faced by two types of companies. A fraction $1-\lambda$ of companies will be unable to pay pack their loans in any circumstances. The other type of companies, a fraction λ, may choose a safe investment returning R_S^1 with certainty in time period 1 from an investment L_1^S, and returning R_S^2 from an investment of L_2^S in time period 2. Alternatively, they can invest into a risky investment that succeeds with probability π and then yields a return of R_R^1 from investment L_1 in time period 1, and then if successful continues with a safe investment in time period 2, yielding R_S^2 (the same return as for the safe investment above) from investing L_S^R.

The bank thus faces a moral hazard and an adverse selection problem. Firstly the adverse selection problem presents itself in the fact that the bank does not know whether it lends to a company that is never able to repay their loans, a fraction $1-\lambda$ of the total companies in the market, while a fraction λ might be able to repay their loan. The moral hazard problem emerges from the choice of investments by the company potentially repaying the loan; this company can make a safe or a risky investment, where only the safe investment is profitable for the bank.

We assume that with financing costs r_D for the deposits of banks, the safe investment is socially desirable as we have

$$\left(1+R_1^S\right)L_1^S-(1+r_D)\,L_1^S+\left(1+R_S^2\right)L_2^S-(1+r_D)\,L_2^S>0, \tag{25.8}$$

while the risky investment is not socially desirable given that

$$\pi\left(1+R_R^1\right)L_1-(1+r_D)\,L_1+\pi\left(\left(1+R_S^2\right)L_2^R-(1+r_D)\,L_2^R\right)<0. \tag{25.9}$$

The safe investment is socially desirable as its returns would exceed that of its funding costs, while the risky investment, on average, would not able to generate sufficient returns to cover its funding costs.

Let us further assume for simplicity that the expected size of the loans in both cases are identical in time period 2, hence

$$\pi L_2^R = L_2^S \tag{25.10}$$

and the expected returns of the risky and safe investments in time period 1 are equal, too, such that

$$\pi\left(1 + R_1^R\right) = 1 + R_1^S. \tag{25.11}$$

The company re-invests the profits of time period 1 to reduce the borrowing in time period 2. Hence the profits for the first and second time period of the safe investment are then given by

$$\begin{aligned}\Pi_C^1 &= \left(1 + R_S^1\right) L_1 - \left(1 + r_L^1\right) L_1, \\ \Pi_C^2 &= \left(1 + R_S^2\right) L_2^S - \left(1 + r_L^2\right) \left(L_2^S - \Pi_C^1\right),\end{aligned} \tag{25.12}$$

where r_L^t denotes the loan rates in time period t. The total profits to the company from the safe investment are thus given by

$$\begin{aligned}\Pi_C &= \Pi_C^1 + \Pi_C^2 \\ &= \left(1 + R_S^2\right) L_2^S - \left(1 + r_L^2\right) L_2^S \\ &\quad - \left(\left(1 + r_L^2\right) - 1\right)\left(\left(1 + R_S^1\right) L_1 - \left(1 + r_L^1\right) L_1\right),\end{aligned} \tag{25.13}$$

where we inserted for Π_C^1 in the expression for P_C^2 using equation (25.12).

Similarly we can obtain the company profits for the risky investment, noting that the investment in the second time period is only made if the initial investment was successful. We then have

$$\begin{aligned}\hat{\Pi}_C^1 &= \pi\left(\left(1 + R_R^1\right) L_1 - \left(1 + r_L^1\right) L_1\right), \\ \hat{\Pi}_C^2 &= \left(1 + R_R^2\right) L_2^R - \left(1 + r_L^2\right)\left(L_2^R - \frac{\hat{\Pi}_C^1}{\pi}\right).\end{aligned} \tag{25.14}$$

For the total company profits we then obtain analogously to above that

$$\begin{aligned}\hat{\Pi}_C &= \hat{\Pi}_C^1 + \hat{\Pi}_C^2 \\ &= \pi\left(\left(1 + R_R^2\right) L_2^R - \left(1 + r_L^2\right) L_2^R\right. \\ &\quad \left. - \left(\left(1 + r_L^2\right) - 1\right)\left(\left(1 + R_R^1\right) L_1 - \left(1 + r_L^1\right) L_1\right)\right).\end{aligned} \tag{25.15}$$

If banks want to ensure that companies choose the socially preferred safe investment, they need to ensure that this is more profitable to do so, $\Pi_C \geq \hat{\Pi}_C$, or from equations (25.13) and (25.15) that

$$\left(\left(1+R_S^2\right)-\pi\left(1+R_R^2\right)\right)L_2^R-\left(1+r_L^2\right)\left(L_2^S-\pi L_2^R\right) \tag{25.16}$$
$$+\left(\left(1+r_L^2\right)-1\right)\left(\left(1+R_S^1\right)L_1-\left(1+r_L^1\right)L_1\right.$$
$$\left.-\pi\left(\left(1+R_R^1\right)L_1-\left(1+r_L^1\right)L_1\right)\right)\geq 0$$

We note that banks would not be able to generate profits from companies that make investments that are not socially desirable; this is because these investment by our assumption in equation (25.9) would not recover its funding costs, which banks have to bear.

The first two terms in these profits are zero due to equations (25.10) and (25.11), and hence the condition for the socially desirable investment being chosen becomes

$$1+r_L^1 \leq \frac{\left(1+R_S^1\right)-\pi\left(1+R_R^1\right)}{1-\pi}. \tag{25.17}$$

Hence, if the loan rate in the first time period is sufficiently low, companies will choose the safe investment. Banks will seek to charge the highest loan rate possible to maximize their profits, and thus this condition will be fulfilled with equality if a bank does not face any competition, but with competition between banks might be lower. The loan rate in the second time period does not affect these incentives, hence for the remainder we can focus mainly on the loan rate in the first time period.

Banks know that companies will choose safe investments if the condition in equation (25.17) is fulfilled, but in the first time period they do not know the type of borrower they are giving a loan to, while for the second time period, only those that are able to repay their loans are continuing and hence they know the type of company they are then lending to. Hence the bank profits are for each period given by

$$\Pi_B^1 = \lambda\left(1+r_L^1\right)L_1-\left(1+r_D\right)L_1, \tag{25.18}$$
$$\Pi_B^2 = \left(1+r_L^2\right)\left(L_2^S-\Pi_C^1\right)-\left(1+r_D\right)\left(L_2^S-\Pi_C^1\right),$$

and total profits are $\Pi_B = \Pi_B^1 + \Pi_B^2$, noting that only a fraction λ of companies continue to period 2. In the first time period loan repayments are only received from those that are able to make such repayments, a fraction λ of all potential companies; we know that given the bank will charge a low loan rate, that companies potentially able to repay their loan will choose the safe investment and hence repayment is guaranteed. In the second time period the safe investment is chosen and the loan is always repaid, but the required loan amount reduced by the retained profits from the first time period.

In order to be willing to provide loans, banks need to be profitable, hence we require that $\Pi_B \geq 0$. This can easily be transformed into

$$\lambda \geq \lambda^* = \frac{(1 + r_D) L_1}{\left((1 + r_L^2) - (1 + r_D)\right)\left(L_2^S - \Pi_C^1\right) + \left(1 + r_L^1\right) L_1}. \tag{25.19}$$

We thus see that banks would only provide loans if there are enough companies able to repay their loans. Hence we can interpret λ as the probability with which the loans are repaid and use this as a measure of the riskiness of loans.

We can easily see that

$$\frac{\partial \lambda^*}{\partial\left(1 + r_L^2\right)} < 0, \qquad \frac{\partial \lambda^*}{\partial\left(1 + r_L^1\right)} < 0. \tag{25.20}$$

We know from the results on oligopolistic markets that the loan rate increases as competition between decreases, hence $\frac{\partial\left(1+r_L^t\right)}{\partial N} < 0$, where we took the number of banks in the market, N, as an indicator of the degree of competition. Using this result we get

$$\frac{\partial \lambda^*}{\partial N} = \frac{\partial \lambda^*}{\partial\left(1 + r_L^t\right)} \frac{\partial\left(1 + r_L^t\right)}{\partial N} > 0 \tag{25.21}$$

and therefore increased competition will reduce the risk of the loans banks give as the probability of loans being repaid increases.

Competition has the effect of reducing risk-taking by banks. The reason for this result is that a lower loan rate in either time period reduces the profits banks can generate and this makes them less willing to take on additional risk as their revenue will be lower, while their costs are not affected. This reduced revenue will require banks to reduce the risks they take in order to remain profitable.

Reading Petersen & Rajan (1995)

25.1.3 Competition in deposit markets

While it is obvious to focus on competition in the loan market to asses the impact on the risk-taking behaviour of banks, it will also be interesting to assess how market power in the deposit market can affect such behaviour. Depositors have access to the money market paying interest r, but when doing so face some costs $c_D \geq 0$. When choosing banks, depositors do not face these costs and hence can charge a lower deposit rate of $1 + r - c_D$ without making depositors worse off. This is equivalent to the lower deposit rate in oligopolistic markets and the size of the discount there reflects the level of competition between banks; thus we can interpret a low value of c_D as a high level of competition, with $c_D = 0$ representing perfect competition, and a high value of c_D representing a low level of competition.

Companies are assumed to be customers of a specific bank and switching to another bank involves costs of c_L. These costs could arise from foregoing the benefits of relationship lending or the move to a bank with different offerings, as in monopolistic competition. The case of $c_L = +\infty$ would correspond to a monopoly as companies would never switch banks and $c_L = 0$ is the case of perfect competition with companies switching banks freely; intermediate values represent imperfect competition between banks and a lower value of c_L corresponds to a higher degree of competition.

We assume that risk and returns of investments are positively related. Hence for the return of companies, we find that $\frac{\partial(1+R)}{\partial\pi} < 0$, such that a higher likelihood of success of the investment, π, a low risk, will be associated with a low return on investment, R. We further assume that this relationship gets stronger the higher the risks are, hence $\frac{\partial^2(1+R)}{\partial\pi^2} < 0$ and $\frac{\partial^3(1+R)}{\partial\pi^3} < 0$.

Monopolistic loan markets in loan markets Let us first consider the case of a monopolistic bank by setting $c_L = +\infty$. Companies succeed with probability π in their investments and are only then able to repay the loan L with interest r_L. Hence their profits are given by

$$\Pi_C = \pi\left((1+R)\,L - (1+r_L)\,L\right) \tag{25.22}$$

and the optimal level of risk is given by the first order condition that maximises the company profits, hence

$$\frac{\partial\Pi_C}{\partial\pi} = \left(\pi\frac{\partial\,(1+R)}{\partial\pi} + \left((1+R) - (1+r_L)\right)\right)L = 0. \tag{25.23}$$

For banks with limited liability and deposit rate $1 + r_D = 1 + r - c_D$ who finance their loans fully by deposits, we have their profits as

$$\begin{aligned}\Pi_B &= \pi\left((1+r_L)\,L - (1+r_D)\,L\right) &(25.24)\\ &= \left(\pi\,(1+R) + \pi^2\frac{\partial\,(1+R)}{\partial\pi} - \pi\,(1+r_D)\right)L,\end{aligned}$$

where for the second equality we solved equation (25.23) for $1 + r_L$ and inserted the result. Equation (25.23) gives us a constraint for the profit maximization of the bank as the company will react to the loan rate r_L by adjusting the level of risk, π, such that their profits are maximised. Banks as monopolists are able to maximize their own profits, taking into account the reaction of the companies seeking loans. The first order condition for the optimal profits of the bank is then given by

$$\begin{aligned}\frac{\partial\Pi_B}{\partial\pi} &= \left(2\pi\frac{\partial\,(1+R)}{\partial\pi} + \pi^2\frac{\partial^2\,(1+R)}{\partial\pi^2}\right. &(25.25)\\ &\qquad\left. + (1+R) - (1+r_D)\right)L\\ &= 0.\end{aligned}$$

Totally differentiating this expression for π and c_D after using that $1 + r_D = 1 + r - c_D$, we get the relationship between the risk-taking of the banks, π and the competition, c_D, as

$$\frac{d\pi}{dc_D} = -\frac{1}{2\frac{\partial(1+R)}{\partial\pi} + 4\pi\frac{\partial^2(1+R)}{\partial\pi^2} + \pi^2\frac{\partial^3(1+R)}{\partial\pi^3}} > 0. \tag{25.26}$$

We thus see that as competition in the deposit market increases through a lower c_D, the risk of default, $1-\pi$, increases. Hence there is a positive relationship between the level of competition in the deposit market and the level of risk taking by the bank. The higher deposit rate that needs to be paid in more competitive markets, the bank seeks to recover these costs through a higher loan rate; this higher loan rate then, however, induces companies to increase the risk they are taking to preserve their own profitability in face of these higher costs.

Banks could, rather than providing loans, invest the proceeds from deposits into the money market at rate r, having only paid $r - c_D$ for the deposits. Hence, banks will always be able to make a profit of at least $c_D L$, representing the difference between these interest rates. We therefore require that the profits from lending have to exceed the profits from investing into the money market and hence

$$\Pi_B \geq c_D L, \tag{25.27}$$

or when inserting from for the profits from equation (25.24) that the market power in the deposit market is not too high as we can solve this expression for

$$c_D \leq \underline{c}_D = \frac{\pi(1+R) + \pi^2\frac{\partial(1+R)}{\partial\pi} - \pi(1+r)}{1-\pi}. \tag{25.28}$$

Thus the market power in the deposit market must not be too large as otherwise the investment into the money market is too attractive compared to granting loans.

Perfect competition in loan markets Turning to the case of perfect competition, banks will only break even compared to the investment into the money market such that equation (25.27) is fulfilled with equality. With equation (25.24) this condition can be solved for

$$1 + r_L = (1+r) + \frac{1-\pi}{\pi}c_D, \tag{25.29}$$

which serves as a constraint on the profit maximization of the company. Inserting equation (25.29) into equation (25.22) and differentiating, we obtain the first order condition for a profit maximum of the companies as

$$\begin{aligned}\frac{\partial\Pi_C}{\partial\pi} &= \left(\pi\frac{\partial(1+R)}{\partial\pi} + (1+R) - (1+r_D)\right)L \\ &= 0,\end{aligned} \tag{25.30}$$

where we used that $1 + r_D = 1 + r - c_D$. Totally differentiating this expression for π and c_D, we get the relationship between the risk-taking of the banks, π and the competition, c_D, as

$$\frac{d\pi}{dc_D} = \frac{1}{2\pi\frac{\partial(1+R)}{\partial\pi} + \pi\frac{\partial^2(1+R)}{\partial\pi^2}} < 0. \tag{25.31}$$

The sign of this relationship arises from our assumption that $\frac{\partial(1+R)}{\partial\pi} < 0$ and $\frac{\partial^2(1+R)}{\partial\pi^2} < 0$. Hence there is a negative relationship between the level of competition in the deposit market and the level of risk taking by the bank. Given the higher deposit rate that needs to be paid in more competitive markets and the therefore reduced profits the bank can make, the bank seeks to ensure their profitability by not increasing risks too much and facing bankruptcy; this results in banks taking less risks as competition in the deposit market increases.

We can solve equation (25.30) for

$$1 + r_D = \pi\frac{\partial\,(1+R)}{\partial\pi} + (1+R)\,, \tag{25.32}$$

which we can insert into equation (25.25) to obtain

$$\frac{\partial\Pi_B}{\partial\pi} = \left(\pi^2\frac{\partial^2\,(1+R)}{\partial\pi^2} + \pi\frac{\partial\,(1+R)}{\partial\pi}\right)L < 0. \tag{25.33}$$

As equation (25.25) represents the condition for the optimal risk in the monopolistic case and we inserted the solution of the competitive case, we see that for the for the optimal risk level in the monopolistic case, π_M, compared to the optimal risk level in the case of perfect competition, π_C, we observe $\pi_M < \pi_C$. The negative sign in equation (25.33) implies that π_C is exceeding π_M and hence we have established the relationship between the optimal risks in the cases of a monopoly in the loan market and perfect competition.

We thus find that competitive loan markets are less risky than monopolistic loan markets, but we also see from equation (25.31) that in contrast to the monopolistic case, the risk increases as deposit markets become less competitive.

Imperfect competition in loan markets After these extreme cases of a monopoly and perfect competition in loan markets, we now look at the case of imperfect competition in loan markets, the case where $0 < c_L < +\infty$.

In the case of perfect competition, we can use equation (25.29) and then solve equation (25.30) for the loan rate to be set as

$$1 + r_L = 1 + R + \pi\frac{\partial\,(1+R)}{\partial\pi} + \frac{c_D}{\pi}. \tag{25.34}$$

Inserting this result into equation (25.22), we obtain the company profits in perfect competition as

$$\Pi_C^C = -\left(\pi^2\frac{\partial\,(1+R)}{\partial\pi} + c_D\right)L. \tag{25.35}$$

In order to entice a company to change banks, the bank seeking this company to switch would engage in competitive behaviour and hence charge competitive loan rates, resulting in risks as derived in perfect competition. Hence for a company to stay with the existing bank we need that the profits this company obtains, Π_C^M, to exceed the profits if switching to a banks setting competitive loan rates, Π_C^C, taking into account the switching costs c_L, hence we require that $\Pi_C^M \geq \Pi_C^C - c_L L$.

If the bank would not face the threat of the company switching to another bank, it would act as a monopolist and charge loan rates accordingly. From equation (25.25) we can insert into equation (25.22) and obtain

$$\Pi_C^M = -\pi^2 \frac{\partial (1+R)}{\partial \pi} L. \tag{25.36}$$

We can now determine how these profits change if we change the degree of competition in the deposit market, as represented by c_D. This gives us

$$\begin{aligned} \frac{\partial \Pi_C^C}{\partial c_D} &= -\left(2\pi \frac{\partial \pi}{\partial c_D} \frac{\partial (1+R)}{\partial \pi} + 1\right) L < 0, \\ \frac{\partial \Pi_C^M}{\partial c_D} &= -2\pi \frac{\partial \pi}{\partial c_D} \frac{\partial (1+R)}{\partial \pi} L > 0, \end{aligned} \tag{25.37}$$

using the results on the effect of competition on the optimal risk-taking behaviour of banks from equations (25.26) and (25.31) as well as our assumptions that $\frac{\partial(1+R)}{\partial \pi} < 0$. As we decrease c_D, the monopolistic left hand side of the condition $\Pi_C^M \geq \Pi_C^C - c_L L$ reduces while the competitive right hand size increases such that for a sufficiently small c_D^* this constraint becomes binding. The larger c_L is, the smaller c_D^* will be as this condition will be fulfilled sooner.

If the condition that $\Pi_C^M \geq \Pi_C^C - c_L L$ is binding, then

$$c_L = \pi_M^2 \frac{\partial (1+R_M)}{\partial \pi} - \pi_C^2 \frac{\partial (1+R_C)}{\partial \pi} - c_D. \tag{25.38}$$

Totally differentiating this expression for π_M and π_C, we obtain

$$\frac{d\pi_M}{d\pi_C} = \frac{2\pi_C \frac{\partial(1+R_C)}{\partial \pi} + \pi_C^2 \frac{\partial^2(1+R_C)}{\partial \pi^2}}{2\pi_M \frac{\partial(1+R_M)}{\partial \pi} + \pi_M^2 \frac{\partial^2(1+R_M)}{\partial \pi^2}} > 0. \tag{25.39}$$

Whenever the constraint is binding, we have

$$\frac{\partial \pi_M}{\partial c_D} = \frac{\partial \pi_M}{\partial \pi_C} \frac{\partial \pi_C}{\partial c_D} < 0 \tag{25.40}$$

using the signs in equations (25.26) and (25.39). From the condition $\Pi_C^M = \Pi_C^C - c_L L$, we can solve for

$$\frac{\partial (1+R_C)}{\partial \pi} = \frac{\pi_M^2}{\pi_C^2} \frac{\partial (1+R_M)}{\partial \pi} - \frac{c_L + c_D}{\pi_C^2} \tag{25.41}$$

and insert this result into the first order condition for the perfect competition, equation (25.30), such that

$$\frac{\partial \Pi_C}{\partial \pi} = \left(\frac{\pi^2}{\pi_C} \frac{\partial (1+R)}{\partial \pi} - \frac{c_L}{\pi_C} + (1+R) - (1+r_L) \right) L. \tag{25.42}$$

If we assumed that $\pi_M = \pi_C$, this becomes

$$\frac{\partial \Pi_C}{\partial \pi} = -\frac{c_L + c_D}{\pi_C} L < 0. \tag{25.43}$$

using equation (25.30). Hence we find that it would be optimal to have $\pi_M < \pi_C$ as $\pi_M = \pi_C$ is above the optimum for the company. Hence we see that if the constraint to prevent companies switching banks becomes binding, $c_D < \hat{c}_D$, the risk π is decreasing in market power, and decreasing for larger market power.

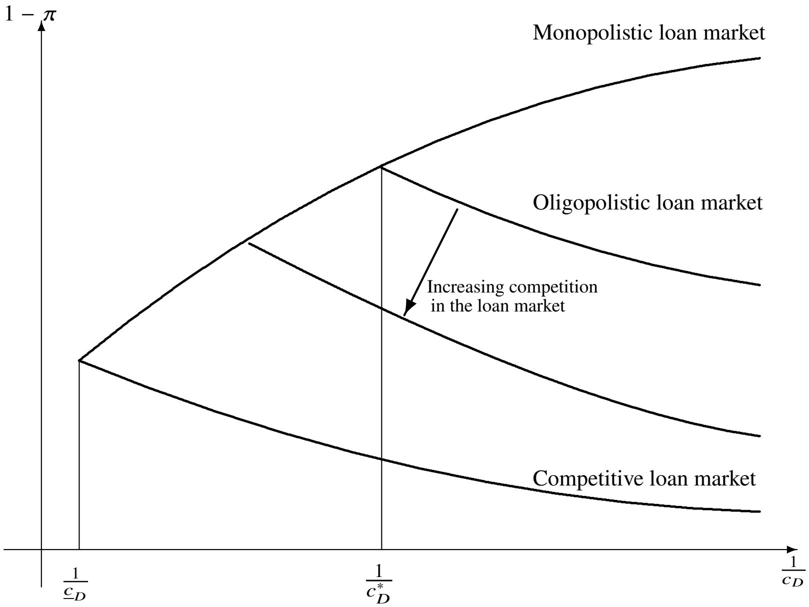


Fig. 25.1: Risk taking with competition in deposit markets

We thus see that in the general case of imperfectly competitive loan markets, in low levels of competition in the deposit market, the banks will behave like monopolists in the loan market and the risk banks take as the competition in the deposit market increases, will also increase. This is due to the market power banks have in the deposit market, they can set loan conditions such that companies will not consider switching to another banks due to the costs involved when doing so. As competition in deposit market increases, the loss of market power makes it ever more difficult to prevent companies from switching banks and there will come a point where banks need to act competitively in the loan markets to avoid companies switching banks. Thus banks will act like a banks in perfect competition and reduce the risk taken if competition in the deposit market increases further. This leads to a situation where for low levels of competition in deposit markets the risks banks take are increasing in the deposit market competition, but for higher levels of deposit market competition the risks will be decreasing. This change from on increase in risks to a decrease in risks if the competition in deposit markets increases to a decrease will come sooner of the competition in the loan market is more intense. The reason is that in this case the costs of switching banks for companies are lower and hence banks need to act competitively sooner to prevent companies from switching to other banks. Figure 25.1 summarizes these results.

Summary We have seen that competition in the market for deposits can affect the level of risks banks optimally take when providing loans. In the most realistic case of imperfectly competition in the loan market, an increase in the competition in deposit markets leads to an increase in the risk banks take when providing loans if this competition is low, but if competition for deposits is sufficiently high, a further increase in competition reduces the risks banks take.

With this result it is clear that competition in deposit markets can affect the optimal risk-taking in loan markets and the influence any measures to increase competition in deposit markets have, will depend on the initial level of competition. If competition in deposit markets is low, an increase in competition can lead to an increase in the risks banks are taking more risks; if competition in deposit markets is already high, it may lead to banks taking lower risks. Increasing competition in loan markets will either leave the risk-taking of banks unaffected if competition in deposits markets is low, or it will reduce risk-taking if competition in deposit markets is sufficiently high.

Reading Arping (2017)

Résumé

The relationship between competition and the risks that banks take is complex. We have seen that on the one hand competition might increase the incentives to take risks in order to compensate for the loss in profits due to competition; this was only the case, however, if the costs of these increased risks were not fully internalised. Whenever the additional costs of the risks are fully internalised, the incentives to take risks do not change. On the other hand, we have also seen that risks are reducing with increased competition as banks facing higher competition seek to protect their lower profit margins by taking on less risks. Competition in deposit markets will also affect the risk taking in loan markets, but the effect will depend on the degree of competition in the loan market; in monopolistic loan markets, the risks banks take are increasing as competition in deposit market increases and for competition loan markets this relationship is reversed. To complicate the analysis even more, imperfectly competitive loan markets will result in higher risks being taken with deposit markets being less competitive but increasing competition, but lower risks are taken if the deposit market is more competitive and increases the level of competition further.

While the effect of competition on loan and deposit rates are mostly obvious in that loan rates reduce and deposit rates increase, the effect on loan risks reflects a range of influences. The overall impact of competition on the risk-taking behaviour of banks will depend on the relative importance of the individual factors, which generally vary across markets, across time, but also across individual banks.

25.2 Diversification and bank failures

Competition between banks may influence the decision on the optimal risks individual loans pose to banks, but banks can also manage the risks they are taking through the types of loans they provide. Banks may provide loans to similar companies and thus the probability of a a bank failing might be higher than a bank diversifying their loan portfolio and thereby reducing the risk of bank failure considerably. We will therefore focus now on a model assessing the risk of bank failures rather then the risks associated with a single loan.

Let us assume that companies have access to investment opportunities whose returns R are increasing in the risk they are taking. If π denotes the probability the investment succeeds, then we have $\frac{\partial(1+R)}{\partial\pi} < 0$. We also assume $\frac{\partial^2(1+R)}{\partial\pi^2} < 0$ and as we increase the risk, $1-\pi$, the effect on returns is increasing.

With companies financing their investment with a loan the size of L_i at a loan rate of r_L, we have the their profits given as

$$\Pi_C^i = \pi\left(\left(1+R\right)L_i - \left(1+r_L\right)L_i\right). \tag{25.44}$$

The first order condition for a profit maximum when choosing the optimal risk level of the investment is given by

$$\frac{\partial\Pi_C^i}{\partial\pi} = \left(\pi\frac{\partial\left(1+R\right)}{\partial\pi} + \left(1+R\right) - \left(1+r_L\right)\right)L_i = 0. \tag{25.45}$$

Totally differentiating this expression with respect to π and $1+r_L$ gives us after re-arranging that

$$\frac{d\left(1+r_L\right)}{d\pi} = 2\frac{\partial\left(1+R\right)}{\partial\pi} + \pi\frac{\partial^2\left(1+R\right)}{\partial\pi^2} < 0 \tag{25.46}$$

and we see that safer investments, a higher π, will require lower loan rates to be optimal for companies.

Banks are lending to a large number of companies and we assume that their defaults have a correlation of ρ. Define

$$\xi_i = -\Phi^{-1}\left(1-\pi\right) + \sqrt{\rho}z + \sqrt{1-\rho}\varepsilon_i, \tag{25.47}$$

where z and ε_i are independently distributed random variables that follow a standard normal distribution, $z \sim N(0,1)$ and $\varepsilon_i \sim N(0,1)$, and $\Phi\left(\cdot\right)$ denotes the cumulative density function of the standard normal distribution. We can interpret z, which is common to all companies as a common factor determining the companies' ability to repay loans, such as macroeconomic conditions, while ε_i denotes a company-specific factor, for example their market position or balance sheet strength. The initial term, $-\Phi^{-1}\left(1-\pi\right)$, can be interpreted as a constant factor that accounts for the probability of default of the average company; with z and ε_i being normalised to zero, these factors will only account for deviations from this average company.

We then obtain

$$\begin{aligned} Prob\left(\xi_i < 0\right) &= Prob\left(\sqrt{\rho}z + \sqrt{1-\rho}\varepsilon_i < \Phi^{-1}\left(1-\pi\right)\right) \\ &= \Phi\left(\Phi^{-1}\left(1-\pi\right)\right) \\ &= 1-\pi, \end{aligned} \tag{25.48}$$

where the second equality uses that as $E\left[\sqrt{\rho}z + \sqrt{1-\rho}\varepsilon_i\right] = 0$ and $Var\left[\sqrt{\rho}z + \sqrt{1-\rho}\varepsilon_i\right]$ 1 and the term follows a normal distribution. We can thus interpret $Prob\left(\xi_i < 0\right)$ as the probability of default of an individual borrower.

Let us now define the failure rate in a portfolio of such loans, $1-\gamma$, for a given value of the common factor z, as

$$\begin{aligned} 1-\gamma &= Prob\left(\xi_i < 0|z\right) \\ &= Prob\left(\varepsilon_i < \frac{\Phi^{-1}\left(1-\pi\right) - \sqrt{\rho}z}{\sqrt{1-\rho}}|z\right) \\ &= \Phi\left(\frac{\Phi^{-1}\left(1-\pi\right) - \sqrt{\rho}z}{\sqrt{1-\rho}}\right). \end{aligned} \tag{25.49}$$

This expression captures the failure rate of a portfolio as we only consider the idiosyncratic risk as captured by ε_i, which in a sufficiently large portfolio will be diversified away. The distribution of $1-\gamma$, the failure rate in this large portfolio, will be driven by the common factor z and we obtain

$$\begin{aligned}
F(1-\gamma) &= Prob\left(\Phi\left(\frac{\Phi^{-1}(1-\pi)-\sqrt{\rho}z}{\sqrt{1-\rho}}\right)<1-\gamma\right) \\
&= Prob\left(z>\frac{\Phi^{-1}(1-\pi)-\sqrt{1-\rho}\Phi^{-1}(1-\gamma)}{\sqrt{\rho}}\right) \\
&= 1-\Phi\left(\frac{\Phi^{-1}(1-\pi)-\sqrt{1-\rho}\Phi^{-1}(1-\gamma)}{\sqrt{\rho}}\right) \\
&= \Phi\left(\frac{\sqrt{1-\rho}\Phi^{-1}(1-\gamma)-\sqrt{\rho}-\Phi^{-1}(1-\pi)}{\sqrt{\rho}}\right),
\end{aligned} \tag{25.50}$$

using the fact that z is assumed to be normally distributed.

Banks lend the amount $L_i = \frac{L}{N}$ to each company and have limited liability such that any losses do not need to be covered. The expected profits are given by

$$\Pi_B^i = \max\left\{\gamma\left(1+r_L\right)L_i-\left(1+r_D\right)L_i, 0\right\}, \tag{25.51}$$

with γ denoting the fraction of loans repaid and r_D denotes the deposit rate; we assume that loans are fully financed by deposits. A bank will fail if the repayment from loans do not cover their obligation from repaying deposits, $\gamma\left(1+r_L\right)L_i < \left(1+r_D\right)L_i$, or

$$1-\gamma > 1-\gamma^* = \frac{\left(1+r_L\right)-\left(1+r_D\right)}{1+r_L}. \tag{25.52}$$

As the actual fraction of loans repaid, γ, is stochastic, the bank will seek to maximize their expected profits, that can written as

$$E\left[\Pi_B^i\right] = \left(\int_0^{1-\gamma^*}\left(\gamma\left(1+r_L\right)-\left(1+r_D\right)\right)f\left(1-\gamma\right)d\gamma\right)L_i. \tag{25.53}$$

We take into account that the bank can only realise profits if the failure rate from loans, $1-\gamma$, is not too high.

For convenience let us define

$$\hat{\Pi}_B^i = \int_0^{1-\gamma^*}\left(\gamma\left(1+r_L\right)-\left(1+r_D\right)\right)f\left(1-\gamma\right)d\gamma \tag{25.54}$$

such that

$$E\left[\Pi_B^i\right] = \hat{\Pi}_B^i L_i. \tag{25.55}$$

Using standard competition arguments, we know that as banks seek to lend more to each company by increasing L_i, the loan rate will fall, $\frac{\partial(1+r_L)}{\partial L} < 0$, thus reducing the profits $\hat{\Pi}_B^i$ per unit of loans provided. We assume now, however, that as the loan rate falls, the default rates on loans, $1-\pi$, reduces as indicated by equation (25.46), but this reduction does not fully outweigh the loss of revenue from the lower interest rate. Thus we assume $\frac{\partial\hat{\Pi}_B^i}{\partial L_i} < 0$ and $\frac{\partial^2\hat{\Pi}_B^i}{\partial L_i^2} < 0$ in line with our assumption that $\frac{\partial^2(1+R)}{\partial\pi^2} < 0$.

The first order condition of profit maximization from equation (25.55) becomes

$$\frac{\partial E\left[\Pi_B^i\right]}{\partial L_i} = L_i\frac{\partial\hat{\Pi}_B^i}{\partial L_i} + \hat{\Pi}_B^i = 0. \tag{25.56}$$

Totally differentiating this relationship with respect to the aggregate lending L and the number of banks N, we get using $L_i = \frac{L}{N}$ and after re-arranging

$$\frac{dL}{dN} = \frac{L}{N}\frac{\frac{\partial\hat{\Pi}_B^i}{\partial L}}{L\frac{\partial^2\hat{\Pi}_B^i}{\partial L^2} + (N+1)\frac{\partial\hat{\Pi}_B^i}{\partial L}} > 0. \tag{25.57}$$

The sign becomes obvious if we recall that from our discussion above that we assume that $\frac{\partial\hat{\Pi}_B^i}{\partial L_i} < 0$ and $\frac{\partial^2\hat{\Pi}_B^i}{\partial L_i^2} < 0$, and as $L_i = \frac{L}{N}$ we have $\frac{\partial L_i}{\partial L} > 0$. We then can derive that

$$\frac{\partial(1+r_L)}{\partial N} = \frac{\partial(1+r_L)}{\partial L}\frac{\partial L}{\partial N} < 0, \tag{25.58}$$

implying that the loan rate is decreasing as competition between banks, as measured by the number of banks N, increases.

From equation (25.52) we get that the probability of a bank failing is given by

$$\begin{aligned}\hat{\pi} &= Prob\,(1-\gamma > 1-\gamma^*) \qquad (25.59)\\ &= 1 - F\,(1-\gamma^*)\\ &= \Phi\left(\frac{\Phi^{-1}(1-\pi) - \sqrt{1-\rho}\,\Phi^{-1}(1-\gamma^*)}{\sqrt{\rho}}\right).\end{aligned}$$

From this expression we get with $\phi(\cdot)$ denoting the density of the normal distribution, that

$$\begin{aligned}\frac{\partial\hat{\pi}}{\partial L} = \frac{\phi}{\sqrt{\rho}}&\left(\frac{\partial\Phi^{-1}(1-\pi)}{\partial\pi}\frac{\partial\pi}{\partial(1+r_L)}\frac{\partial(1+r_L)}{\partial L}\right. \qquad (25.60)\\ &\left.-\sqrt{1-\rho}\frac{\partial\Phi^{-1}(1-\gamma^*)}{\partial(1-\gamma^*)}\frac{\partial(1-\gamma^*)}{\partial L}\right).\end{aligned}$$

From this expression we easily obtain that

$$\frac{\partial \Phi^{-1}(1-\pi)}{\partial \pi} = -\frac{1}{\phi\left(\Phi^{-1}(1-\pi)\right)} < 0, \tag{25.61}$$
$$\frac{\partial \Phi^{-1}(1-\gamma^*)}{\partial (1-\gamma^*)} = -\frac{1}{\phi\left(\Phi^{-1}(1-\gamma^*)\right)} > 0,$$
$$\frac{\partial \Phi^{-1}(1-\pi)}{\partial L} = \frac{1+r_D}{(1+r_L)^2}\frac{\partial (1+r_L)}{\partial L} < 0.$$

As we know from equation (25.46) that $\frac{\partial \pi}{\partial (1+r_L)} < 0$, in equation (25.60) the first term in brackets is positive, as is the second term; hence the overall sign will depend on the relative size of each term. The first term denotes risk shifting as more competition leads to more overall lending from as indicated in equation (25.57), which lowers the loan rate as $\frac{\partial(1+r_L)}{\partial L} < 0$, which in turn lowers the default rate as companies seek less risky investments, shown in equation (25.46). The second term gives the marginal effect of more competition reducing the loan rate, which reduces bank profits and makes banks thus more likely to fail.

Which of these two effects dominates, can in general not be determined. Given that

$$\frac{\partial \hat{\pi}}{\partial N} = \frac{\partial \hat{\pi}}{\partial L}\frac{\partial L}{\partial N}, \tag{25.62}$$

and from equation (25.57) $\frac{\partial L}{\partial N} > 0$, the overall impact of competition (N) on bank failure rates $(1-\hat{\pi})$ will thus depend on the relative importance of these two affects.

If the return on investments, R, is not affected by the risk, thus if we had $\frac{\partial(1+R)}{\partial \pi} = 0$, implying that $\frac{\partial(1+r_L)}{\partial \pi} = 0$, the risk-shifting term vanishes and we have $\frac{\partial \hat{\pi}}{\partial L} > 0$, hence competition would increase the risk of banks failing. In the case of perfectly correlated defaults, $\rho = 1$, the second term vanishes and as $\frac{\partial \hat{\pi}}{\partial L} < 0$ increasing competition reduces the risk of bank failure. For uncorrelated defaults, $\rho = 0$, the number of defaulting loans are fixed at $1-\pi$ by the law of large numbers; thus banks would set rates such that defaults are never occurring and banks are perfectly safe. We see that a well diversified portfolio of loans with lower correlations might not be beneficial to the risks of banks failing as a portfolio with a high correlation between loan defaults, and thus very similar loans, reduces bank failures if competition increases.

While nothing can be said about the sign of $\frac{\partial \hat{\pi}}{\partial L}$ in general, we can show that for a large number of banks, N, thus a high degree of competition, we find $\frac{\partial \hat{\pi}}{\partial L} > 0$. From equation (25.56) we get that

$$\lim_{N\to+\infty} \frac{L}{N}\frac{\partial \hat{\Pi}_B^i}{\partial L_i} + \hat{\Pi}_B^i = \hat{\Pi}_B^i = 0 \tag{25.63}$$

and hence with equation (25.54) we see that $\gamma(1+r_L) - (1+r_D) = 0$. As $1+r_L$ is decreasing in N as shown in equation (25.58), we need the probability of a bank surviving, γ to increase until it reaches $\gamma = 1$, such that $1+r_L = 1+r_D$ as a limit. With $1+r_L$ decreasing, equation (25.46) implies that π increases, but we reasonably can assume that $\pi < 1$ at all times, even if $1+r_L = 1+r_D$. Hence from equation

(25.61) we see that $\frac{\partial \Phi^{-1}(1-\pi)}{\partial \pi} < +\infty$. On the other hand as γ approaches 1, we observe that $\frac{\partial \Phi^{-1}(1-\gamma^*)}{\partial (1-\gamma^*)} = \frac{1}{\phi\left(\Phi^{-1}(0)\right)} = +\infty$ as the density of the normal distribution at infinity is zero. Using these results, we find that for sufficiently large competition, a high N, we have $\frac{\hat{\pi}}{\partial L} > 0$ and hence more competition would increase the risk of banks failing.

We see that with equations (25.46) and (25.58)

$$\frac{\partial \pi}{\partial N} = \frac{\partial \pi}{\partial (1 + r_L)} \frac{\partial (1 + r_L)}{\partial N} > 0 \tag{25.64}$$

and increased competition would in this model reduces the risk of individual loans given as they are more likely to be repaid. The impact on the failure of the bank, however is much more complex to assess. For highly correlated loan defaults and low degrees of competition, a further increase in competition will decrease the risk of banks failing, while higher degrees of competition, investments whose returns are less sensitive to the risk taken, or low default correlations, will actually increase the risk of bank failure when increasing competition. The detailed delineation of this effect will depend on the specific characteristics of the loans and the existing level of competition between banks.

We thus find that the effect of increasing competition on the risk of bank failure does not only depend on the exiting level of competition between banks, but also the correlation between the loans banks are providing. Overall, higher degrees of existing competition favour an increase in bank failures as does a well diversified loan portfolio.

Reading Martinez-Miera & Repullo (2010)

Conclusions

We have seen that competition between banks does not only affect loan and deposit rates offered, but also the risks banks are taking. This effect on risks typically arises through the impact different loan rates have on the optimal decision by companies on the risks they are seeking to take with the loans they have been provided with. Banks will, of course, take into account these incentives of companies to adjust their risk-taking as loan rates change.

If the costs of banks from taking additional risks are not internalised by the bank, for example through an appropriately priced deposit insurance, banks have incentives to recover any lower profit margins from increased competition by increasing the risks they are taking. On the other hand, increased competition and lower profit margins may make banks more cautious in their lending decisions in order to reduce the potential losses from defaulting companies.

The risk-taking behaviour of banks is not only affected by competition in the loan market, but the deposit market also has an influence. We have seen that and increase in competition in deposit markets from low levels will increase risk-taking by banks, while when competition in deposit markets is already high, risks will reduce. High profits in deposit markets allows for potentially high profit margins and thus can induce more competition in the loan market, affecting the behaviour of companies and thus the overall risk banks are exposed to.

Finally, we established that for the risk of bank failure, not only are the risks of each loan relevant, but also the correlations between defaults in a portfolio of loans. In addition to the level of competition between banks, the effect of an increase in competition will also depend on the correlations of default; a low level of correlation would lead to an increase in the risk of bank failure, while a high correlation would lead to a reduced risk of bank failure.

From this discussion it is apparent that competition does not only affect the prices banks charge, loan and deposit rates, but also the risks they are taking. How these risks are affected will be determined by a number of factors and the overall effect will depend on the relative strength of them, making any analysis of the impact from any policies to increase competition difficult perform.

Review

Increasing competition between banks or between banks and alternative lenders as well as shadow banks does not simply reduce loan rates and increase deposit rates. Such price changes, especially in the loan market have the effect of changing the behaviour of borrowers beyond a simple increase in demand due to a lower price; it can affect the risk-taking of borrowers and thus the risks of loans that banks provide. The details of such influences is often complex and will depend on the specific assumption of the model, many of which provide opposing results. As is common, each model will be able to capture some aspect that influence the effect of competition on banks, and it is the aggregation of all these influences, the judgement of how strong one factor is compared to another, that will drive the overall result.

Taking into account the generally seen as positive result of competition on increasing deposit and reducing loan rates, need to be balanced against the incentives for banks to take additional risks in order to maintain their profitability, although in some circumstances banks with reduced profitability might even become more cautious in their lending decision and reduce the risks they are taking. Any secondary measures to counter additional risk taking might be counter productive in that it might encourage shadow banking or make even stronger existing benefits to alternative lenders, undermining the efforts or even having the opposite effect.

While we have seen that by-and-large we can view deposit and lending markets as separate and assess them as two distinct 'products', we cannot neglect that banks losing deposits to shadow banks will not be able to provide many loans as deposits act as a constraint on their ability to do so. Similarly, a bank not being able to provide many loans due to competition from other lenders will not want to attract large deposits as they would not generate enough surplus from loans to provide depositors with an adequate return. Thus although in many instances loan and deposit markets can be viewed as being separated, they are inherently interlinked and any competitive measures in one market can easily have effects in the other, even if it is only due constraints they impose.

This part has shown the intrinsic difficulty in the competition between banks, complicated by asymmetric information between banks and borrowers as well as the moral hazard that may exist in lender-borrower relationships. What in many product markets might be seen as a predictable effect of increasing (or reducing) competition, might in banking markets have consequences that on first sight seem counterintuitive and will only become understandable if the specific situation a bank finds itself in is properly evaluated.

Part V
The conduct of banking activities

We often implicitly or explicitly assume that the balance sheet of a bank is given, employees make decisions honestly and in the best interest of the bank, whose owners are external shareholders. Such assumptions are convenient as the investigation focusses on other aspects of the decisions banks make as well as their implications. It is, however, important to establish how these aspects of banks and their decision-making are established.

Banks typically hold a fraction of the deposit they hold as cash as well as holding some equity in addition to the funding through deposits. We find that the cash held is in nearly all instances well in excess of what is required to ensure that banks can honour the withdrawal of deposits. Equally, taken aside regulatory requirements, banks hold more equity than needed. We will address the reason for holding cash and equity in chapter 26 and also consider the risks arising from changing interest rates due to the imbalance between short-term deposits and long-term loans. This will give us an insight balance sheet structure that suits the needs of a bank best. While we often abstract from such requirements for convenience in many model, it is nevertheless essential to gain an understanding of the importance of these decisions.

Another commonly made assumption is that employees and managers of banks make decisions which are in the best interest of the bank they are working for. We will in chapter 27 investigate what impact the remuneration has on the decision of employees and how banks use the remuneration of employees to ensure that they work in the best interest of their bank. This will allow us to consider the pay bankers are receiving and how their decisions may be distorted if they maximize their own remuneration rather than the profits of the bank. Given that decisions in banks are made by employees rather than the bank owners directly, we will also have to consider incentives for employees further than only their own interest through the maximisation of their own remuneration received from the bank; employees might distort decisions in response to other payments they might receive, for example from borrowers to grant a loan they would not otherwise be granted. Such corruption can have significant impact on the performance of the bank. However, as we will see in chapter 28, it not only employees that have incentives to behave dishonestly or even illegally; such behaviour might also be beneficial to the bank.

While we often only assess the provision of a loan, bank have to make a decision whether to grant new loans at all. Instead of seeking to attract additional borrowing or deposits, banks might instead focus on improving the quality of services they provide to existing customers. While both activities can and will be pursued simultaneously, banks might choose to focus more on one aspect at the expense of the other. In chapter 29 will therefore look at the strategy banks pursue to grow their business optimally. But banks do not only invest into their business lines, it is often found that they are also engaged in many commercially not viable activities, such as the sponsoring of cultural events. We will also investigate why banks may choose to do so, even if the investment could generate a significantly higher return when invested into their core business.

The common assumption is that banks are owned by private individuals, often, but not necessarily, shareholders. However many banks are organised as mutual banks, who are owned by the customers they seek to serve. Other banks are at least partially owned by governments or other political organisations and some of the loans that are granted would not be granted by privately owned banks, they are politically motivated. In chapter 30 we will investigate the implications of these alternative ownership forms for the provision of loans.

A final aspect to consider in chapter 31 is that we usually assume that banks compete on the price for loans, deposits, and other services. The implicit assumption therein is that borrowers and depositors can compare the prices between banks and choose the bank that offers them the best conditions. Often, however, the pricing structure is so complex that comparing prices between banks is very difficult; we therefore will investigate why this is the case and what drives banks to adopt such intransparent pricing structures.

Chapter 26

Optimal balance sheets

Banks make active decisions on the amount of loans they provide and how much deposits they accept. They have not only to balance the need for cash reserves to meet any demands from deposit withdrawals with generation of profits through the provision of loans. In chapter 26.1 we will first assess the risk exposure banks have that arises out of the mismatch between short-term deposits and long-term loans.

Banks do not only have to decide whether to provide a specific loan, but they also have to decide on the overall amount of lending as well as how to finance the loans they provide. Any funds available not given out as loans will be retained as cash reserves and might also be invested into other, more long-term assets; the latter we neglect in our analysis. Thus banks will have to decide how much of the existing funds they retain as cash reserves and how much they provide as loans. We will investigate the demand for cash reserves in chapter 26.2, based on the costs of adjusting cash reserves if and as needed and also consider how cash reserves can be used to incentivise banks to reduce the risks they are taking.

Banks can fund loans, and cash reserves, either through deposits or through equity. Chapter 26.3 will establish that banks do not only hold equity due to its regulatory requirement, but they can use it to signal the risks they are taking when providing loans and they hold equity to show their commitment to choose loans such that their loan portfolio has a low risk. It is thus that financing loans through a larger amount of equity can be used to show that banks have taken low risks or that they are committed to do so.

A central role of banks is to assess the creditworthiness of loan applicants and to this effect they have to develop systems that allows them to do so adequately. In chapter 26.4 we will assess whether it is preferable for banks to choose a basic risk assessment which does not distinguish risks well, or a more advanced risk assessment, which comes at a higher cost.

26.1 Asset and liability management

The main assets of a bank are the loans provided to companies, with cash reserves, fixed assets, and other investment usually playing only a minor role. Similarly, deposits are the main liability of banks, while equity is only small position; some banks will also have considerable borrowing from central banks or in the interbank loan market. If we focus on loans as assets and deposits as liabilities, we immediately see that the key role of banks in transforming short-term deposits into long-term loans as discussed in chapter 4.1, will lead to a mismatch in the maturity of assets and liabilities. While such a mismatch opens the possibility of bank run, see chapter 15, it also poses a risk in term of the profitability of banks. If we assume that loan rates are fixed for the duration of a loan while those of deposits are adjusted frequently to the prevailing market conditions, then an increase in deposit rates due to changing market conditions could lead to profits margins of banks being eroded as loan rates cannot be adjusted similarly. Conversely, a reduction in deposit rates will increase the profitability of banks as long as long rates on existing loans do not need to be adjusted downwards as well. We will therefore here determine the exposure of banks to such interest rate risk and assess how banks can reduce their risk exposure.

A assume that the current value of a loan is given by V_L and this loan has a time to maturity of T_L time periods. Each time period the bank receives interest r_L on the notional face value L of this loan and the fixed repayment L at maturity. Hence, ignoring the default risk, the value of such a loan is given by the present value of these future cash flows such that

$$V_L = \sum_{i=1}^{T_L} \frac{r_L L}{(1+\hat{r}_L)^i} + \frac{L}{(1+\hat{r}_L)^{T_L}}, \tag{26.1}$$

where $\hat{r}_L$ denotes the current level of the loan rate, which might be different from when the bank made the loan. The time to maturity of the loan, T_L, is the time of the loan until it is fully repaid, not the length of the loan when originated. Furthermore we assume in this valuation that loans are only repaid at maturity and borrowers do now make partial repayments throughout the life time of the bond. Such an arrangement is common for loans to individual borrowers, while for corporate loans it is more common that only interest is paid until the loan is due for repayment; this is also the common arrangement for bonds. Taking into account early repayments in the above valuation formula can easily be achieves and would not alter the results obtained here significantly.

The value of this loan now changes a the current loan rate changes. We can determine the marginal change in the bond value as

$$\frac{\partial V_L}{\partial (1+\hat{r}_L)} = -L\left(\sum_{i=1}^{T_L} i \frac{r_L}{(1+\hat{r}_L)^{i+1}} + T_L \frac{1}{(1+\hat{r}_L)^{T_L+1}}\right). \tag{26.2}$$

We now define the duration as the elasticity of the loan value and the current loan rate:

$$D_L = -\frac{\partial V_L}{\partial (1 + \hat{r}_L)} \frac{1 + \hat{r}_L}{V_L}. \tag{26.3}$$

An alternative interpretation of the duration is that it measures the average time the bank needs to recover its initial outlay of L from the payments that are made on the loan. To see this, define the present value of future payments the bank gets, relative to the current value, as $\omega_i = \frac{r_L L}{(1+\hat{r}_L)^i} \frac{1}{V_L}$ and include the repayment of the face value at $i = T_L$ in the ω_{T_L}. Thus we can rewrite the duration as

$$D_L = \sum_{i=1}^{T_L} i\omega_i, \tag{26.4}$$

which will represent the weighted time, where the weights are given by the present value of the payments made on the loan.

Similarly we obtain for deposits that its duration is given by

$$D_D = -\frac{\partial V_D}{\partial (1 + \hat{r}_D)} \frac{1 + \hat{r}_D}{V_D}, \tag{26.5}$$

where V_D is the value of deposits and $\hat{r}_D$ the current deposit rate.

Let us now consider the value of loans $\tau > 0$ time periods into the future. The value of the loan will have grown by the accumulated interest due and this increased value then needs to be discounted to the present value, thus we obtain this future value as

$$V_L^\tau = V_L \frac{(1 + r_L)^\tau}{(1 + \hat{r}_L)^\tau} \tag{26.6}$$

and one time period further ahead this value, prior to interest being paid, increases by the current loan rate, such that

$$\hat{V}_L^\tau = (1 + \hat{r}_L) V_L^\tau. \tag{26.7}$$

This

$$\begin{aligned} \frac{\partial \hat{V}_L^\tau}{\partial (1 + \hat{r}_L)} &= \frac{\partial V_L}{\partial (1 + \hat{r}_L)} \frac{(1 + r_L)^\tau}{(1 + \hat{r}_L)^{\tau-1}} + (1 - \tau) V_L \left(\frac{1 + r_L}{1 + \hat{r}_L}\right)^\tau \\ &= V_L^\tau \left((1 - \tau) - D_L\right), \end{aligned} \tag{26.8}$$

using equations (26.5) and (26.6). If we now let the time period ahead reduce such that $\tau \to 0$, then $V_L^\tau \to V_L$ and equation (26.8) becomes

$$\frac{\partial V_L^\tau}{\partial (1 + \hat{r}_L)} = V_L (1 - D_L). \tag{26.9}$$

For deposits we obtain similarly

$$\frac{\partial V_D^\tau}{\partial (1 + \hat{r}_D)} = V_D (1 - D_D). \tag{26.10}$$

As the value of a bank (equity) is given by the difference between the value of the assets (loans) and liabilities (deposits), $E = V_L - V_D$, the bank could immunize itself from movements of the interest rates if $\frac{\partial V_L}{\partial(1+\hat{r}_L)} = \frac{\partial V_D}{\partial(1+\hat{r}_D)}$ as then any changes in the value of assets and liabilities cancel each other out. Assuming (approximately) equal interest rates, this implies from equations (26.9) and (26.10) that we need

$$V_L\,(1 - D_L) = V_D\,(1 = D_D)\,. \tag{26.11}$$

As typically the value of equity is negligible compared to the value of assets and liabilities and hence $V_L \approx V_D$, this result implies that to eliminate interest rate risk we require that $D_L \approx D_D$. The net exposure to interest rate risk is then given by $D_L - D_D > 0$. As deposits are typically short-term and loans more long-term, interest rate risk cannot easily be eliminated as we will normally find that $D_L \gg D_D$. It is also not the nature of banks to eliminate interest rate risk by adjusting the maturity of liabilities and/or assets, given their central role in the transformation of short-term, deposits into long-term loans.

Banks might find, however, that their exposure to interest rate is too large and that they want to reduce this exposure. While they can use derivatives to hedge such risks, in particular swaps, they might also seek to adjust the duration of their assets and liabilities. When providing loans, banks can seek to provide more short-term loans or loans with interest rates fixed for short time periods or even variable interest rates; this would reduce the duration of loans and thus close the gap between the duration of assets loans and liabilities, reducing the net risk exposure to interest rate risk. Short-term loan can be promoted by offering more attractive loan rates for such loans compared to loans with longer fixed interest rates. Similarly, banks may provide more generous deposit rates if these are fixed for a longer time period, increasing the duration of liabilities.

An alternative way to reduce the duration of assets is to reduce the amount of loans and retain a larger proportion of funds as cash reserves, which have a very short duration. On the other hand, banks might reduce their reliance on short-term deposit by increasing the amount of equity they hold, which has a very long duration. Both measures will reduce the duration gap between long-term assets and short-term liabilities.

Reading Keiding (2016b, Chapter 3.3)

26.2 Demand for cash reserves

Banks typically retain a proportion of deposits as cash reserves. These funds are not always held in cash or central bank funds but might be invested into other highly liquid and low-risk securities, such as treasury bills; for this reason cash reserves are often referred as liquidity reserves, but for convenience we retain the term cash reserves here. Such cash reserves can be held to allow the withdrawal of deposits as part of the liquidity insurance banks provide, which was discussed in chapter 4.1, but they can also serve additional purposes as we will discuss here.

The withdrawal of deposits for consumption purposes is not as deterministic as models suggest, but instead these are more stochastic and at least in parts difficult to predict. Therefore, to account for the possibility of larger than expected deposit withdrawals, banks will hold additional cash reserves. We will see in chapter 26.2.1 how banks optimally protect themselves against such adverse deposit withdrawals and chapter 26.2.2 will show how such cash reserves are determined if the bank wants to avoid any penalty rates when raising additional cash reserves; chapter 26.2.3 then explores the opportunity cost of retaining cash. However, cash reserves are not only used to reduce the impact of unexpected deposit withdrawals, but can also be used to induce banks to reduce the risks of the loans they are providing as we will be discussing in chapter 26.2.4. Thus cash reserves provide incentives for banks to engage in the monitoring of loans.

26.2.1 Cash reserves as insurance gainst liquidity shocks

Cash reserves are usually held such that banks can repay depositors that seek to withdraw their funds. We will investigate here under which conditions banks would prefer to hold cash reserves and thus forego profits from providing more loans, in order to prevent the bank failing due to the unexpected withdrawal of deposits. In doing so, we will also consider the case where the government provides a bailout to a failing bank.

Banks provide loans L with a probability that this loan is repaid of π_i; this repayment rate will depend on the state of the economy. If the economy is performing well the repayment rate will be high at π_H and if the economy is performing less well, for example during a recession, the repayment rate will be low at $\pi_L < \pi_H$. The economy is in a good state with probability p and in a bad state with probability $1 - p$. Banks may face an exogenous shock with probability γ that would require banks to make payments of C due to depositor withdrawals; thus banks may want to hold cash reserves of C to be able to make these payments. These shocks are assumed to affect all banks in a banking system equally. If faced with such a liquidity shock, we assume that banks have an alternative investment opportunity that will yield a fixed return of B; this investment will, however, not allow depositors to be repaid.

We assume that this alternative investment fulfills the condition

$$0 < \pi_L (1 + r_L) L - (1 + r_D) D < B < \pi_H (1 + r_L) L - (1 + r_D) D, \qquad (26.12)$$

where r_L denotes the loan rate, r_D the deposit rate and D the amount of deposits. We see instantly that in this scenario the bank would choose B rather than provide loans if the the repayment rate is low, but would prefer to provide loans if the repayment rate is high.

We will now consider the optimal choices by banks on holding cash reserves, first in the vase that bank will not be bailed out by the government.

No bailouts Let us first consider a situation in which banks are not bailed out. If the bank does not hold cash reserves C, it will fail as it cannot repay depositors. Its expected profits are then given by

$$\Pi_B = (1-\gamma)\left(p\pi_H(1+r_L)L + (1-p)\pi_L(1+r_L)L - (1+r_D)D\right). \tag{26.13}$$

The bank only obtains profits if no liquidity shock occurs, $1-\gamma$, and then obtains the loans back at the high and low repayment rates, respectively. From these proceeds the bank repays its depositors.

As depositors are only repaid if the bank does not face the liquidity shock, we get the profits of depositors as

$$\Pi_D = (1-\gamma)(1+r_D)D - D \tag{26.14}$$

and hence if banks can extract all surplus from depositors such that $\Pi_D = 0$, the deposit rate will be determined by

$$1 + r_D = \frac{1}{1-\gamma}. \tag{26.15}$$

If we assume that loans are fully financed by deposits, $L = D$, we easily can transform equation (26.14) to become

$$\Pi_B = ((1-\gamma)(p\pi_H + (1-p)\pi_L)(1+r_L) - 1)D. \tag{26.16}$$

A bank holding cash reserves of C would not fail, even if faced with a liquidity shock, and hence their profits will be

$$\hat{\Pi}_B = \gamma p(\pi_H(1+r_L)L - (1+r_D)D) + \gamma(1-p)B + (1-\gamma)\left(p\pi_H(1+r_L)L + (1-p)\pi_L(1+r_L)L + C - (1+r_D)D\right). \tag{26.17}$$

The first term denotes the profits if a liquidity shock is observed and the high repayment rate is realised; in this case the cash reserves are used and profits realized. If the low repayment rate is realized, banks change from providing loans to the alternative investment and obtain B. the second term shows that in the absence of a liquidity shock, the profits are realized and the cash reserves retained. As the cash reserves banks hold cannot be used to provide loans, we have $L = D - C$.

Depositors are only repaid if the bank does not fail; banks fail if a liquidity shock occurs and the low repayment rate is realized, in the latter case deposits are not repaid due to banks conducting the alternative investment. We thus have the profits of depositors given by

$$\hat{\Pi}_D = (1-\gamma(1-p))(1+r_D)D - D. \tag{26.18}$$

If again banks can extract all surplus from depositors such that $\Pi_D = 0$, the deposit rate will be determined by

$$1 + r_D = \frac{1}{1 - \gamma (1 - p)}. \tag{26.19}$$

Inserting this deposit rate into equation (26.17), we obtain the bank profits as

$$\begin{aligned}\hat{\Pi}_B = &((p\pi_H + (1-p)\pi_L)(1+r_L) - 1) D + \gamma (1-p) B \\ &- ((p\pi_H + (1-\gamma)(1-p)\pi_L)(1+r_L) + 1) C.\end{aligned} \tag{26.20}$$

Banks will retain cash reserves of C if it is more profitable to do so, thus $\hat{\Pi}_B \geq \Pi_B$, which we can solve for

$$C \leq C^* = \frac{p\pi_H (1+r_L) D + (1-p) B}{(p\pi_H + (1-\gamma)(1-p)\pi_L)(1+r_L) + 1} \tag{26.21}$$

Thus if the cash reserves required to meet the withdrawal demand of depositors is not too high, it is optimal for banks to maintain such cash reserves.

The effect of bailouts Banks holding the necessary cash reserves cannot fail and will therefore not need a bailout; only banks that do not hold cash reserves and fail will require a bailout. We assume that bailouts only happen if more than one bank fails and only if this occurs due to a liquidity shock as the government will be concerned about the cumulative impact of multiple failing banks; here we only consider the case of two banks. If the other bank does not face a liquidity shock, then it would only fail if it faces a liquidity shock itself and conducts the alternative investment to obtain B, which requires the low repayment rate to be realized. As the other bank also has a probability of $1 - p$, the low repayment rate occurs for both banks with probability $(1 - p)^2$.

Any bailout needs to guarantee the bank a payment of B, as otherwise with the low repayment rate the bank would rather make the alternative investment than accept the bailout. The bank experiencing a high repayment rate would obtain $\pi_H (1 + r_L) L$, but give up $\pi_L (1 + r_L) L$ in exchange for the bailout. Hence total payouts are $B + p (\pi_H - \pi_L)(1 + r_L) L$. A bailout is provided if both banks fail and hence if both obtain the low repayment rate, the profits consist of the profits without a bailout, $\hat{\Pi}_B$, as given in equation (26.20), and the bailout value if there is a liquidity shock. Hence we have

$$\Pi_B^* = \hat{\Pi}_B + \gamma (1-p)^2 (B + p(\pi_H - \pi_L)(1+r_L) L). \tag{26.22}$$

To prevent a bailout due to a liquidity shock, the banks need to retain cash reserves, which they will if it is more profitable to do so. We therefore require that $\hat{\Pi}_B > \Pi_B^*$. This condition is fulfilled if

$$C \leq C^{**} = \gamma p \frac{\left(\pi_H - (1-p)^2 (\pi_H - \pi_L)\right)(1+r_L) D + (1-p) B}{(p\pi_H + (1-\gamma)(1-p)\pi_L)(1+r_L) + 1}. \tag{26.23}$$

Again, as long as the liquidity shock is sufficiently low, banks will hold cash reserves to prevent their failure. As can easily see that $C^{**} < C^{*}$, the size of the liquidity shock that induces banks to hold cash reserves is smaller in the presence of a bailout.

Similarly, if the other bank does not hold sufficient cash reserves and may fail if a liquidity shock occurs, the bank itself will fail and obtain a bailout if the low repayment rate is realised, $1 - p$, hence the bank's profits are given by

$$\Pi_B^{**} = \Pi_B + \gamma (1 - p) (B + p (\pi_H - \pi_L) L) \tag{26.24}$$

and again $\hat{\Pi}_B > \Pi_B^{**}$ ensures that the bank holds cash reserves C; this condition solves for

$$C \leq C^{***} = \gamma p \frac{\pi_H - (1 - p) (\pi_H - \pi_L)}{(p\pi_H + (1 - \gamma) (1 - p) \pi_L) (1 + r_L) + 1} (1 + r_L) D, \tag{26.25}$$

We now can show that

$$C^{***} < C^{**} < C^{*}, \tag{26.26}$$

where the first inequality is only satisfied if we have $B > p (\pi_H - \pi_L) (1 + r_L) D$.

We thus see that if $C < C^{***}$ the liquidity shocks are sufficiently small such that banks will hold cash reserves and a bailout due to the liquidity shock is not necessary. If $C > C^{**}$, banks will not hold cash reserves if banks are bailed out. In the interim region where $C^{***} < C < C^{**}$, the bank will only hold cash reserves if the other bank does, too; we thus require a coordination of the decisions of the two banks. We furthermore observe that as $C^{**} < C^{*}$, the presence of bailouts reduces the maximal feasible size of the liquidity shock that can be accommodated by banks holding cash reserves.

We only considered bailouts that arise due to the liquidity shock. Given the incentives in equation (26.12), banks would always fail if the low repayment rate prevails as they switch to the alternative investment and thus cannot repay their depositors. In this case, even though banks do not repay their depositors, they are not bailed out, this only happens in the presence of a liquidity shock.

Summary We have thus seen that banks will hold cash reserves to be able to withstand a liquidity shock, provided the liquidity shock is not too large; the reason for holding such cash reserves is for banks to avoid their failure and thus the loss of their profits. If banks are bailed out due to a liquidity shock, the size of the shock that banks are willing to accommodate is reduced due to the bailout received, reducing the incentives to commit cash reserves to the bank's survival. With bailouts only provided if multiple banks are failing, a coordination problem emerges as it would be beneficial if banks do not hold cash reserves and are then bailed out if needed, but if a sufficient number of banks hold cash reserves, then no bailout will occur, while if fewer would hold cash reserves a bailout would benefit all banks.

Reading Ratnovski (2009)

26.2.2 Penalty rates

If a bank does not hold sufficient cash reserves, either to meet their obligations to depositors or for regulatory purposes, they have to raise additional funds quickly. Raising liquidity costly may only be possible at a punitive interest rate, either imposed by the central bank to provide this liquidity, but also in the interbank market where the predicament of the bank might easily be known, or the market for term-deposits from institutional investors. Facing such costs will be taken into account by banks when deciding on the amount they hold as cash reserve.

Let us assume that banks face a random withdrawal of deposits of V, whose distribution $F(\cdot)$ is known. Banks holding cash reserves $R < V$ can access loans at an interest rate $\hat{r}$, which we assume to be a higher rate than the loan rate r_L the bank can obtain from lending out funds to act as a deterrent against holding too little cash reserves. With cash reserves, not all deposits D can be lent out, but only an amount of $L = D - R$. Hence the profits of a bank are given by

$$\begin{aligned} \Pi_B &= \pi \left(1 + r_L\right) \left(D - R\right) - \left(1 + r_D\right) D \\ &\quad - \left(1 + \hat{r}\right) E \left[\max \left\{0; V - R\right\}\right] \\ &= \pi \left(1 + r_L\right) \left(D - R\right) - \left(1 + r_D\right) D \\ &\quad - \left(1 + \hat{r}\right) \int_V^R \left(V - R\right) dF\left(V\right). \end{aligned} \tag{26.27}$$

The first term accounts for the loan repayments, which are made with probability π and the second term is the repayment to depositors. The final term is the expected value of the repayment on the liquidity shortfall, $V - R$, including the punitive interest rate $\hat{r}$.

We can now obtain the optimal cash reserves that optimize the profits of banks by solving the first-order condition

$$\begin{aligned} \frac{\partial \Pi_B}{\partial R} &= -\pi \left(1 + r_L\right) + \left(1 + \hat{r}\right) \int_V^D dF\left(V\right) \\ &= -\pi \left(1 + r_L\right) + \left(1 + \hat{r}\right) Prob\left(V > R\right) \\ &= 0, \end{aligned} \tag{26.28}$$

which solves for

$$1 - F\left(R\right) = Prob\left(V > R\right) = \pi \frac{1 + r_L}{1 + \hat{r}}. \tag{26.29}$$

Thus the optimal cash reserves would increase as the penalty rate $\hat{r}$ increases. Similarly, as the loan rate r_L or probability the loan is repaid, π, increases, the higher opportunity costs of holding cash reserves rather than providing loans, will reduce optimal cash holdings. If we were to assume that cash reserves pay interest, these opportunity costs would decrease accordingly and optimal cash reserves would be increasing.

Thus we can see that if banks face the prospects of a shortage in cash reserve, they will optimally balance the lost revenue from not being able to provide more loans against the costs of obtaining additional cash reserves if and as needed. Therefore, it is unlikely that banks hold sufficient cash reserves be able to meet the uncertain demand of depositors withdrawing, but banks will most likely hold less cash reserves and in the case of a particularly large deposit withdrawal will rely on accessing additional funds, even if this is coming at a high cost.

Reading Freixas & Rochet (2008b, Chapter 8.2.1)

26.2.3 Stochastic cash flows

Banks experience a continuous inflow of deposits, payments received into deposit accounts from other banks, as well as outflow of deposits, payments from deposit accounts to accounts at other banks. While many of these payments are predictable, such as the payments of wages or utility bills, the timing of other payments will be less certain. Such payments might include discretionary spending by consumers or investments by companies. The demands on cash reserves such that banks can make expected payments between banks can be planned ahead of time, however those payments that are not regular and therefore less easily planned require additional cash reserves.

Let us assume that banks each payment a bank receives from another bank is of a fixed size V; furthermore we assume that there banks have to raise additional cash reserves or invest excess cash reserves into other assets M times during the T days we consider. Additional deposits and deposit withdrawals are equally likely and hence on average cash reserves are not changing. Milbourne (1983) shows that an optimal strategy would be for the bank to invest their cash reserves into higher yielding assets if the cash reserves R exceed a certain threshold, $R > \overline{R}$. If $R < \underline{R}$ and thus cash reserves fall below a given threshold, assets need to be liquidated to increase cash reserves. Here we set for simplicity $\underline{R} = 0$ such that cash reserves are only replenished if they are fully used. In both cases, when investing into assets and when selling them, we ensure that after this transaction the cash reserves are returned to a benchmark $R = \hat{R}$. We here assume that each buying or selling of assets imposes costs of c on the bank and investing into cash reserves causes the bank to lose out on returns at a rate of $\pi(1 + r_L)$, the amount that would be returned when providing additional loans.

With M denoting the number of transactions during the T time periods, the daily costs to the bank are then given by

$$\Pi_B = \pi(1 + r_L)(D - E[R]) + E[R] - (1 + r_D) D - c\frac{E[M]}{T} \tag{26.30}$$

where $E[\cdot]$ denotes the expected value. The first term denotes the returns on the loans the bank provides, repaid with probability π including interest r_L, which are based on the amount of deposits held less the cash reserves; the second term denotes the cash reserves, where we for simplicity assume no interest is payable on. From these funds the banks has to repay deposits, including interest r_D, and has to bear the cost of the of investing into assets and releasing invested funds.

The time length between transfer i and $i+1$ is denoted by Δt_i; then we clearly have from the definition of having M payments during T days that

$$\sum_{i=1}^{M} \Delta t_i \leq T < \sum_{i=1}^{M+1} \Delta t_i. \tag{26.31}$$

There will be M transactions until T days have passed, but $M+1$ transactions will only be made after T days. As the time length between transactions are independent if we assume that payments themselves arrive randomly, we have $\sum_{i=1}^{M} \Delta t_i = E[M] E[\Delta t_i]$. Using this result in equation (26.31), we easily get

$$\frac{1}{E[\Delta t_i]} - \frac{1}{T} \leq \frac{E[M]}{T} < \frac{1}{E[\Delta t_i]}. \tag{26.32}$$

If we increase the number of days, that is let T go to infinity, such that the time horizon of the bank becomes very long, then from equation (26.32) we immediately see that

$$\frac{E[M]}{T} = \frac{1}{E[\Delta t_i]}. \tag{26.33}$$

With N payments made or received each day and $\frac{\overline{R}}{V}$ denoting the number of payments received that are required to reach the upper boundary and $\frac{\hat{R}}{V}$ to reach the reset point, we can derive from statistics that

$$E[\Delta t_i] = \frac{\hat{R}\left(\overline{R} - \hat{R}\right)}{V^2 N} \tag{26.34}$$

as the average time to either reach the upper barrier $\overline{R}$ or the lower barrier $\underline{R} = 0$ from the reset point $\hat{R}$. We can also show that the average reserves are given by

$$E[R] = \frac{\overline{R} + \hat{R}}{3}. \tag{26.35}$$

Inserting equations (26.34) into equation (26.33) and then the result of this as well as equation (26.35) into equation (26.30), we get

$$\begin{aligned} \Pi_B = {} & \pi(1+r_L)D - (1+r_D)D - c\frac{V^2 N}{\hat{R}\left(\overline{R} - \hat{R}\right)} \\ & - (\pi(1+r_L) - 1)\frac{\overline{R} + \hat{R}}{3}, \end{aligned} \tag{26.36}$$

and the costs are minimized if

$$\frac{\partial \Pi_B}{\partial \hat{R}} = c \frac{V^2 N}{\hat{R}\left(\overline{R} - \hat{R}\right)^2} - \frac{\pi\left(1 + r_L\right) - 1}{3} = 0, \tag{26.37}$$
$$\frac{\partial \Pi_B}{\partial \overline{R}} = c \frac{V^2 N}{\hat{R}^2\left(\overline{R} - \hat{R}\right)^2}\left(\overline{R} - 2\hat{R}\right) - \frac{\pi\left(1 + r_L\right) - 1}{3} = 0,$$

which solves for

$$\hat{R} = \sqrt[3]{\frac{3}{4}\frac{cV^2 N}{\pi\left(1 + r_L\right) - 1}}, \tag{26.38}$$
$$\overline{R} = 3\hat{R}.$$

Using this result we get that from equation (26.35) that the average amount of cash held is

$$E\left[R\right] = \frac{4}{3}\sqrt[3]{\frac{3}{4}\frac{cV^2 N}{\pi\left(1 + r_L\right) - 1}}. \tag{26.39}$$

We see that the amount of cash reserves is increasing in the costs of each transaction, c, and decreasing the higher the lost interest on other investments is, r_L or the more likely loans are being repaid, π. More payments on each day, N, makes the reaching of either boundary more likely and hence more cash reserves are being held. Finally, larger payments, V, also make reaching these boundaries more likely, increasing cash reserves.

We see from equation (26.38) that the reset point for cash reserves, $\hat{R}$, is not in the middle of the upper boundary $\overline{R}$ and the lower boundary $\underline{R} = 0$, but rather it is below this midpoint. The reason for this is the fact that the costs of transferring cash reserves into other assets and releasing additional cash reserves, the third term in equation (26.36), is quadratic in $\hat{R}$ with the minimum at $\hat{R} = \frac{1}{2}\overline{R}$. The lost return due to holding cash reserves, the final term in equation (26.36), is linear in $\hat{R}$, though. As at $\hat{R} = \frac{1}{2}\overline{R}$ the marginal costs for the transfers between assets and cash reserves are zero as the second line in equation (26.37) shows. As the marginal effect for the lost return is positive, it is optimal to reduce $\hat{R}$ below $\frac{1}{2}\overline{R}$ as this reduces overall costs. This is because just below $\frac{1}{2}$ the transfer costs are only marginally increasing while those of the lost returns are reducing. At $\hat{R} = \frac{1}{3}\overline{R}$ these two effects balance each other out.

This ratio is independent of any other parameters in our model. The reason can again be found in the costs banks face, equation (26.36). As the optimal thresholds balance the transfer costs of the third term and the lost returns of the final term, we see that in the transfer term $\hat{R}$ and $\overline{R} - \hat{R}$ have equal importance. For the final term we have $\overline{R} + \hat{R} = \left(\overline{R} - \hat{R}\right) + 2\hat{R}$ and hence $\hat{R}$ has twice the weight of $\overline{R} - \hat{R}$. This weight is independent of any other parameters and thus in the optimal solution this will be reflected in that $\overline{R} - \hat{R} = 2\hat{R}$, or $\overline{R} = 3\hat{R}$.

Banks are holding cash reserves, in excess of those reserves they know that are required for anticipated withdrawals, in order to reduce the costs of having to raise additional cash reserves, the costs c. Such costs might include any punitive interest when obtaining a loan from the central bank or the interbank market, rather than having to liquidate existing assets. They will limit the amount of cash held as they lose out on the higher returns of providing loans, even though making such loans is also costly, for example the assessment of the creditworthiness of the borrower. Banks hold cash reserves that optimally balance these costs against the lost revenue from not providing additional loans.

Reading Miller & Orr (1966)

26.2.4 Cash as an incentive to monitor loans

Banks may retain some of the funds they have obtained from deposits and equity as cash reserves to reduce their exposure to the risks of loans. Having such a smaller exposure might provide banks with incentives to monitor loans in order to reduce the risks they pose. The reduced profitability of the bank due to holding larger cash reserves can make it optimal for banks to monitor loans and thereby secure their profitability, despite the higher costs.

We assume that there are two possible states of the economy, the economy might be performing well, H, or the economy might not be performing well, L, with associated probabilities of p and $1 - p$, respectively. A loan L is repaid with probability π_i, depending on the state of the economy, where in a well-performing economy the loan is more likely to be repaid, $\pi_H > \pi_L$. However, the bank can exert effort of monitoring borrowers such that any loans are repaid with certainty; this effort induces costs of c_i on the banks where the costs in a well performing economy are lower, $c_H < c_L$. Banks invest the deposits they have obtained, D, and equity, E, into loans, L, and cash reserves, R, such that $D + E = L + R$.

After learning the state of the economy, banks can decide to reduce the amount of loans. By selling loans with an amount of $\Delta L \leq L$, they generate additional cash $\Delta R = (1 - \lambda)\,\Delta L$; λ denotes the loss from selling loans as these markets are not fully liquid. The cash and loan positions after this adjustment are then $\hat{R} = R + \Delta R$ and $\hat{L} = L - \Delta L = L - \frac{\Delta R}{1-\lambda}$, respectively. If depositors for simplicity do not obtain any interest, the bank profits with the bank monitoring, and thus loans being repaid with certainty, are given by

$$\Pi_B^M = (1 + r_L)\,\hat{L} + \hat{R} - D - c_i\hat{L} \tag{26.40}$$

and without monitoring, loans may not be repaid, such that bank profits become

$$\Pi_B = \pi_i \left((1 + r_L)\,\hat{L} + \hat{R} - D\right) \tag{26.41}$$

as the bank will only obtain its value if the loans are repaid. Banks monitor if it is more profitable to do so, $\Pi_B^M \geq \Pi_B$, which after inserting for $\hat{L}$ and $\hat{R}$ and noting that $L = E + D - R$, gives us

$$\xi_i (E + D) + (1 - \xi_i) R + \left(1 - \frac{\xi_i}{1 - \lambda}\right) \Delta R \geq D, \tag{26.42}$$

where $\xi_i = (1 + r_L) - \frac{c_i}{1 - \pi_i}$ represents the return on the loan after taking into account the monitoring costs c_i.

If $\xi_i > 1$, the coefficients associated with R and ΔR in equation (26.42) are negative, hence holding cash from the outset or liquidating loans to obtain cash makes this constraint more binding and hence does not provide any additional incentive to monitor loans. As the amount invested into loans reduces as we increase cash holdings, and only loans generate returns, cash holdings will be minimized such that $R = \Delta R = 0$. Thus equation (26.42) becomes $\xi_i (E + D) \geq D$, which is always fulfilled with $\xi_i > 1$.

As $\xi_H > \xi_L$ due to $\pi_H > \pi_L$ and $c_H < c_L$, let us now assume that $\xi_H > 1 > \xi_L$, implying from the above that if the economy is performing well, H, no cash should be held. If the economy is not performing well, L, equation (26.42) will be binding as the amount of cash is kept to a minimum. Thus we have after making some transformations that

$$\frac{\xi_L}{1 - \xi_L} E + \frac{1}{1 - \xi_L} \left(1 - \frac{\xi_L}{1 - \lambda}\right) \delta R_L = D - R_L, \tag{26.43}$$

where ΔR_L and R_L indicate the choices if the economy is not performing well.

If banks exert effort, their ex-ante profits are given by

$$\begin{aligned} \Pi_B &= p\left((1 + r_L) L + R - D\right) + (1 - p)\left((1 + r_L) \hat{L} + \hat{R} - D\right) \\ &\quad - \left(pc_H + (1 - p) c_L\right) L \\ &= (1 + r_L) E + r_L (D - R) - (1 - p)\left(\frac{1}{1 - \lambda} - 1\right) \Delta R_L \\ &\quad - \left(pc_H + (1 - p) C_L\right) L \\ &= \left(1 + \frac{1 + 2\xi_L}{1 + \xi_L} r_L\right) E + \frac{1 - \xi_L (1 - \lambda) - 2\lambda}{(1 - \xi_L)(1 - \lambda)} \Delta R_L \\ &\quad - \left(pc_H + (1 - p) c_L\right) L, \end{aligned} \tag{26.44}$$

where the last equality uses equation (26.43). The profits are determined for the two possible states with their respective probabilities. We clearly see that if $1 - \xi_L (1 - \lambda) - 2\lambda < 0$, the optimal choice for the bank is that $\Delta R_L = 0$ as a positive value would reduce bank profits. Hence, if $1 > \xi_L > \frac{1 - 2\lambda}{1 - \lambda}$ the bank will not sell loans to raise additional cash.

From equation (26.43) we then get that $\frac{\xi_L}{1 - \xi_L} E = D - R$. As $L = E + D - R$, we obtain the optimal amount of lending as

$$L = \frac{1}{1 - \xi_L} E. \tag{26.45}$$

Depositors would retain their money in the bank when learning of a poorly performing economy if liquidating the loans $\hat{L}$, at a loss to be assumed $\hat{\lambda} > \lambda$, and the cash this generates, is less than their deposits, i.e. $\left(1-\hat{\lambda}\right)\hat{L}+\hat{R}<D$. This expression solves for

$$\frac{1-\hat{\lambda}}{\hat{\lambda}}E=\frac{\lambda-\hat{\lambda}}{\hat{\lambda}\left(1-\lambda\right)}\Delta R_L<D-R. \tag{26.46}$$

As $\Delta R_L=0$ and using that $D-R=\frac{\xi_L}{1-\xi_L}$, we easily obtain

$$\frac{1-\hat{\lambda}}{\hat{\lambda}}<\frac{\xi_L}{1-\xi_L} \tag{26.47}$$

or

$$\xi_L>1-\hat{\lambda} \tag{26.48}$$

such that depositors retain their funds with the bank.

Thus, if the return on monitored bank loans is sufficiently low, $\xi_L<1-\hat{\lambda}$, depositors would withdraw their funds and banks could not raise deposits as such behaviour would be anticipated. If $1-\hat{\lambda}<\xi_L<\frac{1-2\lambda}{1-\lambda}$ banks will liquidate loans to generate cash and if $\xi_L>\frac{1-2\lambda}{1-\lambda}$, then no additional cash is generated. In the intermediate range, equation (26.43) becomes

$$\Delta R_L=\frac{(D-R)\left(1-\xi_L\right)-\xi_L E}{1-\lambda-\xi_L}\left(1-\lambda\right). \tag{26.49}$$

As a high value for ΔR_L maximizes the bank profits in equation (26.44), we set $R=0$, allowing the bank to provide the largest possible amount of loans and then, if needed, raise a large amount of cash reserves. Thus while initially no cash is held, it later is generated from the sale of loans. Formally, using equation (26.49), we get from equation (26.44) that

$$\frac{\partial\Pi_B}{\partial R}=-\frac{1-\xi_L\left(1-\lambda\right)-2\lambda}{1-\lambda-\xi_L}+\left(pc_H+\left(1-p\right)c_L\right), \tag{26.50}$$

which we assume to be negative due to the first term being negative; thus we indirectly assume that the monitoring costs c_H and c_L are sufficiently small.

Cash is used here to ensure that banks monitor borrowers and we see from equation (26.42) that is the expected performance of the loans is sufficiently low, a low ξ_i, then holding cash reserves, R, as well as raising additional cash reserves through selling loans, ΔR, makes the condition for monitoring to be profitable for banks more easily fulfilled.

Cash is held as an incentive for banks to monitor the loans they have provided and thereby reduce the risk banks are taking. If the returns on loans are high regardless of the state of the economy, then no cash reserves are required and none are raised at a later stage as the return loans provide are sufficiently high such that monitoring of loans occurs even in the absence of cash reserves. It is if the return on loans in a low-performing economy are sufficiently low that banks will retain some cash and if the low state of the economy is realised, will sell some of these loans to reduce their exposure to the loans. It is this reduced exposure to loans that then incentives banks to monitor their loans and maintain their profitability on this smaller portfolio. It is not necessary to impose minimum cash reserves on banks in the form of regulation as holding cash reserves is optimal for the banks themselves.

Reading Calomiris, Heider, & Hoerova (2015)

Résumé

We have seen that banks optimally balance the lost profits from retaining cash reserves rather than providing loans against the costs of raising additional cash reserves if needed. This leads to banks holding cash reserves as a precautionary measure, even if they are not sure that these are needed. If additional cash reserves are needed, those held may not be sufficient to meet the demand and additional cash reserves need to be raised. It is thus that the cash reserves held are a compromise between the costs and benefits, but will only rarely be exactly the correct amount that is actually needed.

In addition, holding cash reserves can be optimal to provide incentives for banks to monitor the loans they provide and as well as limiting the risk exposure of the bank to such loans by reducing the loan amount due to the cash reserves retained. This is an acting as an insurance against adverse movement in the economy which will reduce the repayments from the loans provided. Therefore using cash reserves allows banks to limit the risks they are taking to an optimal level.

26.3 Equity holdings

When discussing the holding of equity in banks, this is commonly approached as an assessment of the regulation of equity rather than the amount of equity that would be optimal for banks. However bank equity can do more than merely reduce the risk of a bank failing, either through the increased cushion against losses or the incentives to change the quality of loans that are provided. We will discuss in chapter 26.3.1 how banks can use equity to distinguish themselves as a bank providing low-risk loans and thus posing a low risk to depositors from banks that take higher risks and are therefore more risky to depositors. These banks use their holding of equity as a signal to convey to depositors these risks.

But equity cannot only be used to convey risks, banks may also change their policy on which types of loans they provide as we will discuss in chapter 26.3.2. There banks with higher amounts of equity seek to reduce the risks they take by diversifying their portfolio of loans. The larger losses banks with higher equity face if loans are not repaid, will affect the amount of risks a bank is willing to take. We will also investigate in chapter 26.3.3 how government support to the banking system as a whole can affect the leverage of a bank, and thus how much capital is held.

26.3.1 Equity as a signalling device

The role of equity is to provide a cushion against losses from investments, such as the provision of loans, such that creditors, depositors in the case banks, are not facing losses immediately; losses up the size of the equity are covered first by the owners of the bank.n The larger equity is, the more the losses that owners can face, which might well affect the risks they are willing to take. Reversing this logic, the risks a bank is exposed to might affect how much equity they are willing to invest. If the risks a bank are taking cannot be reliably and credibly communicated to the public, then equity might be used as a signal to support their assertion about the level of risk the bank is taking. Risks are not easily assessed and for outsiders of banks any statement by the bank about the level of risk cannot be verifies easily. Therefore using a reliable and observable signal, such as the amount of equity held, will aid this communication.

We assume that banks have a probability π_i of a loan L_i being repaid with interest r_L. There are two types of banks with $\pi_H > \pi_L$, thus one bank provides a low-risk loan with a high probability of it being repaid (H) and the other provides a high-risk loan with a lower probability of it being repaid (L). Banks have equity E and raise deposits $D_i = \kappa_i E$, where κ_i is the leverage. The total amount lent out is financed by deposits and equity, thus $L_i = D_i + E = (1 + \kappa_i)\,E$. The profits of banks are given by

$$\begin{aligned}\Pi_B^i &= \pi_i \left((1 + r_L)\,L_i - \left(1 + r_D^i\right) D_i\right) - E \qquad (26.51)\\ &= \left(\pi_i \left((1 + r_L)\,(1 + \kappa_i) - \left(1 + r_D^i\right) \kappa_i\right) - 1\right) E,\end{aligned}$$

where r_D^i denotes the deposit rate. We assume that banks charge the same loan rate, regardless of the risks involved.

For both banks to attract deposits the returns they offer to depositors must be identical, thus $\pi_H \left(1 + r_D^H\right) = \pi_l \left(1 + r_D^L\right)$, which solves for

$$1 + r_D^L = \frac{\pi_H}{\pi_L}\left(1 + r_D^H\right) > 1 + r_D^H. \qquad (26.52)$$

We have thus established the relationship between the deposit rates of the two banks. In order to prevent banks to switch their type by preferring one type over the other, we require that the profits of both banks are identical, thus $\Pi_B^H = \Pi_B^L$. Solving this expression we obtain

$$\kappa_L = \frac{(\pi_H - \pi_L)\,(1 + r_L) + \pi_H \left(r_L - r_D^H\right) \kappa_H}{\pi_L \left(r_L - r_D^L\right)} > \kappa_H. \qquad (26.53)$$

We thus see that the leverage of the banks providing low-risk loans is lower than that of the bank providing high-risk loans; this is equivalent to a bank with the same total assets holding more equity.

To ensure the bank providing low-risk loans chooses the lower deposit rate r_D^H and leverage κ_H while the bank providing high-risk loans chooses the deposit rate

r_D^L and leverage κ_L, we need that the profits of doing so is higher than the profits choosing the alternative values. We thus require that

$$\begin{aligned}
&(\pi_H\left((1+r_L)(1+\kappa_H)-(1+r_D)\kappa_H\right)-1)E \\
&\qquad \geq \left(\pi_H\left((1+r_L)(1+\kappa_L)-\left(1+r_D^L\right)\kappa_L\right)-1\right)E, \\
&\left(\pi_L\left((1+r_L)(1+\kappa_L)-\left(1+r_D^L\right)\kappa_L\right)-1\right)E \\
&\qquad \geq (\pi_L\left((1+r_L)(1+\kappa_H)-(1+r_D)\kappa_H\right)-1)E,
\end{aligned} \tag{26.54}$$

which gives us

$$\frac{r_L-r_D^L}{r_L-r_D^H}\kappa_L \leq \kappa_H \leq \frac{r_L-r_D^L}{r_L-r_D^H}\kappa_L, \tag{26.55}$$

which can only be fulfilled if

$$\kappa_H = \frac{r_L-r_D^L}{r_L-r_D^H}\kappa_L. \tag{26.56}$$

Let us finally consider that not only do depositors need to receive the same returns on their funds across the two banks, but need to be willing to provide such deposits in the first place. Assume now that by providing low-risk loans the bank can only commit to use a fraction λ of the loan repayments they obtain to meet their depositor demands. This might be due to the possibly of moral hazard by bank managers in appropriating some of the funds available to them. For depositors to provide funds to the bank, the funds available from the loans, taking into account the leverage of the bank, has to exceed the amount the bank commits to repay in case the loans are repaid. Thus we require that

$$\lambda\pi_H(1+r_L)(1+\kappa_H)E \geq \pi_H\left(1+r_D^H\right)\kappa_H E, \tag{26.57}$$

which easily solves for

$$1+r_D^H \leq \lambda(1+r_L)(1+\kappa_H). \tag{26.58}$$

Thus the deposit rate must not exceed a certain value to be credible. If banks are competing for depositors, they will be required to offer the highest possible deposit rate and thus equation (26.58) is fulfilled with equality. If we combine this with equations (26.52), (26.53), and (26.56) we can solve explicitly for the deposit rates and leverages of both bank types.

If we assume that banks cannot credibly communicate which type of loans they have provided, depositors can use the leverage of the banks, or their equity holding, as a signal that provides this information and makes the lower deposit rate of banks providing low-risk loans sustainable. Banks with a low leverage (high equity ratio) will have provided low-risk loans and banks with high leverage (low equity ratio) will have provided high-risk loans; they cannot mimic the leverage and deposit rate of the other type of bank as the condition in equation (26.54) is fulfilled.

Equity may be used by banks to credibly communicate the risks they are taking when providing loans. Thus equity serves as a signal to depositors, and other market participants, whether the bank is providing high-risk or low-risk loans and thus the risk depositors would face.

Reading Biswas & Koufopoulos (2022)

26.3.2 Equity an incentive to diversify

Banks can provide loans to borrowers with similar characteristics, for example by specialising into specific segments of the loan market, or they could provide loans across the entire market, which can be said to be a well diversified portfolio of loans. We will look at the incentives that the provision of equity gives to banks choosing one or the other of these two strategies.

Let us assume that a bank can provide two loans that allow diversification through negative correlation. The first loan yields a return of $\pi_H^1 (1 + r_L)$ if the economy is performing well, H, and nothing if the economy is not performing well, L. The other loan gives a return og $\pi_H^2 (1 + r_L)$ if the economy is performing well and $\pi_L^2 (1 + r_L)$ if the economy is not performing well. We assume $\pi_H^1 > \pi_L^2 > \pi_H^2$ and with deposit rates r_D that $\pi_L^2 (1 + r_L) > 1 + r_D > \pi_H^2 (1 + rL)$. Both loans have the same expected returns as with p denoting the probability of the economy performing well, we assume

$$p\pi_H^1 = p\pi_H^2 + (1 - p)\,\pi_L^2. \tag{26.59}$$

The first loan is more risky than the second loan as the as both loans have the same expected return, but the first loan has a very high or a very low success rate, depending on the performance of the economy, while the second loan's success rate shows less variability as its success rates are always of intermediate value. The returns on two loans are also negatively correlated as the first loan has a high success rate if the economy is performing well, while the second loan has a high success rate if the economy is not performing well.

Banks provide loans to the amount of L_i for loan i; these two loans are fully financed with deposits D and equity E, such that $L_1 + L_2 = D + E$. The profits of the bank are then given by

$$\begin{aligned} \Pi_B &= p\pi_H 1\,(1 + r_L)\,L_1 + \left(p\pi_H^2 + (1 - p)\,\pi_L^2\right)(1 + r_L)\,L_2 \qquad (26.60) \\ &\quad - (1 + r_D)\,D - \frac{1}{2}c\,(L_1 + L_2)^2 \\ &= p\pi_H^1\,(1 + r_L)\,(L_1 + L_2) - (1 + r_D)\,(L_1 + L_2) \\ &\quad + (1 + r_D)\,E - \frac{1}{2}c\,(L_1 + L_2)^2, \end{aligned}$$

using (26.59) and $L_1 + L_2 = D + E$ in the second equation. The first term denotes the expected repayment from the first loan and the second term that of the second loan. From this we deduct the funding costs due to depositors. The final term represents the costs of running the bank, which we assume to be quadratic in the total lending. We clearly see that due to the identical returns the two assets are perfect substitutes and the optimal amount of lending, $L = L_1 + L_2$, is given from the maximization of the bank profits, which gives us the first order condition as

$$\frac{\partial \Pi_B}{\partial L} = p\pi_H^1 (1 + r_L) - (1 + r_D) - cL = 0, \tag{26.61}$$

where we considered only the first loan for simplicity. Hence we get the optimal amount of lending as

$$L^* = \frac{p\pi_H^1 (1 + r_L) - (1 + r_D)}{c}. \tag{26.62}$$

Equation (26.61) assumed that banks cannot fail and deposits are always repaid. In order to ensure that a bank cannot fail, we need to ensure that however the economy is performing, the bank is making profits. We need for the economy performing well and not well, respectively, that

$$\begin{aligned} \Pi_B^H &= \pi_H^1 (1 + r_L) L_1 + \pi_H^2 (1 + r_L) L_2 \\ &\quad - (1 + r_D)(L_1 + L_2 - E) \\ &\geq 0, \\ \Pi_B^L &= \pi_L^2 (1 + r_L) L_2 - (1 + r_D)(L_1 + L_2 - E) \\ &\geq 0, \end{aligned} \tag{26.63}$$

where we neglected the costs of running the bank by interpreting them as sunk costs. Solving equation (26.62) for the size of the second loan, L_2, and inserting into equation (26.63), we get the constraints on the size of the first loan as

$$\begin{aligned} L_1 &\leq \overline{L}_1 \\ &= \frac{\left(p\pi_H^1 (1 + r_L) - (1 + r_D)\right)\left(\pi_L^2 (1 + r_L) - (1 + r_D)\right) + c(1 + r_D) E}{c\pi_L^2 (1 + r_L)}, \\ L_1 &\geq \underline{L}_1 \\ &= \frac{\left(p\pi_H^1 (1 + r_L) - (1 + r_D)\right)\left((1 + r_D) - \pi_H^2 (1 + r_L)\right) + c(1 + r_D) E}{c\left(\pi_H^1 - \pi_H^2\right)(1 + r_L)}. \end{aligned}$$

In order to ensure that a bank would never fail, both conditions needs to be fulfilled and we need to choose $\underline{L}_1 \leq L_1 \leq \overline{L}_1$ for the first loan. As we can show that $\overline{L}_1 < L^*$, the bank would choose to invest into the first loan as well as the second loan; this implies a diversified portfolio of loans. Inserting equation (26.62) into equation

(26.60) gives us the profits of the bank as

$$\Pi_B^* = \frac{\left(p\pi_H^1 (1 + r_L) - (1 + r_D)\right)^2}{2c} + (1 + r_D) E. \tag{26.64}$$

Let us now assume that the bank would default if the economy is not performing well, then equation (26.60) for the bank profits would change to

$$\begin{aligned} \Pi_B = p \Big(\pi_H^1 (1 + r_L) L_1 + \pi_H^2 (1 + r_L) L_2 \\ - (1 + r_D)(L_1 + L_2 - E)\Big) - \frac{1}{2} c (L_1 + L_2)^2 . \end{aligned} \tag{26.65}$$

In this case the bank would generate profits to their owners only if the economy is performing well, which happens with probability p. As by assumption $\pi_H^1 > \pi_H^2$, the bank would only invest into the first loan, thus $L_2^* = 0$, hence the optimal loan amount is given from the first order condition

$$\frac{\partial \Pi_B}{\partial L_1} = p \left(\pi_H^1 (1 + r_L) - (1 + r_D)\right) - cL_1 = 0, \tag{26.66}$$

which solves for

$$L_1^* = \frac{p \left(\pi_H^1 (1 + r_L) - (1 + r_D)\right)}{c} > L^*, \tag{26.67}$$

where the last inequality arises from comparison with equation (26.62) and noting that $\pi_H^1 > \pi_H^2$. Inserting this result back into equation (26.65), we get the bank profits as

$$\Pi_B^{**} = \frac{p^2 \left(\pi_H (1 + r_L) - (1 + r_D)\right)^2}{2c} + p (1 + r_D) E. \tag{26.68}$$

Banks prefer the undiversified larger loan, that would risk the bank failing if the economy is not performing well, provided that $\Pi_B^{**} > \Pi_B^*$. This condition solves for

$$E < E^* = \frac{2p\pi_H^1 (1 + r_L) - (1 + p)(1 + r_D)}{2c}. \tag{26.69}$$

Hence if banks hold small amounts of equity, they have an incentive to provide larger loans, more risky loans, and to not diversify their holdings, increasing the overall risk of the bank failing. As we had assumed that $\pi_H^1 > \pi_L^2$, a bank cannot fail if the economy is performing well and survive if the economy is not performing well and hence we can ignore this case.

Banks will provide more risky loans and do not diversify their portfolio in order to benefit from the higher returns these loans generate to the bank. The losses in form of the loss of equity if the loan is not repaid are outweighed by the profits if the loan is repaid. Once equity increases, these losses become too large to exceed the benefits if the loan is repaid and banks change towards providing smaller, less risky loans and diversify to reduce their losses, even if this comes at the price of reduced profits if the loans are repaid.

We can conclude that banks with a low amount of equity have an incentive to provide more risky and larger loans, as well as not diversify their loan portfolio. An increase of capital (equity) requirement above E^* would ensure that banks reduce the risks and hence the failure of their bank.

Reading Challe, Mojon, & Ragot (2013)

26.3.3 Government subsidies and leverage

Banks might be in receipt of government support in a variety of forms, free deposit insurance, central bank funding at low costs, loan guarantees, subsidies to provide essential services in rural, or poorer areas. Such support, call subsidies here, may affect the decision-making of banks and we will be looking at the impact such subsidies have on the leverage of banks, and hence indirectly on the amount of capital they are holding.

Let us assume that there are two types of banks, small banks and large banks. We have an infinite number of small banks who overall have a market share of $1 - \mu$, with the remaining fraction μ of the market shared by N larger banks. The total lending by all banks is fixed at L and thus shared between banks according to their individual market share μ_i. These loans are identical and repaid with probability π, including interest r_L. Loans are fully financed by deposits D that require payment of interest r_D, and equity E that requires a rate of return r_E. The total repayments on the loan, $\pi\,(1 + r_L)$, are a random variable with a distribution $F(\cdot)$ that is not known in advance. Thus banks do not know with certainty the revenue they will be receiving from the loans they have granted.

Another source of income to banks is a government subsidy of T, that is shared by banks according to their market share μ_i. This subsidy may be the provision of free deposit insurance, access to central bank loans below prevailing market rates, or the provision of guarantees against losses on their loan portfolio..

A bank would default defaults if its income from repaid loans and the government subsidy do not allow to repay depositors in full, thus $\pi\,(1 + r_L)\,L + T < (1 + r_D)\,D$. This condition becomes

$$\pi\,(1 + r_L) < \frac{(1 + r_D)\,D - T}{L}. \tag{26.70}$$

If the bank defaults, it will have to liquidate its assets, and we assume that such a sale yields a fraction λ of the loan value; hence banks raise the amount of $\lambda\pi\,(1 + r_L)\,L + T$. Taking into account the uncertainty about the repayment of the loans, depositors will receive this amount if the bank fails and the condition in equation (26.70) is fulfilled and if the bank does not fail, their deposits will be repaid in full. The depositors will therefore receive payments of

$$\Pi_D = \int_0^{\frac{(1+r_D)D-T}{L}} (\lambda\pi(1+r_L)L+T)\,dF(\pi(1+r_L)) \tag{26.71}$$
$$+\int_{\frac{(1+r_D)D-T}{L}}^{+\infty} (1+r_D)\,DdF(\pi(1+r_L))\,.$$

In the case that the bank fails, the owners of the bank obtain no payment and if the bank does not fail, the bank equity increases by the profits the bank has made. With a required return of r_E which is used as a discount factor, the additional value to the owners of the bank is given by

$$\Pi_E = \frac{\int_{\frac{(1+r_D)D-T}{L}}^{+\infty} (\pi(1+r_L)L+T-(1+r_D)D)\,dF(\pi(1+r_L))}{1+r_E}. \tag{26.72}$$

The total wealth generated by the banks consists of the repaid loans, reduced by any losses from default and the excess payments to shareholders. Thus

$$\Pi_W = \int_0^{+\infty} \pi(1+r_L)\,LdF(\pi(1+r_L)) \tag{26.73}$$
$$-(1-\lambda)\int_0^{\frac{(1+r_D)D-T}{L}} \pi(1+r_L)\,LdF(\pi(1+r_L))$$
$$-\left(1-\frac{1}{1+r_E}\right)\int_{\frac{(1+r_D)D-T}{L}}^{+\infty} (\pi(1+r_L)L+T$$
$$-(1+r_D)D)\,dF(\pi(1+r_L))\,.$$

Using the Leibniz rule, we can obtain the first order condition for the optimal size of the government subsidy that maximizes this wealth as

$$\frac{\partial\Pi_W}{\partial T} = (1-\lambda)((1+r_D)D-T)\,f\left(\frac{(1+r_D)D-T}{L}\right) \tag{26.74}$$
$$-\left(1-\frac{1}{1+r_E}\right)\left(\frac{(1+r_D)D-T}{L}\right)$$
$$=0$$

If we now assume that $x\frac{\partial f(x)}{\partial x}+f(x)>0$, then we obtain

$$\begin{aligned}\frac{\partial^2 \Pi_W}{\partial T^2} = &-(1-\lambda)\left(f\left(\frac{(1+r_D)D-T}{L}\right)\right.\\&\left.+((1+r_D)D-T)\frac{\partial f\left(\frac{(1+r_D)D-T}{L}\right)}{\partial(\pi(1+r_L))}\right)\\&-\left(1-\frac{1}{1+r_E}\right)\frac{1}{L}f\left(\frac{(1+r_D)D-T}{L}\right)\leq 0,\end{aligned} \tag{26.75}$$

$$\begin{aligned}\frac{\partial^2 \Pi_W}{\partial T\partial((1+r_D)D_i)} = &\mu_i\left((1-\lambda)\left(f\left(\frac{(1+r_D)D-T}{L}\right)\right.\right.\\&\left.+((1+r_D)D-T)\frac{\partial f\left(\frac{(1+r_D)D-T}{L}\right)}{\partial(\pi(1+r_L))}\right)\\&\left.+\left(1-\frac{1}{1+r_E}\right)\frac{1}{L}f\left(\frac{(1+r_D)D-T}{L}\right)\right)\geq 0,\end{aligned}$$

where we used that the deposits of each banks reflect their market share, $D_i = mu_iD$, in the second expression. Totally differentiating the first order condition in equation (26.74), we obtain that

$$\frac{\partial^2 \Pi_W}{\partial T^2}dT+\frac{\partial^2 \Pi_W}{\partial T\partial((1+r_D)D_i)}d((1+r_D)D_i)=0$$

and thus

$$\frac{\partial T}{\partial((1+r_D)D_i)} = -\left(\frac{\partial^2 \Pi_W}{\partial T^2}\right)^{-1}\frac{\partial^2 \Pi_W}{\partial T\partial((1+r_D)D_i)}, \tag{26.76}$$

where the first term is negative from the first expression in equation (26.75). If the bank is small, then $\mu_i = 0$ and hence the second term is zero from the second expression in equation (26.75) and hence small banks' decisions to increase leverage, a higher amount of deposits relative to equity for a given amount of loans this bank gives, does not affect the size of the government subsidy this bank receives. In contrast, large banks with $\mu_i > 0$, affect the size of subsidies with a higher leverage increasing its size. Thus indirectly, these subsidies encourage banks to increase their leverage and thus increase the subsidy they receive.

Banks decide their optimal leverage by maximizing the total value of the bank, $\Pi = \Pi_D + \Pi_E$. This is because competition between banks requires them to also generate value to depositors, leading them to maximize the joint profits and then distribute these according to the competitive pressures. Inserting from equations (26.71) and (26.72), we get the first order condition for the optimal amount of repayment to depositors as

$$\begin{aligned}
\frac{\partial \Pi}{\partial\left(\left(1+r_D\right)D_i\right)} &= \frac{\partial T}{\partial\left(\left(1+r_D\right)D_i\right)} \\
&\quad -\frac{1-\frac{\partial T}{\partial((1+r_D)D_i)}}{L}\left(\left(1+r_D\right)D-T\right) \\
&\quad \times f\left(\frac{\left(1+r_D\right)D-T}{L}\right) \\
&\quad -\left(1-\frac{1}{1+r_E}\right) \\
&\quad \times \int_{\frac{(1+r_D)D-T}{L}}^{+\infty}\left(\frac{\partial T}{\partial\left(\left(1+r_D\right)D\right)}-1\right)dF\left(\pi\left(1+r_L\right)\right) \\
&= \frac{\partial T}{\partial\left(\left(1+r_D\right)D_i\right)}\left(\frac{1-\lambda}{L}\left(\left(1+r_D\right)D-T\right)f\left(\frac{\left(1+r_D\right)D}{L}\right)\right. \\
&\quad \left.+\left(1-\left(1-\frac{1}{1+r_E}\right)\left(1-F\left(\frac{\left(1+r_D\right)D-T}{L}\right)\right)\right)\right) \\
&\quad +\frac{\partial \Pi}{\partial\left(\left(1+r_D\right)D\right)} \\
&= 0
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial \Pi}{\partial\left(\left(1+r_D\right)D\right)} &= -\frac{1-\lambda}{L}\left(\left(1+r_D\right)D-T\right)f\left(\frac{\left(1+r_D\right)D-T}{L}\right) \\
&\quad +\int_{\frac{(1+r_D)D-T}{L}}^{+\infty}dF\left(\pi\left(1+r_L\right)\right) \\
&= -\frac{1-\lambda}{L}\left(\left(1+r_D\right)D-T\right)f\left(\frac{\left(1+r_D\right)D-T}{L}\right) \\
&\quad +\left(1-F\left(\frac{\left(1+r_D\right)D-T}{L}\right)\right).
\end{aligned}$$

and $f(x) = \frac{\partial F(x)}{\partial x}$ denotes the density of the distribution of loan repayments. For small banks, $\mu_i = 0$, this first order condition reduces to $\frac{\partial \Pi}{\partial((1+r_D)D)} = 0$ as we have seen above in equation (26.76) that $\frac{\partial T}{\partial((1+r_D)D_i)} = 0$. As with $\frac{\partial T}{\partial((1+r_D)D_i)} \geq 0$, the first term in the first order condition is positive for large banks that have $\mu_i > 0$, and hence if large banks were to use the same leverage as small banks, we would have $\frac{\partial \Pi}{\partial((1+r_D)D_i)>0}$, thus large banks will choose a higher leverage than small banks. As we see from equations (26.75) and (26.76) that $\frac{\partial T}{\partial((1+r_D)D_i)}$ is increasing in the size of banks, μ_i, it is obvious that the larger the bank, the larger the leverage. This is because the higher the leverage of the bank, the higher the government subsidy is and thus they seek to secure a larger subsidy. The higher subsidy reduces the likelihood of bank failures, off-setting the additional risks emerging from a higher leverage.

Define Π_i as the value of bank i, i.e. $\Pi_i = \mu_i \Pi$, we then easily get the effect of a higher leverage, more deposits, of another bank on these profits as

$$\begin{aligned}\frac{\partial \Pi_i}{\partial \left((1+r_D) D_j \right)} = \frac{\partial T}{\partial \left((1+r_D) D_j \right)} & \left((1-\lambda) \frac{(1+r_D) D - T}{L} \right. \\ & \times f \left(\frac{(1+r_D) D - T}{L} \right) \\ & - \left(1 - \left(1 - \frac{1}{1+r_E} \right) \right) \\ & \left. \times \left(1 - F \left(\frac{(1+r_D) D - T}{L} \right) \right) \right), \end{aligned} \tag{26.77}$$

where the term in brackets is positive. If bank the other bank, bank j, is a large bank, then we have from above that $\frac{\partial T}{\partial \left((1+r_D) D_j \right)} > 0$ and thus increasing the leverage of a large bank increases the leverage of all other banks, including small banks. For small banks we had $\frac{\partial T}{\partial \left((1+r_D) D_j \right)} = 0$ and hence them increasing their leverage does not affect the leverage of other banks. We thus see that in the presence of subsidies large banks have a higher leverage than small banks, but their increased leverage increases the leverage of smaller banks, too, compared to the absence of any subsidy.

Subsidies here are not paid specifically to failing banks only to aid their survival, but given as general support to all banks. This is a realistic scenario in times of financial stress as help is given to the banking sector in general, for example by providing free deposit guarantees, or giving emergency loans to banks through the central bank, with such loans available to all banks. In this case the large banks can increase their leverage, and thus the risks they are taking as the cushion against any losses in form of the equity they hold becomes smaller relative to the amount of loans banks provide; they do so in the knowledge that subsidies will be increased due to their higher leverage, causing a form of moral hazard from the provision of such subsidies. While small banks are not affected due to their size, they nevertheless benefit from the higher subsidies, causing them to increase leverage as well. Banks compete for the subsidy and as it is optimally provided such that larger banks obtain a larger share of this subsidy, they will seek to increase their size through a higher leverage.

Reading Dávila & Walther (2020)

Résumé

We have seen that banks can distinguish themselves from other banks that provide more risky loans by holding larger equity. This higher fraction of equity, relative to the amount of loans provided, serves as a signal for to convey the information that they are providing loans with lower risks. Such information cannot be communicated credibly directly to investors or depositors. banks communicating this information credibly and distinguishing themselves from banks that take higher risks benefit from being able to pay lower deposit rates, reflecting their lower risk. This benefit compensates the bank for the higher costs arising from equity; a bank taking on higher risks will find the costs of higher equity outweighs the benefits from mistakenly being identified as a low-risk bank.

Equity is not only held to signal to the market the riskiness of the loans they have provided, it will also provide an incentive to reduce these risks due to the potentially higher losses if loans are not repaid. This would reinforce the role of equity as a signal for low-risk banks as not only does it provide this information for a given risk level, but it gives an incentive to provide loans such that the overall risk to the bank is lower. While banks being subject to minimum equity (capital) requirements has been extensively discussed in the regulatory literature, but we see here that equity may be held voluntarily by banks themselves.

We have finally seen that any government support for the banking system, or any other form of external support, will induce banks to increase their leverage, or, equivalently, reduce the amount of capital they hold. This is so that they can obtain a larger share of the support as they hare larger in size. This competition between banks to obtain the support leads then for all banks to increase their leverage, and thus indirectly increase the risks they are taking.

26.4 Credit risk assessment

The risks banks are taking by providing loans have to be evaluated by the banks and banks have access to different methods to determine these risks. We will here investigate whether a basic methodology that does not distinguish risks properly is preferable or whether it is the use of a more advanced methodology that assesses risks more precisely.

Let us assume there are two possible states of the economy, H with probability p and L with probability $1-p$. A loan L is repaid with interest r_L with probability π_i and financed using deposits D and equity E. The repayment rate in state H is higher than in state L, $\pi_H > \pi_L$ and we can therefore interpret the high state as an economy performing well and the low state of an economy performing less well.

There two types of loans, one is such that idiosyncratic risk is diversified and, depending on the state, the bank obtains $\pi_i\,(1+r_L)\,L$; this represents a given amount that is repaid to the bank and does not expose the bank to any risk. The other loans do not diversify and the bank obtains the full repayment of $(1+r_L)\,L$ with probability π_i and no payment with probability $1-\pi_i$. If providing loans without information, the former are obtained with probability λ and the latter with probability $1-\lambda$. We can therefore interpret diversified loans as being low-risk as they always pay the expected amount, $\pi_i\,(1+r_L)$, and the undiversified loans as being risky as they pay either $1+r_L$ or nothing; having no information on the type of loan due to having a basic credit risk assessment only, implies that loans are given randomly and they can be either diversified, low-risk, or undiversified, high-risk.

We assume that for diversified loans, the bank makes profits only in state H, hence we assume that

$$\pi_H\,(1+r_L)\,L-(1+r_D)\,D \geq 0 \geq \pi_L\,(1+r_L)\,L-(1+r_D)\,D. \tag{26.78}$$

If the non diversified loans are successful, then only in state H does the bank make a profit; we therefore assume that

$$\begin{aligned}\lambda\pi_H(1+r_L)L+(1-\lambda)(1+r_L)L-(1+r_D)D\geq 0 \\ \geq \lambda\pi_L(1+r_L)L+(1-\lambda)(1+r_L)L-(1+r_D)D\end{aligned} \tag{26.79}$$

Finally, we assume that if the non-diversified loans fail, the banks make losses regardless of the state of the economy, therefore we require that

$$\lambda\pi_L(1+r_L)L-(1+r_D)D\leq\lambda\pi_H(1+r_L)L-(1+r_D)D\leq 0, \tag{26.80}$$

where the repayments from the diversified loans are not sufficient to generate a profit for the bank. We can combine these conditions to require

$$\begin{aligned}\frac{1}{\lambda}\frac{1+r_D}{1+r_L}\frac{D}{L} &\geq \pi_H \\ &\geq \frac{1+r_D}{1+r_L}\frac{D}{L} \\ &\geq \frac{1}{\lambda}\left(\frac{1+r_D}{1+r_L}\frac{D}{L}-(1-\lambda)\right) \\ &\geq \pi_L.\end{aligned} \tag{26.81}$$

We can now assess the impact the sophistication of the risk assessment has on the socially preferred choice of loan provision.

Basic risk assessment A bank that cannot distinguish between these two types of loans will be successful if the loan is repaid in state H only, hence

$$\begin{aligned}\Pi_B &= p\pi_H(\lambda\pi_H(1+r_L)L+(1-\lambda)(1+r_L)L-(1+r_D)) \\ &= p\pi_H((\lambda\pi_H+(1-\lambda))(1+r_L)-(1+r_D))L \\ &\quad +p\pi_H(1+r_D)E,\end{aligned} \tag{26.82}$$

using that $L = D + E$. The banks makes a profit only if state H occurs, p, and the loan is repaid, π_H; in all other cases the bank fails and due to limited liability, bank owners do no obtain any payments. The bank receives then the repayments from the diversified loans, λ and the undiversified loans, $1-\lambda$.

The social welfare is given by the expected outcome of the loan provision, less an costs c arising from the failing of the bank, which we assume to be quadratic in the loan amount. We consider the cases of both states occurring, p and $1-p$ and note that the bank fails in all cases unless state H occurs and the loan is repaid, $1-p\pi_H$. Thus we have

$$\begin{aligned}\Pi_W &= p\pi_H\left(\lambda\pi_H\left(1+r_L\right)L+\left(1-\lambda\right)\left(1+r_L\right)L-\left(1+r_D\right)D\right) \\ &\quad+\left(1-p\right)\left(\lambda\pi_L\left(1+r_L\right)L+\left(1-\lambda\right)\pi_L\left(1+r_L\right)L\right) \\ &\quad-\left(1+r_D\right)D-\left(1-p\pi_H\right)cL^2 \\ &= \left(\left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)\right. \\ &\quad\left.-\left(1+r_D\right)\right)L+\left(1+r_D\right)E-\left(1-p\pi_H\right)cL^2.\end{aligned} \tag{26.83}$$

The socially optimal loan amount is then given when solving the first order condition

$$\begin{aligned}\frac{\partial\Pi_W}{\partial L} &= \left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)-\left(1+r_D\right)-2c\left(1-p\pi_H\right)L \\ &= 0,\end{aligned} \tag{26.84}$$

which gives us

$$L^* = \frac{\left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)-\left(1+r_D\right)}{2c\left(1-p\pi_H\right)}. \tag{26.85}$$

Inserting this expression back into equation (26.83) we get the social welfare as

$$\Pi_W^* = \frac{\left(\left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)-\left(1+r_D\right)\right)^2}{4c\left(1-p\pi_H\right)}+\left(1+r_D\right)E. \tag{26.86}$$

The optimal lending L^* can be implemented using capital requirements as from equation (26.82) we easily get $\frac{\partial\Pi_B}{\partial L}>0$ when using equation (26.79) and banks would use the largest loan volume possible. If we define the leverage κ through $D=\kappa E$ and hence $L=D+E=\left(1+\kappa\right)E$, we easily obtain the capital requirement to be $\kappa^*=\frac{L^*-E}{E}$.

We can now compare the social optimum and how to implement it with the different levels of sophistication when assessing loan risks.

Advanced risk assessment If banks can distinguish the loan types and they invest into the diversified loans, they make a profit in state H only due to our assumption in equation (26.78) and thus we have these profits given by

$$\hat{\Pi}_B = p\left(\pi_H\left(1+r_L\right)\hat{L}-\left(1+r_D\right)\hat{D}\right). \tag{26.87}$$

Social welfare in this case is given by

$$\begin{aligned}\hat{\Pi}_W &= p\pi_H\left(1+r_L\right)\hat{L}+\left(1-p\right)\pi_L\left(1+r_L\right)\hat{L} \\ &\quad-\left(1+r_D\right)\hat{D}-\left(1-p\right)c\hat{L}^2 \\ &= \left(\left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)-\left(1+r_D\right)\right)\hat{L} \\ &\quad+p\pi_H\left(1+r_D\right)E-\left(1-p\right)c\hat{L}^2,\end{aligned} \tag{26.88}$$

where we have taken into account the losses in state L and the costs of a bank failing. Note that we had assumed that only diversified loans are provided by banks, hence we only consider such loans for the social optimum and the bank profits. The socially optimal loan amount is obtained from the first order condition

$$\frac{\partial \hat{\Pi}_W}{\partial \hat{L}} = (p\pi_H + (1-p)\,\pi_L)\,(1+r_L) - (1+r_D) - 2c\,(1-p)\,\hat{L} \tag{26.89}$$

solving for

$$\hat{L}^* = \frac{(p\pi_H + (1-p)\,\pi_L)\,(1+r_L) - (1+r_D)}{2c\,(1-p)}. \tag{26.90}$$

Inserted back into the social welfare, we easily obtain

$$\begin{aligned} \hat{\Pi}_W^* = {} & \frac{((p\pi_H + (1-p)\,\pi_L)\,(1+r_L) - (1+r_D))^2}{4c\,(1-p)} \\ & + (1+r_D)\,E \\ > {} & \Pi_W^*. \end{aligned} \tag{26.91}$$

The last inequality arises from the fact that $1 - p < 1 - p\pi_H$. Thus it is socially better to use the advanced risk assessment and invest into diversified (low-risk) loans only.

Finally, if investing into non-diversified loans, we know that banks are profitable only if the loan succeeds, hence the bank profits are given by

$$\hat{\hat{\Pi}}_B = (p\pi_H + (1-p)\,\pi_L)\left((1+r_L)\,\hat{\hat{L}} - (1+r_D)\,\hat{\hat{D}}\right) \tag{26.92}$$

and social welfare is given by

$$\begin{aligned} \hat{\hat{\Pi}}_W = {} & ((p\pi_H + (1-p)\,\pi_L)\,(1+r_L) - (1+r_D))\,\hat{\hat{L}} \\ & + (1+r_D)\,E - (1 - (p\pi_H + (1-p)\,\pi_H))\,c\hat{\hat{L}}. \end{aligned} \tag{26.93}$$

The undiversified loan is repaid with the respective probabilities in the two possible states.

Then the socially optimal loan amount is given by solving the first order condition

$$\begin{aligned} \frac{\partial \hat{\hat{\Pi}}_W}{\partial \hat{\hat{L}}} = {} & (p\pi_H + (1-p)\,\pi_L)\,(1+r_L) - (1+r_D) \\ & -2c\,(1 - (p\pi_H + (1-p)\,\pi_L))\,c\hat{\hat{L}} = 0, \end{aligned} \tag{26.94}$$

which yields

$$\hat{\hat{L}}^* = \frac{(p\pi_H + (1-p)\,\pi_L)\,(1+r_L) - (1+r_D)}{2c\,(1 - (p\pi_H + (1-p)\,\pi_L))} \tag{26.95}$$

and the social welfare becomes

$$\hat{\hat{\Pi}}_W^* = \frac{(p\pi_H + (1-p)\pi_L)(1+r_L) - (1+r_D)}{4c(1-(p\pi_H + (1-p)\pi_L))} + (1+r_D)E > \Pi_W^*, \tag{26.96}$$

with the inequality due to $1 - (p\pi_H + (1-p)\pi_L) < 1 - p\pi_H$. Thus it is socially better to use the advanced risk assessment and invest into undiversified (high-risk) loans only. With social welfare higher for investing into diversified and undiversified loans when using the advanced risk assessment rather than using the basic risk assessment, it is always socially optimal to use the advanced risk assessment.

The diversified loans, low-risk loans, are socially preferred over the undiversified loans, high-risk loans, if the social welfare is higher, thus we require that $\hat{\Pi}_W > \hat{\hat{\Pi}}_W$, which using equations (26.91) and (26.96) requires

$$p > p^* = \frac{\pi_L}{1-(\pi_H - \pi_L)}. \tag{26.97}$$

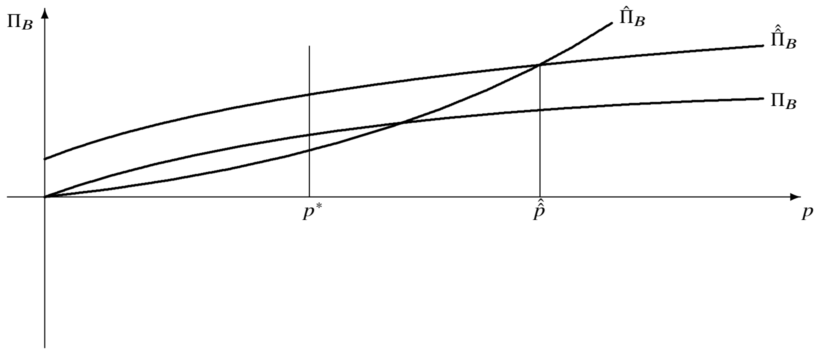
Hence it is socially optimal if banks know the type of loans they invest in and if $p > p^*$ for them to invest into the diversified (low-risk) loans, while for $p < p^*$ they should invest into undiversified (high-risk)loans. This preference arises through the size of the loans that the banks should provide.

Having established that the advanced risk assessment is socially preferred to the basic risk assessment and which type of loan provision is optimal, we can now consider the choices made by banks.

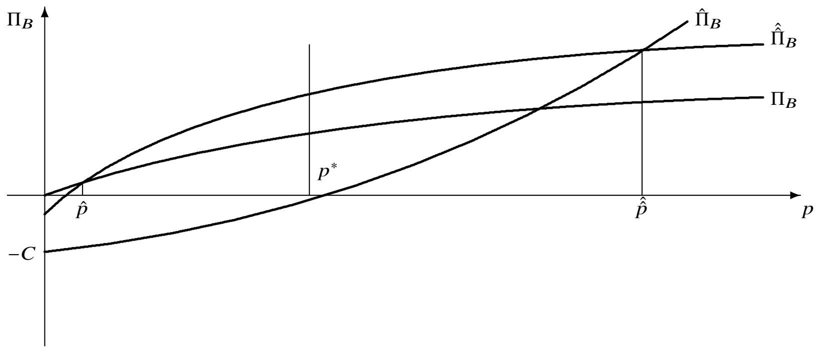
Bank choice Using equations (26.82), (26.87), and (26.92), we can show that for the profits of banks we have $\frac{\partial \Pi_B}{\partial p} \geq 0$, $\frac{\partial \hat{\Pi}_B}{\partial p} \geq 0$, and $\frac{\partial \hat{\hat{\Pi}}_B}{\partial p} \geq 0$; hence the banks in all situations prefer a higher probability of state H to occur. Furthermore we can obtain that for $p = 0$ we have $\Pi_B = \hat{\Pi}_B = 0$ and $\hat{\hat{\Pi}}_B > 0$; in the case that $p = 1$, we have $L^* = \hat{\hat{L}}^*$ and $\hat{L}^* = +\infty$, such that $\hat{\Pi}_B = +\infty$ and $\hat{\hat{\Pi}}_B > \Pi_B$ as inserting these numbers easily shows.

Let us now introduce costs to the bank of C that allow them to use the advanced risk assessment that allows them to distinguish the two types of loans; these costs are assumed to be fixed and will be deducted from the bank profits as shown in equations (26.82), (26.87), and (26.92). These costs would also affect the social welfare by having to be born by banks, but we assume that they would not change the result that the more advanced risk assessment is preferable; the optimal choice between diversified and undiversified loans is not affected as long as we assume that these costs are fixed and not depend on the amount of loans provided. Figure 26.1 illustrates the results for the case of no costs, low costs, and high costs. We will evaluate the resulting bank choices graphically.

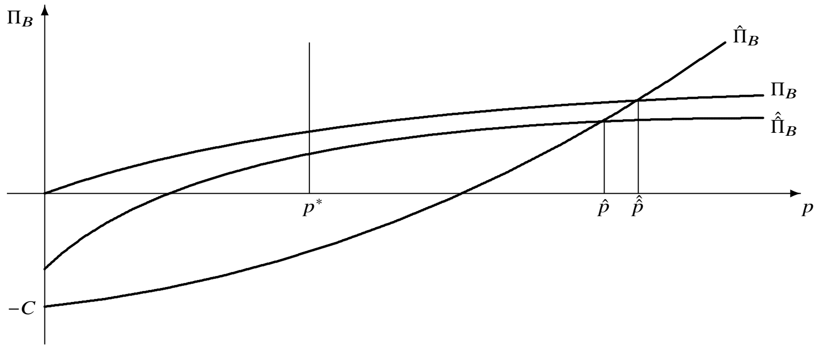
Socially optimal would be to always use the sophisticated risk management and switch to diversified loans for any $p > p^*$ as defined in equation (26.97). As at this point $\hat{\Pi}_W = \hat{\hat{\Pi}}_W$, we can derive that $p^* = p\pi_H + (1-p)\pi_L$ and $\hat{L} = \hat{\hat{L}}$, implying



(a) No costs of advanced risk assessment



(b) Low costs of advanced risk assessment



(c) High costs of advanced risk assessment

Fig. 26.1: Optimal choice of risk assessment regime

that $\hat{\Pi}_B < \hat{\hat{\Pi}}_B$. Hence we see that $p^* < \hat{\hat{p}}$ and the switch to low-risk diversified happens for too high values of p, even with no costs to introduce the advanced risk assessment.

We see that if banks do not face any costs to introduce the advanced risk assessment, they will always employ the more sophisticated risk management system and for $p < \hat{\hat{p}}$ invest into non-diversified (high-risk) loans while for $p > \hat{\hat{p}}$ into diversified (low-risk) loans. The sophistication of the risk assessment banks choose is consistent with the social optimum, but the choice of loans is not socially optimal over a wide range between p^* and $\hat{\hat{p}}$, where banks choose the high-risk undiversified loans but socially optimal would be to choose the low-risk diversified loans.

We see that with low costs of introducing the advanced risk assessment, banks will not employ the sophisticated risk management for $p < \hat{p}$, making their choice socially sub-optimal. Their choice of high-risk undiversified loans extends to an ever higher level of p, the likelihood of state H. Thus with costs increasing the choices by banks with respect to the sophistication of their risk assessment as well as the choice of loan type. With $\hat{p}$ denoting the threshold to apply the advanced risk assessment, we see that while it would not be chosen for all values of p, it will be widely chosen, coinciding with the socially preferred choice and thus allowing banks to make their optimal choice of the loan type.

If we increase the costs for introducing advanced risk assessments even further, banks would use the advanced risk assessment only for very high values of p, thus only if the good state of the economy, H is sufficiently high, making the choice of the socially preferred risk assessment unlikely. This also implies that the bank cannot actively choose their preferred type of loan, which for most parts would be the high-risk undiversified loan. Only once they make use the advanced risk assessment will they be able to make an active choice and by that time choose the low-risk diversified loan.

Summary We have established that banks, when facing additional costs to introduce a more sophisticated risk assessment regime that allows them to differentiate better the risks of loans, will do so only if the likelihood of the economy being in a good state is sufficiently high; this implies that more advanced risk assessments will only be introduced in a well-performing economy, while during recessions banks will prefer to retain a basic risk assessment that does not allow to differentiate risks much. It is not the costs of the more advanced risk assessment that prevent its adoption during a low-performing economy and hence reduced profits of the bank, instead this is driven by the profits of banks, regardless of the state of the economy.

If banks were required to introduce the more advanced risk assessment, then the cost considerations of its introduction can be neglected and the analysis were as in the case of this regime being cost-free. In this case, the choice of the loan type would not be socially optimal at all times, but would be more aligned with the social optimum and in particular, low-risk diversified loans will be taken more frequently than in the case where due to inadequate risk assessment no distinction can be made; loan decisions would be closer to the social optimum.

Reading Feess & Hege (2012)

Conclusions

We have seen that the different durations of assets (loans) and liabilities (deposits) exposes banks to interest rate risk as interest rates fluctuate over time. While such risks can be reduced by reducing the duration of assets through granting more short-term loans, loans with variable interest rates or invest more into cash reserves, as well as increasing the duration of liabilities by attracting more deposits with fixed terms or using more equity, the risk can never be fully eliminated as it is an essential role of the bank to transform short-term deposits into long-term loans.

Ignoring the importance of interest rate risk, cash reserves are held to meet the demands from depositors withdrawing their funds. We have found that banks optimally balance the costs of a shortage of cash reserves against the lost profits from holding more cash reserves due to lower lending. Banks may face punitive interest rates when seeking to raise cash reserves at short notice, or they face transaction costs when raising such funds or investing excess cash reserves and these costs have to be set against the lost revenue from not lending out the cash reserves. A central element in these considerations is that cash reserves fluctuate randomly due to random withdrawals and deposits by depositors. While known and anticipated deposit movements can be accounted for, it is the random fluctuation around such predictable cash requirements that necessitate additional cash reserves. But cash cannot only be used to cover the withdrawal of deposits, the lower profits when lending less due to higher cash reserve can also incentive the bank to ensure that those loans they provide are of low risk, such that these lower profits are not jeopardised by losses.

As much as cash reserves can be used to lower the risks banks are taking, so can the use of more equity to finance loans, replacing deposits, provide an incentive to lower the risks of the loan portfolio the banks hold. The higher loss to the bank from the loss of equity in case of loans defaulting will provide an incentive for banks to lower the risks they are taking. But high equity can also be used to signal to depositors that a bank takes on low risks and if the risks they are taking cannot be directly communicated, depositors can use the leverage of the bank to indicate the level of risk they are taking and hence the deposit rate they require. Thus, as cash reserves, equity might be hold to commit to a provide low-risk loans and thus expose depositors to low risks only.

Risk assessments should be conducted using the most advanced methods available to banks, but we have seen that due to cost reasons, banks will only want to introduce these in times of low risks, namely when the economy is performing well and default rates are overall low. It would, however, be more beneficial if banks were to introduce more advanced risk assessment regimes in times of poor economic performance and higher default rates as then they would be able to choose the risks of the loans they provide more accurately, even if the social optimum will not be reached without further regulatory intervention.

Chapter 27

Remuneration policies

It is not banks that make decision to provide a loan, but employees being employed by the bank. The decisions by employees will be driven by the consequences these decisions have for their personal situation giving rise to the potential for moral hazard in the decision-making. Such moral hazard is addressed by providing employees with remuneration that is structured such that it aligns the interests of the bank and their employees. In this chapter we will look into aspects of remuneration and what consequences they have for the provision of loans by banks.

We will at first in chapter 27.1 assess how managers with different abilities are matched to banks of different sizes, what implications this has for the remuneration of managers, and how any interference with the salaries employees receive might affect the provision of loans by a bank. The impact different forms of remuneration have on the loan risks of banks are explored in more detail in chapter 27.2 and chapter 27.3 will then also assess the impact remuneration can pay in preventing systemic banking crises. In chapter 27.4 look at the relationship between regulators and banks to assess why regulators are only able to attract less able employees than banks. Banks in distress or with low profitability and hence low bonus payments to employees might suffer a brain-drain in that employees leave the bank for other banks. We will discuss in chapter 27.5 how the use of a bonus pool can limit such a brain drain.

27.1 The impact of ability

Employees, and especially managers at banks, will have different abilities to generate profits. This might be due their differences in assessing the risks of companies or the ability to negotiate loan terms that are better for the banks, such as securing higher loan rates or obtaining larger collateral than other managers. This ability of managers would be reflected in their remuneration, but also in the bank they work for.

Let us assume that there are N banks each of different size such that the amount of loans they provide is order with $L_1 \geq L_2 \geq \ldots \geq L_N$. There are also N managers to match with these banks, who have different abilities to generate profits. The profits per unit of loans a manager j is able to obtain is given by

$$\Pi_M^j = \pi_j \left(1 + r_L^j\right) - \left(1 + r_D^j\right) \tag{27.1}$$

with π_j denoting the success rate of loans the manager can generate, r_L^j the interest rate on loans and r_D the funding cost of the deposits used to fund the loan. Manager j working for bank i gets paid a fraction λ_i^j of the profits they generate for their bank, hence the salary of a manager is then given then by

$$w_i^j = \lambda_i^j \Pi_M^j L_i, \tag{27.2}$$

and the profits the bank retains after paying the manager are

$$\Pi_B^{i,j} = \left(1 - \lambda_i^j\right) \Pi_M^j L_i. \tag{27.3}$$

For convenience we assume that $\Pi_M^1 \geq \Pi_M^2 \geq \ldots \geq \Pi_M^N$ and the ability of managers to generate profits is ordered. All managers have an outside option of employment paying them a salary of w^j. Thus total payments of the bank to the manager must meet this outside option to prevent the manager from leaving the bank, hence

$$w_j = \lambda_i^j \Pi_M^j L_i \tag{27.4}$$

for any bank i. A bank would not pay the manager more than their outside option as this would reduce the bank's profits.

Allocation of managers If deciding whether to employ managers j or $j-1$, thus a manager of higher or lower ability, the most bank i is willing to pay the better manager j would be such that the profit the bank makes is identical to the profits it would make when employing the manager with the lower ability, $j-1$. Thus the highest possible salary will be determined such that $\Pi_B^{i,j} = \Pi_B^{i,j-1}$, which we can solve for

$$\lambda_i^{j-1} = 1 - \frac{\Pi_M^j}{\Pi_M^{j-1}} \left(1 - \lambda_i^j\right). \tag{27.5}$$

Similarly, for bank $i-1$ to compete for manager $j-1$, the maximum pay they would consider is such that hiring the better manager j gives them the same profits. $\Pi_B^{i-1,j-1} = \Pi_B^{i-1,j}$. Noting that we can write

$$\begin{aligned}\Pi_B^{i-1,j} &= \left(1 - \lambda_{i-1}^j\right) \Pi_M^j L_{i-1} \\ &= \Pi_M^j L_{i-1} - \lambda_{i-1}^j \Pi_M^i L_{i-1} \\ &= \Pi_M^j L_{i-1} - w_i^j.\end{aligned} \tag{27.6}$$

Using this expression we obtain the maximal pay for manager $j-1$ employed by bank $i-1$ as

$$\lambda_{i-1}^{j-1} = 1 - \frac{\Pi_M^j}{\Pi_M^{j-1}} + \frac{w_i^j}{\Pi_M^{j-1} L_{i-1}} = 1 - \frac{\Pi_M^j}{\Pi_M^{j-1}} \left(1 - \frac{w_i^j}{\Pi_M^j L_{i-1}} \right). \tag{27.7}$$

In order for manager $j-1$ to prefer bank $i-1$ as employer over bank i, we need that $w_{i-1}^{j-1} \geq w_i^{j-1}$. Noting that $\lambda_i^j \Pi_M^j L_i = w_j$, the outside option of the manager, we can transform this expression into

$$\left(\Pi_M^{j-1} - \Pi_M^j \right) (L_{i-1} - L_i) \geq 0 \tag{27.8}$$

As by assumption bank $i-1$ is larger than bank i, $L_{i-1} > L_i$, and the profits generated by manager $j-1$ are larger than those generated by manager j, $\Pi_M^{j-1} \geq \Pi_M^j$, this is always fulfilled. This result implies that a larger bank always attracts better managers. As this needs to be fulfilled for all banks and all managers, the only solution is that $i = j$ and hence the best performing manager will be employed by the largest bank.

Optimal pay Noting that banks and managers are matched, $i = j$, we can rewrite the fraction of profits the manager receives as salary, equation (27.5), as

$$\lambda_{i+1}^i = 1 - \frac{\Pi_M^{i+1}}{\Pi_M^i} \left(1 - \lambda_{i+1}^{i+1} \right). \tag{27.9}$$

Competition for managers means that a smaller bank $i+1$ would offer a manager at the larger bank i the same salary, thus $w_i^j = w_{i+1}^j$. Inserting equation (27.9) into the expression for w_{i+1}^j in equstion (27.4), this relationship can be solved for

$$\lambda_i^i = \frac{\Pi_M^i - \Pi_M^{i+1}}{\Pi_M^i} \frac{L_{i+1}}{L_i} + \frac{\Pi_M^{i+1}}{\Pi_M^i} \frac{L_{i+1}}{L_i} \lambda_{i+1}^{i+1}. \tag{27.10}$$

Iteratively inserting for λ_{i+1}^{i+1} and assuming that $\lambda_N^N = 0$, such that the least able manager receives no salary, this solves for

$$\lambda_i^i = \frac{1}{\Pi_M^i L_i} \sum_{j=i+1}^{N} \left(\Pi_M^{j-1} - \Pi_M^j \right) L_j. \tag{27.11}$$

As $L_j < L_i$ for any $j > i$ by our assumption that banks' sizes are ordered, we can use equation (27.11) to obtain

$$\lambda_i^i < \frac{1}{\Pi_M^i} \sum_{j=i+1}^{N} \left(\Pi_M^{j-1} - \Pi_M^j \right) = \frac{\Pi_M^i - \Pi_M^N}{\Pi_M^i} < 1. \tag{27.12}$$

We observe that the fraction of profits paid to managers are such that they will never take the entire surplus, unless the least able manager does not produce any profits at all, $\Pi_M^N = 0$. The fraction of the profits that are paid to a manager consists of additional profits they generate compared to the least able manager; this can be interpreted as the value-added this manager produces to the bank compared to the worst-performing manager.

Using equation (27.10) we easily get that

$$\frac{\lambda_i^i}{\lambda_{i+1}^{i+1}} = \frac{\Pi_M^{i+1}}{\Pi_M^i} \frac{L_{i+1}}{L_i} + \frac{\Pi_M^i - \Pi_M^{i+1}}{\Pi_M^i} \frac{L_{i+1}}{\lambda_{i+1}^{i+1} L_i} < 1. \tag{27.13}$$

Hence, we observe that more able managers obtain a smaller proportion of the profits they generate than less able managers. However, as more able managers are employed by larger banks, their total pay is still increasing as we can easily show that $w_i^i > w_{i+1}^{i+1}$. Similarly the profits of larger banks are higher as we can easily show that $\Pi_B^i \geq \Pi_B^{i+1}$.

We have thus seen that managers are rewarded for the value they add over the value the worst-performing manager generates. While the fraction of the generated profits they obtain as remuneration reduces, the larger size of the bank employing them increases the total remuneration of better managers. This required that better managers are employed by larger banks, which is the case as larger banks are able to offer managers higher salaries without reducing their own profitability, and thus the largest bank is able to offer the highest salary to the best manager, who will take up this offer, leaving the second-largest banks to employ the second-best manager, and so on until the smallest bank employs the worst=performing manager.

Pay caps Let us now assume simplify the analysis by considering only two banks. We further assume there two asset types the bank can invest into, but that the total assets of both banks, A, are identical and banks might only differ in the allocation between these two assets. As an additional constraint, assume that the total supply of each asset across the two banks is also fixed. With the size of asset k of bank i denoted L_i^k, we have

$$\begin{aligned} L_i^1 + L_i^2 &= A, \\ L_1^k + L_2^k &= A. \end{aligned} \tag{27.14}$$

We can interpret these two assets as two types of loans, one might be a low-risk loan and the other a high-risk loan.

With no direct preferences of the bank for one asset over the other, a symmetric equilibrium would be that $L_i^k = \frac{1}{2}A$ and both banks hold the asset equally. However, this cannot be an equilibrium as the banks compete for the better of the two managers for this asset class. We assume that managers differ in their ability to manage the first assets, but there is no manager with superior skills available for the second asset and hence banks are not competing for any managers in this asset class and no salary is offered. Given the results we have obtained above, if a bank were to increase the asset allocation for one asset, it has an advantage to attract the more able manager for this asset as naturally the other bank would be smaller in that asset class and would only be able to attract the less-able manager. Hence a symmetric allocation of the assets cannot be an equilibrium in this case.

Let us denote by Π_k^i the profits the manager i generates with asset k. With only two banks we know from above that

$$\begin{aligned} \lambda_1^k &= 1 - \frac{\pi_k^2}{\pi_k^1}, \\ \lambda_2^k &= 0, \end{aligned} \tag{27.15}$$

where λ_i^k denotes the payment to manager i in asset k.

The profits for the two banks are now given as

$$\begin{aligned} \Pi_B^1 &= \left(1 - \lambda_1^1\right) \Pi_1^1 L_1^1 + \Pi_2^2 L_1^2, \\ \Pi_B^2 &= \left(1 - \lambda_2^1\right) \Pi_1^2 L_2^1 + \Pi_2^1 L_2^2, \end{aligned} \tag{27.16}$$

where we assume for simplicity that bank 1 obtains the higher allocation of asset 1 and thus the better manager, while asset 2 no salary is payable by our assumption that no able managers are available. Note that we cannot have a situation where a bank holds a larger amount of both assets as we assumed that the total assets of both banks are identical, thus it requires a bank holding more of one asset to hold less of the other asset. We will, however have such an asymmetric allocation with each bank attracting one of the more able managers for one asset and one of the less able managers for the other asset.

Let us now assume that a pay cap is imposed on the manager of asset 1 such that

$$\overline{\lambda}_1 \leq 1 - \frac{\Pi_1^2}{\Pi_2^1}, \tag{27.17}$$

which means that the bank is not able to pay this manager the remuneration that would emerge in equilibrium. Regulators may impose such pay caps with the aim to reduce what they see as excessive pay or they seek to reduce incentives to take risks that may generate large profits by rewarding managers less for taking such risks as they will no longer fully participate in the profits their strategies generate.

Inserting $\overline{\lambda}_1$ for λ_1^1 in equation (27.16), as well as using equation (27.15) and noting that $L_i^2 = A - L_i^1$, we get the profits of the two banks as

$$\begin{aligned}
\Pi_B^1 &= \left(1-\overline{\lambda}_1\right)\Pi_i^1 L_1^1 + \Pi_2^2\left(A - L_1^1\right) \\
&= \left(1-\overline{\lambda}_1\right)\left(\Pi_1^1 - \Pi_2^2\right)L_1^1 + \Pi_2^2 A, \\
\Pi_B^2 &= \Pi_1^2 L_2^1 + \Pi_2^2\left(A - L_2^1\right) \\
&= \left(\Pi_1^2 - \Pi_2^2\right)L_2^1 + \Pi_2^2 A \\
&= \left(\Pi_1^2 - \Pi_2^2\right)A + \Pi_2^2 A - \left(\Pi_1^2 - \Pi_2^2\right)L_1^1 \\
&= \Pi_1^2 A - \left(\Pi_1^2 - \Pi_2^2\right)L_1^1.
\end{aligned} \tag{27.18}$$

In order for banks to be indifferent between assets 1 and 2 being the larger, these two profits need to be identical, $\Pi_B^1 = \Pi_B^2$, which solves for

$$\left(\Pi_i^1\left(1-\overline{\lambda}_1\right) - \Pi_2^2 + \Pi_1^2 - \Pi_2^2\right)L_1^1 + \left(\Pi_2^2 - \Pi_1^2\right)A = 0 \tag{27.19}$$

and hence

$$L_1^1 = \frac{\Pi_1^2 - \Pi_2^2}{\Pi_1^1\left(1-\overline{\lambda}_1\right) + \Pi_1^2 - 2\Pi_2^2}A. \tag{27.20}$$

If the constraint on $\overline{\lambda}_1$ is not binding, the pay for λ_1^1 can be inserted for v to get $L_1^1 = A$, thus the assets are allocated asymmetrically across banks with bank 1 holding the entire first asset and bank 2 holding the entire second asset.

It ia easy to see that L_1^1 is decreasing as $\overline{\lambda}_1$ decreases and thus the pay cap becomes more stringent, hence the more restrictive pay on an asset becomes, the less the bank allocates to it. We therefore find that regulation on pay can affect the size of a business line affecting those banks that employ the most able managers for these business lines; the reduced business at that bank will then be picked up by banks with less able managers. Of course, this can be a desired effect in that banks can be induced by limits on pay to reduce their exposure in some markets and thus force banks to diversify their business and opening markets up to other banks. This will in general also have implications for the risks of banks, such as forcing banks to have less exposure to high risk assets and improved diversification through a more balanced asset structure.

Summary We have shown that large banks attract the best managers due to the ability to pay them more from an increased asset base, although the fraction of profits that are paid out as salary is reducing the more able managers are. A pay cap on managers for a specific asset class or business line that is binding has the effect of reducing the exposure of the bank to that asset class or business line as banks compete less fiercely for managers and thus do not seek such a large exposure to this asset class in order to generate sufficient profits that allows them to pay the high salary the manager would demand. This can have the effect of forcing banks to diversify their business lines and will allow banks with less able managers to enter this market.

Reading Thanassoulis (2014)

27.2 The impact on risk taking

If managers are rewarded for the profits they are taking, this might affect their willingness to take on risks. We will therefore consider different ways managers can be rewarded and evaluate the risk they are taking when making lending decisions.

We assume that banks can invest in one of two types of loans. The first type if loans is safe in that the loan will always be repaid; the rate charged on such a loan is r_L^1 and the interest paid to depositors of such a bank are given as r_D^1; deposits finance loans fully. If the manager is paid w_1, then the expected profits of this bank are given by

$$\Pi_B^1 = (1 + r_L) L - \left(1 + r_D^1\right) L - w_1. \tag{27.21}$$

The second type of loan will only be repaid with probability π and if the loan fails, the bank will fail as well. If the bank fails it will loose all its equity E, managers will be paid their remuneration as agreed and the remaining equity is distributed to depositors. Hence with a loan rate r_L^2, deposit rate r_D^2, and $\overline{w}_2$ denoting the remuneration of the manager if the loan is repaid, we have the expected profits of this bank given by

$$\Pi_B^2 = \pi \left(\left(1 - r_L^2\right) L - \left(1 + r)D^2\right) L - \overline{w}_2\right) - (1 - \pi) E. \tag{27.22}$$

If the loan is not repaid and the bank loses its equity, the remuneration of the manager does not affect the profits generated to the bank.

The expected return for the depositors are, with $\underline{w}_2$ denoting the remuneration of the manager if the loan is not repaid,

$$\Pi_D^2 = \pi \left(1 + r_D^2\right) L + (1 - \pi) \left(E - \underline{w}_2\right). \tag{27.23}$$

Depositors receive the agreed interest if the loan is repaid and it the loan is not repaid they seize the equity of the bank after the manager has been repaid.

Similarly. the payments to managers are $\overline{w}_2$ if the loan is repaid and $\underline{w}_2$ if the loan is not repaid, hence

$$\Pi_M^2 = \pi\overline{w}_2 + (1 - \pi) \underline{w}_2. \tag{27.24}$$

Trivially for the safe investment these become

$$\Pi_D^1 = \left(1 + r_D^1\right) L, \tag{27.25}$$

$$\Pi_M^1 = w_1$$

as depositors are always repaid their deposits and the managers always receives the agreed remuneration.

Fixed salary We assume now that managers are paid a fixed salary of w_1 and w_2, respectively, which does not depend on the outcome of the loan. As $\underline{w}_2 = \overline{w}_2 = w_2$, we can first analyze the deposit rates. For depositors to be indifferent between the two types of banks we need $\Pi_D^1 = \Pi_D^2$ and hence

$$1 + r_D^2 = \frac{1 + r_D^1}{\pi} - \frac{1 - \pi}{\pi} \frac{E - w_2}{L}. \tag{27.26}$$

Depositors need to be indifferent between bank types as otherwise all deposits will be with one of these only, not allowing banks to choose the level of risk they prefer.

For managers to be indifferent between working for either bank of the two banks, we need $w_1 = w_2$. Assuming that the bank choosing the safe loan makes no profits due to competition, $\Pi_B^1 = 0$, we get

$$w_1 = w_2 = \left(1 + r_L^1\right) L - \left(1 + r_D^1\right) L \tag{27.27}$$

when solving from equation (27.21). Managers need to be indifferent between bank types as otherwise managers would only want to work for one of these, not allowing banks to choose the level of risk they prefer as there is no manager for the other bank type.

In order for a bank to prefer the manager taking out the risky loan to the safe loan we need that its profits are higher doing so, hence $\Pi_B^2 \geq \Pi_B^1 = 0$, which when using equations (27.26) and (27.27), solves for

$$1 + r_L^2 \geq \frac{1 + r_L^1}{\pi}. \tag{27.28}$$

Thus if the interest on the risky loan is sufficiently high, the bank would prefer the risky loan to be provided.

Profit sharing It is common that managers are paid a fraction λ_i of the profits they generate. In this case the remunerations become

$$\begin{aligned} w_1 &= \lambda_1 \left(\left(1 + r_L^1\right) L - (1 + r_D) L\right), \\ \overline{w}_2 &= \lambda_2 \left(\left(1 + r_L^2\right) L - \left(1 + r_D^2\right) L\right), \\ \underline{w}_2 &= 0. \end{aligned} \tag{27.29}$$

If the risky loan is not repaid, the bank makes a loss and hence will not pay a salary to their manager.

For deposits to be indifferent between the two types of banks, $\Pi_D^1 = \Pi_D^2$, we need when using equations (27.23) and (27.25) that

$$1 + r_D^2 = \frac{1 + r_D^1}{\pi} - \frac{1 - \pi}{\pi}\frac{E}{L}. \tag{27.30}$$

For managers to be indifferent between the two types of banks we need $\Pi_M^1 = \Pi_M^2$, which using equation (27.29), implies that

$$\lambda_2 = \frac{\left(1 + r_L^1\right) - \left(1 + r_D^1\right)}{\pi\left(\left(1 + r_L^2\right) - \left(1 + 1_D^2\right)\right)}\lambda_1 \tag{27.31}$$

As we assume perfect competition for banks providing the risk-free loan and they will thus make not profits, $\Pi_B^1 = 0$, we obtain $\lambda_1 = 1$.

For the bank granting the risky loan, we get from equation (27.22)

$$\Pi_B^2 = (1 - \lambda_2)\,\pi\left(\left(1 + r_L^2\right)L - \left(1 + r_D^2\right)L\right) - (1 - \pi)\,E \geq 0 = \Pi_B^1. \tag{27.32}$$

We want the profits of the bank making the risky loans to be higher than for the bank making risk-free loans as in this case the bank would prefer providing risky loans. Inserting equations (27.30) and (27.31) into these profits, we obtain

$$1 + r_L^2 \geq \frac{1 + r_L^1}{\pi} \tag{27.33}$$

which is identical to the condition in equation (27.28) and hence the constraint is equally binding when managers are remunerated through a share of the profits they generate or are paid a fixed salary.

Salary paid in debt We can also consider that managers are paid in debt. This is comparable to deferred payment as the debt will only be repaid, and thus the manager receiving their remuneration, if the bank does not fail. This debt which is being given to managers has the same priority as deposits and can thus be interpreted as managers being required to deposit their remuneration with the bank.

In this case, the payment depositors receive when the loan is not repaid reduces as the salary to managers increases the amount to $E + w_2$ and depositors only obtain a diluted fraction $\frac{E}{E+w_2}E$ of the available equity. Thus the payments to depositors become

$$\Pi_D = \pi\left(1 + r_D^2\right)L + (1 - \pi)\,\frac{E^2}{E + w_2}. \tag{27.34}$$

Thus for deposits to be indifferent between the two types of banks, $\Pi_D^1 = \Pi_D^2$, we require that

$$1 + r_D^2 = \frac{1 + r_D^1}{\pi} - \frac{1 - \pi}{\pi}\frac{E}{E + w_2}\frac{E}{L}. \tag{27.35}$$

The payments to managers if the loans are repaid are then $\Pi_M^1 = \left(1 + r_D^1\right)w_1$ and $\Pi_M^2 = \pi\left(1 + r_D^2\right)w_2$, for banks investing into risk-free and risky loans, respectively. For managers to be indifferent between the two types of banks we need $\Pi_M^1 = \Pi_M^2$, which, after using equation (27.35) becomes

$$w_2 = \frac{1 + r_D^1}{\left(1 + r_D^1\right) - (1 - \pi) \frac{E}{E+W_2} \frac{E}{L}} w_1. \tag{27.36}$$

Requiring the bank investing into the safe loan to make zero profits, we get from $\Pi_B^1 = \left(1 + r_L^1\right) L - \left(1 + r_D^1\right) (L + w_1) = 0$, that

$$w_1 = \left(\frac{1 + r_L^1}{1 + r_D^1} - 1 \right) L \tag{27.37}$$

Inserting this result into equation (27.36), we could solve for w_2 and we find that $w_2 > 0$.

The bank prefer the investment into the risky loan if it is more profitable to do so than investing into the risk-free loan, $\Pi_B^2 = \pi \left(\left(1 + r_L^2\right) L - \left(1 + r_D^2\right) (L + w_2) \right) - (1 - \pi) E \geq 0 = \Pi_B^1$, which solves for

$$\begin{aligned} 1 + r_L^2 &\geq \left(1 + r_D^2\right) \left(1 + \frac{w_2}{L}\right) + \frac{1 - \pi}{\pi} \frac{E}{L} \\ &= \frac{1 + r_D^1}{\pi} \frac{L + w_2}{L} + \frac{1 - \pi}{\pi} \frac{E}{L} \frac{w_2}{L} \frac{L - E}{E + w_2}. \end{aligned} \tag{27.38}$$

Differentiating this expression with respect to the w_2 we obtain that $\frac{\partial\left(1+r_L^2\right)}{\partial w_2} = \frac{1+r_D^1}{\pi_L} + \frac{1-\pi}{\pi} \frac{E}{L} \frac{L-E}{L} \frac{E}{(E+w_2)^2} > 0$ and hence the constraint on the loan rate of the risky loan tightens in w_2. As for $w_2 = 0$ this constraint becomes $\frac{1+r_D^1}{\pi}$, identical to the constraint in the case of fixed salaries and profit sharing. As the constraint is increasing as the remuneration of the manager increases, it will require a higher threshold for the loan rate of risky loans than in with fixed salaries or profit sharing.

We thus find that deferring remuneration by requiring salaries to be kept as debt in the bank makes the provision of risky loans less desirable as a higher loan rate on them is required.

Clawback of salary In some instances the salary paid to a manager might be recovered if the loan is not repaid; such a clawback would leave the manager without remuneration. If the loan is repaid, the manager received a fixed salary w_2. Of course, the payments to the manager of a bank choosing risk-free loans is unaffected by clawbacks as these loans are always repaid. We thus have the payments of to the manager as

$$\begin{aligned} \overline{w}_2 &= w_2, \\ \underline{w}_2 &= 0. \end{aligned} \tag{27.39}$$

For depositors to be indifferent between banks providing risky and risk-free loans, we get similar to equation (27.30) that

$$1 + r_D^2 = \frac{1 + r_D^1}{\pi} - \frac{1 - \pi}{\pi}\frac{E}{L}. \tag{27.40}$$

For the bank providing risk-free loans to make no profits due to perfect competition, we need

$$w_1 = \left(1 + r_L^1\right) L - \left(1 + r_D^1\right) L \tag{27.41}$$

and managers are indifferent between the two types of banks if $\pi w_2 = w_1$. For the bank to choose the risky loan we need those profits to exceed the profits of providing the risk-free loan, $\Pi_B^2 \geq \Pi_B^1 = 0$, which after inserting from equations (27.40) and (27.41) becomes

$$1 + r_L^2 \geq \frac{1 + r_L^1}{\pi}. \tag{27.42}$$

This result is identical to equation (27.28) and hence a clawback does not change the incentives to provide risky loans if compared to a fixed salary or a salary based on the profits managers generate.

Summary If the taking of risk by a bank is not desirable, then the best way to incentivise banks to take low risks is to defer pay to managers by requiring to hold their remuneration as deposits in the bank; in this case the constraint for choosing risky loans is the tightest. This is because the dilution for existing depositors increases the deposit rate, making risky loan less attractive and reducing the profits of banks, in addition to exposing managers to the risks they are taking and thus affecting their remuneration. Managers receive the same remuneration in all cases as the level is set by the profits of the banks providing risk-free loans, which we assumed to be zero due to banks being in perfect competition in this market.

Reading Thanassoulis & Tanaka (2018)

27.3 Preventing systemic banking crises

When a bank fails it can affect the prospects of others banks to fail, be it through direct contagion in the interbank market or the sale of assets such that comparable assets at other banks also fall in value. In particular, if a number of banks fail, such defaults by banks can easily spread in what is often referred to as a systemic banking crisis. It is thus that if banks decide to grant risky loans that they increase the risk of a systemic banking crisis emerges. We will explore here how the use of a clawback mechanism for manager salaries can reduce the incentives of managers to choose risky loans such the risk of a systemic banking crisis is reduced.

Banks have a choice between providing a safe and a risky loan. The safe loan is repaid with certainty and hence with a loan rate r_L^1, a deposit rate r_D, loan amount L and managers being paid a wage w_t^1, we have the bank profits given by

$$\Pi_B^{t,1} = \left(1 + r_L^1\right) L - \left(1 + r_D\right) L - w_t^1. \tag{27.43}$$

The risky loan is repaid with probability π and in case of default the bank receives no payments. We assume the resources of the bank to be sufficient to repay their depositors and the manager's salary. Hence with loan rate r_L^2 and manager salary w_t^2, the expected profits of the bank are

$$\Pi_B^{t,2} = \pi \left(1 + r_L^2\right) L - (1 + r_D) L - w_t^2. \tag{27.44}$$

In a banking system with N banks, of which N_2 choose the risky loan, we assume that if $\gamma < \frac{N_2}{N}$, that is the fraction of banks choosing the risky loan exceeds a threshold γ, a systemic banking crisis can occur with probability $1 - p$. This might be due to the need of banks to sell assets in order to repay deposits and the following deterioration of asset values, or direct contagion by not extending loans between banks. If a banking crisis occurs, all banks choosing the risky loan will make losses such that the final payment they obtain is zero. In addition there are costs borne by the public of C, which may be arising from bailouts or deposit insurance company payments. In addition, managers of all banks, even if they chose the safe loan or did not participate in the banking sector in that time period, will have to repay a fraction λ of their previous earnings. We restrict the analysis to 2 time periods of making such choices.

We assume that risky loans are more profitable to the bank than safe loans, this implies that

$$\pi \left(1 + r_L^2\right) L - (1 + r_D) L > \left(1 + r_L^1\right) L - (1 + r_D) L. \tag{27.45}$$

This higher profitability is even true when taking into account the possibility of a banking crisis. If we assume that managers have outside employment opportunities paying $\hat{w}$, these are not as high as to make it unprofitable for banks to employ them choosing safe loans. Finally, we assume that taking into account the social costs C, makes the threat of a banking crisis undesirable. We require that

$$\begin{aligned} & p \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) \\ & > \left(1 + r_L^1\right) L - (1 + r_D) L \\ & > \hat{w} \\ & > p \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) - (1 - p) C. \end{aligned} \tag{27.46}$$

The profits of the bank from providing the risky loan, taking into account the possibility of a banking crisis, exceeds the profits of banks providing the risk-free loan when no banking crisis can emerge, which is sufficient to pay the manager their salary. If we take into account the social costs of a banking crisis, providing risky loans is not desirable as risk-free loans are more beneficial and even the manager salary cannot be paid.

As we consider the clawback of salary from previous time periods, we will consider a model in which there are two time periods; this allows us to consider the incentives to provide risky loans in both time periods. We will solve the model backwards by first considering the incentives in the second time period.

Incentives to choose risk-free loan in the second time period Taking the behaviour of other banks as given, we can assess the manager's remuneration for choosing a safe loan, a risky loan, or seeking alternative employment. Excluding the current bank, the number of banks choosing risky loans are such that the threshold γ is not breached when this bank chooses a risky loan, is our first scenario. We then consider the case that when choosing the risky loan would breach the threshold γ, and finally when the threshold is already breached without the bank choosing the risky loan. Table 27.1 shows the resulting manager remuneration for all three cases.

	$\frac{N_2-1}{N} < \gamma$	$\frac{N_2-1}{N} = \gamma$	$\frac{N_2-1}{N} > \gamma$
Safe loan	w_2^1	w_2^1	$w_2^1 - (1-p)\lambda w_1^i$
Risky loan	w_2^2	$pw_2^2 - (1-p)\lambda w_1^i$	$pw_2^2 - (1-p)\lambda w_1^i$
Outside employment	$\hat{w}$	$\hat{w}$	$\hat{w} - (1-p)\lambda w_1^i$

Fig. 27.1: Remuneration of managers in the second time period

If there is no risk of a systemic banking crisis, the first column of table 27.1, managers will receive their agreed remuneration. If the bank choosing the risky loan could cause a banking crisis, the second column of figure 27.1, then if the bank chooses the safe loan, no systemic crisis can unfold and the managers receives the agreed remuneration; the same is true if he leaves the profession as then the bank will not provide a risky loan. If the bank decides to provide a risky loan, then in case of no systemic crisis, p, the manager receives the agreed remuneration and if a systemic crisis emerges, no remuneration is paid and the manager has to repay his previous salary. Finally if a systemic crisis is possible regardless of the choice of the bank, then if such a crisis emerges, all managers have to repay past remuneration as indicated in the final column of figure 27.1.

In the case that $\frac{N_2-1}{N} > \gamma$, let us assume a bank offers $w_2^2 = \frac{\hat{w}+2\varepsilon}{\pi}$ for managers taking the risky loan and $w_2^1 = \hat{w} + \varepsilon$ to those choosing the safe loan; where ε is some positive number. From the last column in table 27.1 it is obvious that for any $\varepsilon > 0$ the manager would want to choose the risky loan. For the bank to have this choice as optimal we need that its profits from doing so are higher than from the safe loan, thus $p\Pi_B^{2,2} > \Pi_B^{2,1}$ which takes into account the possibility of a systemic crisis. After inserting for our suggest salary w_2^1, this condition becomes

$$p\left(\pi\left(1+r_L^2\right)L - (1+r_D)L\right) - \left(\left(1+r_L^1\right)L - (1+r_D)L\right) > \varepsilon, \qquad (27.47)$$

which will be fulfilled for some sufficiently small ε from our assumption in equation (27.46). Hence all banks will choose risky loans once the threshold γ is breached and we have $N_2 = N$.

Let us now consider the case that $\frac{N_2}{N} < \gamma$ and no banking crisis can happen, regardless of the choice of the bank. If the bank offers $w_2^2 = \hat{w} + \varepsilon$ and $w_2^1 = \hat{w} + \varepsilon$, we see from the first column in table 27.1 that managers prefer to choose risky loans. Banks will choose the same if $\Pi_B^2 > \Pi_B^1$, or

$$\left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) - \left(\left(1 + r_L^1\right) L - (1 + r_D) L\right) > \varepsilon, \tag{27.48}$$

which by virtue of the assumption in equation (27.45) is fulfilled. Note that no systemic crisis can emerge in this case, regardless of the choice of loans by the bank. Thus all banks will want to choose risky loans until $N_2 = \gamma N$.

Neglecting ε in our setting of salaries by assuming it converges towards zero, we propose to set

$$\begin{aligned} w_2^1 &= \hat{w}, \\ w_2^2 &= \frac{\hat{w}}{p}. \end{aligned} \tag{27.49}$$

In order for $N_2 = \gamma N$ to become an equilibrium, we need some banks to choose risky loans and some safe loans. The same must hold for managers. Hence for the marginal bank that would trip the numbers of banks choosing risky loans above γN, we need for the manager that the wage received if choosing the safe loan exceeds the wage if choosing the risky loan; thus

$$w_2^1 \geq \pi w_2^2 - (1 - p) \lambda w_1^i. \tag{27.50}$$

Similarly, for the bank we require that the profits from the provision of a safe loan is higher than for the provision of a risky loan, using equation (27.50), this requires

$$\begin{aligned} &\left(1 + r_L^1\right) L - (1 + r_D) L - w_2^1 \\ &\quad \geq p \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L - w_2^2\right) \\ &\quad \geq p \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) - w_2^1 - (1 - p) w_1^i, \end{aligned} \tag{27.51}$$

or

$$\lambda w_i^1 \geq \frac{p \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right) - \left(\left(1 + r_L^1\right) L - (1 + r_D) L\right)}{1 - p} \tag{27.52}$$

The total amount clawed back from the manager in a banking crisis must be sufficiently large such that there is no incentive for too many banks to choose risky loans.

As this requirement for the claw back depends on the salary in period 1, we will need now to analyze the salaries of managers in the previous time period to establish if this condition can be fulfilled.

Decisions in the initial time period Let us now consider the condition that in both time periods all banks choose the risk-free loans. The managerial remuneration is shown in table 27.2, where we take into account the clawback from time period 1 if a systemic crisis emerges and note that the salary in banks choosing risky loans is only paid if no banking crisis occurs. If the bank provides the safe loan in the first time period and no systemic crisis can occur in the second time period, then the manager receives the agreed salary. If a systemic crisis can occur, then some of that salary will be clawed back. If risky loans are provided in the first time period, then the salary is only paid if no systemic crisis occurs in the first time period, provided it can happen; in the first time period no previous salaries exist and hence none can be clawed back. Any salary is only paid in the second time period if in the first time period no systemic crisis has occurred as with a systemic crisis all salaries are eliminated. We note that the possibility of $\frac{N_2-1}{N} = \gamma$ is not considered as this would imply some banks providing the risk-free loan while others provide risky loans. As we assume that all banks and managers are homogenous, they would either all select risk-free loans or all select risky loans, and hence we can exclude this possibility from our considerations.

With the fact that banks do not pay higher salaries than required, they would not pay salaries in excess of the outside option $\hat{W}$, hence

$$\begin{aligned} (1 - \lambda (1 - p))\, w_1^1 &= \hat{w}, \\ p\, (1 - \lambda (1 - p))\, w_1^2 &= \hat{w}, \end{aligned} \tag{27.53}$$

which can easily be solved for w_1^1 and w_1^2, respectively.

			$t = 2$	
			$\frac{N_2-1}{N} < \gamma$	$\frac{N_2-1}{N} \geq \gamma$
$t = 1$	Safe loan		w_1^1	$(1 - \lambda (1 - p))\, w_1^1$
	Risky loan	$\frac{N_2-1}{N} < \gamma$	w_1^2	$(1 - \lambda (1 - pi))\, w_1^2$
		$\frac{N_2-1}{N} \geq \gamma$	pW_1^2	$p\, (1 - \lambda (1 - p))\, w_1^2$

Fig. 27.2: Remuneration of managers in the first time period

Banks prefer the risky loan if this is more profitable, thus $p\Pi_B^{1,2} > \Pi_N^{1,1}$, which after inserting from equation (27.53) reduces to the first inequality in equation (27.46) and hence banks prefer to choose risky loans. Obviously, banks could choose not to enter the market at all and make no profits. To avoid this situation, we require bank profits to be positive, $\pi\Pi_B^{1,2} > 0$, which after inserting translates to

$$\lambda < \frac{p\left(\pi\left(1 + r_L^2\right) L - (1 + r_D)\, L\right) - \hat{w}}{(1 - p)\left(\pi\left(1 + r_L^2\right) L - (1 + r_D)\, L\right)}. \tag{27.54}$$

Hence we see that for small clawbacks only the equilibrium enabling a banking crisis can be achieved. If we insert for w_1^2, in equation (27.52) we see that the constraint on λ for a high risk equilibrium in time period 2 is less restrictive.

These results imply that for sufficiently large clawbacks, λ, the equilibrium without banking crises is preferable, thus banks provide risk-free loans. To ensure that a bank would not prefer to opt out of lending in time period 1, receiving no payments, and then choose the risky equilibrium in period 2, these profits need to be less than those of the safe equilibrium in both periods, $p\Pi_B^{2,2} < \Pi_B^{1,1} + \Pi_B^{2,1}$. Noting that $w_1^1 = w_2^1 = \hat{w}$ and $w_2^2 = \frac{\hat{w}}{\pi}$, this becomes

$$\pi\left(1 + r_L^2\right) L < \frac{2\left(1 + r_L^1\right) L - (2 - p)\left(1 + r_D\right) L - \hat{w}}{p}. \tag{27.55}$$

Combining equations (27.52) and (27.54), we see that providing risky loans can be avoided if the clawback is sufficiently large, namely

$$\begin{aligned} &\frac{p\left(\pi\left(1 + r_L^2\right) L - \left(1 + r_D\right) L\right) - \left(\left(1 + r_L^1\right) L - \left(1 + r_D\right) L\right)}{(1 - p)\,\hat{w}} \\ &\quad < \lambda < \frac{p\left(\pi\left(1 + r_L^2\right) L - \left(1 + r_D\right) L\right) - \hat{w}}{(1 - p)\left(\pi\left(1 + r_L^2\right) L - \left(1 + r_D\right) L\right)}. \end{aligned} \tag{27.56}$$

For a feasible solution we need to be able to obtain a λ that meets both conditions and hence we require

$$\begin{aligned} &\left|\hat{w} - \frac{1}{2}p\left(\pi\left(1 + r_L^2\right) L - \left(1 + r_D\right) L\right)\right| \\ &\quad < \sqrt{\left(r_L^1 - r_D\right) L - p\left(1 - \frac{1}{4}p\right)\left(\pi\left(1 + r_L^2\right) L - \left(1 + r_D\right) L\right)^2}. \end{aligned} \tag{27.57}$$

Thus the remuneration of the manager needs to be sufficiently large to provide sufficient incentives to become a deterrent to choose the risky loan as the remuneration might be clawed back. This remuneration must also not be too large to avoid managers being willing to take the risk of their remuneration being clawed back as the amount they can keep if a systemic crisis does not emerge is sufficiently high. If the salary of managers sufficiently large, then we choose a clawback level above in the range indicated by equation (27.56) such that the size of a possible clawback is large enough for managers to prefer choosing the risk-free and thus avoid a systemic crisis with certainty.

Summary We have seen that we can use a clawback of the previous time period's salary to provide incentives for managers and banks to choose the type of loans that does not lead to a systemic banking crisis. This is achieved by requiring a sufficiently large clawback of the salary, which disincentivises managers from seeking high risk strategies that might lead to a systemic crisis. Managers would have to be given a sufficiently high salary to ensure their losses are significant enough in the case where a systemic crisis materialises; the salary must not be so high that their reward in the case that no systemic crisis emerges outweighs these losses.

A key difference to other models, such as the model in chapter 27.2, is that previous period's salaries are clawed back, rather than only the salary from the current time period. This additional clawback provides strong intertemporal incentives to avoid banking crises by providing low-risk loans. It is therefore possible for regulators to require banks to include such clawback conditions in the knowledge that this will provide incentives for banks to choose low-risk loans and hence making systemic banking crises less likely.

Reading Aptus, Britz, & Gersbach (2020)

27.4 Bankers and regulators

Managers may work in a bank, but they might also provide their expertise to a regulator supervising banks. If a bank does not employ an able manager, that manager might subsequently work for a regulator and be more efficient than a less able manager in detecting any misconduct. When determining the salaries of managers, banks should take this possibility into account to maximize their own profits.

We assume there are two types of employees, those with high skills, leading to success rate π_H on any loans the bank provides, and those with low skills, leading to a success rate $\pi_L < \pi_H$ on such loans. The fraction of high-skilled employees is p and consequently a fraction $1-p$ are low-skilled employees. An employee of type i employed by the bank receives a salary of w_B^i if the loan is repaid, and w_R^i if he is employed by the regulator and successfully discovers any misconduct within the bank; if the loan is not repaid, a bank employee receives a salary of $\hat{w}_B^i$ and if not detecting any misconduct when working for the regulator, he will receive a salary of $\hat{w}_R^i$. If working in a bank, it may engage in misconduct and gain additional funds of F from this activity, but if such misconduct is detected has to pay a penalty P. The regulator detects the fraud with probability γ. Misconduct can be the infringement of capital or liquidity requirements, the inadequate assessment of risks the bank faces, or organisational deficits that expose the bank to unnecessary operational risk.

Hence for a manager of ability i employed by a bank choosing loans with risk π_i, the salary is given by

$$\begin{aligned}
\Pi_M^i &= \pi_i \max\left\{(1-\gamma)\left(w_B^i + F\right) + \gamma\left(w_B^i + F - P\right); w_B^i\right\} \\
&\quad + (1-\pi_i)\max\left\{(1-\gamma)\left(\hat{w}_B^i + F\right) + \gamma\left(\hat{w}_B^i + F - P\right)\right. \\
&\quad \left. ; \hat{w}_B^i\right\} \\
&= \pi_i w_B^i + (1-\pi_i)\,\hat{w}_B^i + \max\left\{F - \gamma P, 0\right\}.
\end{aligned} \tag{27.58}$$

The first term covers the case where the loan is repaid with his misconduct undetected $(1-\gamma)$ and detected (γ), respectfully. Similarly, the second term covers the case where the loan the bank provides is not repaid. The employee will only engage in misconduct if this is more profitable than not engaging in misconduct and draw his salary without any additional benefits and potential penalties.

If a fraction α_i of the managers with skill level i available in the market work for the bank, then the number of regulators R that successfully identify misconduct is given by

$$
\begin{aligned}
R &= \pi_H p \left(1-\alpha_H\right) + \pi_L \left(1-p\right)\left(1-\alpha_L\right), \\
B &= p\alpha_H + \left(1-p\right)\alpha_L,
\end{aligned}
\tag{27.59}
$$

where the second equation gives the number of banks. The regulator employs the remaining $1-\alpha_i$ managers, of which a fraction p $(1-p)$ are highly (lowly) skilled. We assume that they detect misconduct with the same probability with which loans are repaid, thus their success rate working for a bank and for the regulator is identical.

We then have the detection rate as

$$
\gamma = \frac{R}{B} = \frac{\pi_H p \left(1-\alpha_H\right)\pi_L \left(1-p\right)\left(1-\alpha_L\right)}{p\alpha_H + \left(1-p\right)\alpha_L}. \tag{27.60}
$$

With R denoting the number of regulators successfully identifying misconduct and B the number of banks, this definition of the detection rate assumes that reach regulator can monitor only a single bank.

We assume that working for regulators gives intrinsic benefits of Δ to managers. These benefits might arise from the accumulation of knowledge as well as industry contacts that can be used in the future to secure more senior positions in the industry. Thus we have the salary for managers at regulators given by

$$
\begin{aligned}
\hat{\Pi}_M^i &= \pi_i \left(w_R^i + \Delta\right) + \left(1-\pi_i\right)\left(\hat{w}_R^i + \Delta\right) \\
&= \pi_i w_R^i + \left(1-\pi_i\right)\hat{w}_R^i + \Delta,
\end{aligned}
\tag{27.61}
$$

where the first term reflects a successful regulator and the second term an unsuccessful regulator. Note that we assume that the benefits Δ accrue regardless of the success of the regulator.

Banks are assumed to compete with each other for managers. Their profits are given by

$$
\Pi_B^i = \pi_i \left(\left(1+r_L\right)L - \left(1+r_D\right)D - w_B^i\right) + \left(1-\pi_i\right)\left(-\hat{w}_B^i\right) = 0, \tag{27.62}
$$

reflecting the repayment of loans L including interest r_L less the funding costs of deposits D, r_D, and payment of the salary to the manager, provided the loan is repaid. If the loan is not repaid, the bank only has to pay the salary of the manager. Perfect competition between banks ensures that these profits are zero. Hence we get that the expected salary is

$$
\pi_i w_B^i + \left(1-\pi_i\right)\hat{w}_B^i = \pi_i \left(\left(1+r_L\right)L - \left(1+r_D\right)D\right).
$$

If we assume that in the case of the loan not being repaid, no salary paid to the manager, $\hat{W}_R^i = 0$, then we obtain

$$w_B^i = (1 + r_L) L - (1 + r_D) D. \tag{27.63}$$

Using equation (27.58), the total payoff to the manager is given by

$$\Pi_M^i = \pi_i ((1 + r_L) L - (1 + r_D) D) + \max \{F - \gamma P, 0\}. \tag{27.64}$$

The regulator is concerned about the costs of a useful report. This productivity is the ratio of successful reports from a manager of type i and the expected salary costs for this manager, hence

$$\rho_i = \frac{\pi_i}{\pi_i w_R^i + (1 - \pi_i) \hat{w}_R^i}. \tag{27.65}$$

In order to attract highly skilled regulators we need that the remuneration of regulators exceeds that of managers in banks, $\hat{\Pi}_M^i \geq \Pi_M^i$. Using equations (27.61) and (27.64), this becomes

$$\begin{aligned} &\pi_H w_R^H + (1 - \pi_H) \hat{w}_R^H \\ &\quad \geq \pi_H ((1 + r_l) L (1 + r_D) D) - \Delta + \max \{F - \gamma P, 0\}. \end{aligned} \tag{27.66}$$

Then using equation(27.65), productivity of the highly-skilled regulator becomes

$$\rho_H = \frac{\pi_H}{\pi_H ((1 + r_L) L - (1 + r_D) D) - \Delta + \max \{F - \gamma P, 0\}}. \tag{27.67}$$

Similarly, attracting a lowly skilled manager, will give us its productivity as

$$\rho_L = \frac{\pi_L}{\pi_L ((1 + r_L) L - (1 + r_D) D) - \Delta + \max \{F - \gamma P, 0\}}. \tag{27.68}$$

We can now see that the productivity of low-skilled managers exceeds that of high-skilled managers, $\rho_L > \rho_H$, if

$$(\pi_H - \pi_L) (\Delta - \max \{F - \gamma P, 0\}) > 0. \tag{27.69}$$

As we had assumed that $\pi_H > \pi_L$, we need $\Delta > \max \{F - \gamma P, 0\}$ for this condition to be fulfilled and the regulator will attract low-skilled managers and the bank will obtain the high-skilled managers. Thus, if the benefits from becoming a regulator are sufficiently high, the bank will attract all high-skilled managers, $\alpha_H = 1$, and the regulators all low-skilled managers, $\alpha_L = 0$.

Inserting for α_i equation (27.60), we can rewrite the condition in equation (27.69) as

$$\Delta > \max \left\{F - \frac{1 - p}{p} \pi_L P, 0\right\}. \tag{27.70}$$

This condition is fulfilled if the benefits of misconduct (F) are low, the penalties when misconduct is detected (P) are high, few high-skilled managers exist (p), and the ability of the low-skilled managers are high (π_L). We can interpret Δ as the acquisition of knowledge at the regulator that allows the manager to switch into being high-skilled and thus in the future change from the regulator to the bank.

We have thus established that as long as the intrinsic benefits of working at a regulator are sufficiently high, such a the value of contacts within the industry but also the accumulation of knowledge and expertise, regulators will attract low-skilled employees, while high-skilled employees will join banks directly. This is the result of regulators not directly concerned with the detection of misconduct, but having an emphasis on 'value for money', expressed through the productivity of their employees, which takes into account their detection of misconduct, but also the salaries they are paid. The lower salaries paid to low-skilled employees makes them more attractive to regulators and the future benefits of working for a regulator, will allow an even lower salary, increasing productivity at the regulator.

Reading Bond & Glode (2014)

27.5 Employee retention

Banks that are facing distress may struggle to retain employees as they seek employment at other banks, fearing for the impact on their earnings in the future or even their employments. As it is commonly the most highly-skilled employees that are leaving, their departure may well increase the possibility of the bank becoming distressed. We will investigate here how banks can ensure that employees are not leaving the bank in such circumstances. A common way that boni are distributed to employees is through the use of a bonus pool. The bank pays into a pool and the proceeds of this pool is then shared by the employees. However, if an employee leaves before the pool is distributed, he will not receive any payments from this pool; similarly, employees joining may be included to benefit from the existing pool. As there can be substantial time between the funds allocated to the pool and it being distributed, any employee leaving would do so under loss of his bonus.

Let us now consider a market with two banks and in each banks employees obtain the same bonus. In total there are N employees across the two banks and each bank employs a fraction γ_i of these employees. If each employee receives a bonus of w_i, then the size of the bonus pool will be $\gamma_i N w_i$. Due to bank 1 heading towards distress, employees are considering leaving the bank and if a fraction λ of employees are actually leaving the bank, the bonus pool would be divided between $\gamma_i N - \lambda\gamma_i N = (1-\lambda))\,\gamma_i N$ remaining employees. Thus the bonus each remaining employee can expect is given by

$$\hat{w}_1 = \frac{\gamma_1 N w_1}{(1-\lambda)\,\gamma_1 N} = \frac{w_1}{1-\lambda}. \tag{27.71}$$

If those employees leaving the bank now join the other bank, we assume that they are participating in the bonus pool of that bank. This means that the number of eligible employees increases from $\gamma_2 N$ to $\gamma_2 N + \lambda\gamma_1 N$. If this bank used to pay a bonus of w_2, then the bonus for an employee joining the second bank is given by

$$\hat{w}_2 = \frac{\gamma_2 N w_2}{\gamma_2 N + \lambda\gamma_1 N} = \frac{\gamma_2 w_2}{\gamma_2 + \lambda\gamma_1}. \tag{27.72}$$

An employee would switch from bank 1 to bank 2 if it would pay him a higher bonus, $\hat{w}_2 \geq \hat{w}_1$, which solves for

$$\lambda \leq \lambda^* = \frac{\gamma_2 (w_2 - w_1)}{\gamma_1 w_1 + \gamma_2 w_2}, \tag{27.73}$$

where we assume that $w_2 > w_1$ given the better prospects of bank 2. Thus if less than a fraction λ^* of employees leave bank 1 to join bank 2, the it is rational for an employee to leave, while if more than a fraction λ^* is leaving, it would be better to remain at bank 1; thus in equilibrium we require that $\lambda = \lambda^*$. It is therefore that banks will lose a fraction λ^* of their employees, but not all employees will leave as the unchanged bonus pool at their existing bank is to be shared among fewer employees, increasing the bonus of those remaining, while at the same time the newly arriving employees at the other bank are diluting the bonus pool there, reducing the bonus available to each employee.

In order to entice employees to switch banks, the bank seeking to employ employees of the other bank may offer them a fixed salary w_2^* rather than a bonus that is based on performance while they were not employed. In this case the employee will switch banks if $w_2^* \geq \hat{w}_1$, which becomes

$$\lambda \leq \lambda^{**} = \frac{w_2^* - w_1}{w_2^*}. \tag{27.74}$$

It is easy to see that $\lambda^{**} > \lambda^*$ if $w_2^* > \gamma_1 w_1 + \gamma_2 w_2$, where we used that $\gamma_1 + \gamma_2 = 1$. Thus, if the bank would offer a payment above the average pay of the two banks, the fraction of employees leaving the distressed banks would be higher. This is because there is no dilution of the existing bonus pool in banks 2, switching employees would accept a lower bonus than existing employees in that bank, but are offered a higher payment than at their existing bank, as we had assumed that $w_2 > w_1$; however, the costs of banks 2 would also increase.

We have seen that the use of bonus pools can help the retention of employees in time of low profitability of the bank as those employees leaving will not be entitled to their share of the bonus pool, increasing the bonus payments to those staying with the bank, enticing them to stay. More employees might be leaving a bank if they are offered a fixed bonus from the other bank, as this does not dilute the bonus pool of the bank they are joining. While a bonus pool cannot eliminate that employees switch banks in time of distress, but it can limit the impact it will have by ensuring that a sizeable fraction of employees are retained, assuming the pay differences between banks are not too high.

Reading Hoffmann & Vladimirov (2025)

Conclusions

We have seen that the most able employees are employed by the largest banks as they can offer them the highest salaries and that restrictions on salaries in some areas of banking will lead to banks reducing the size of such activities as they cannot attract the highly-skilled employees they seek. It is thus that large banks, or banks which have business lines which high market shares will employ the most able managers, leaving smaller banks with less able managers. this will then become a reinforcing mechanism in that large banks can grow even more as they employ the most able managers to attract customers and they make better decisions generating more profits from which the business can be expanded further.

Common practices to align the interests of banks and their employees are for employees to participate in the profits they generate or to reduce the remuneration if the employee is not successful. We have seen that these mechanisms are not effective in reducing the risks employees seek to take, but instead it would be beneficial to delay the payment of bonuses and link the value of these bonuses with the prospects of the bank. This might be achieved by retaining bonuses as deposits such that employees do not obtain any bonus if the bank fails in the future. Alternatively, a clawback of bonuses paid in previous time periods if a systemic crisis were to emerge, will also reduce the incentives to take larger risks.

We have finally seen that banks will attract the best managers, while regulators tasked with identifying any misconduct in banks will be left with less able employees. It were the potential benefits of working for a regulator in terms of learning additional skills and developing a network of contacts within the industry that allowed regulators to reduce the salaries and show a higher productivity due to such low salaries. Highly skilled managers are less attractive to regulators as they are requiring a higher salary due to the competition with banks, which will not fully compensate for their higher skills. Thus we are left with a situation where banks employ the best managers and regulators employ less skilled managers.

Chapter 28

Malpractice and misconduct in banks

Most decisions by banks will be uncontroversial and fully legal; however, in some instances there might instances where decisions might be questionable, morally and legally. In some instances such decisions might be the result of ignorance or carelessness, while in other instances banks or their employees might make such decisions deliberately in full knowledge of the legal and moral position. In this chapter we will investigate some of these decisions and also discuss, where appropriate, any measures that banks or regulators can take to prevent their occurrence.

Banks offer a variety of products such as loans and deposits, alongside other services that might beneficial to their customers. However, each contract can have many different specifications, for example a loan might have fixed or variable interest rates, be long-term or short-term, be repayable at maturity or have regular fixed payments, can be repaid early or can only be repaid at maturity, besides many other possible specifications. Not all contract types are suitable for all customers, but banks might have an interest in customers taking one contractual form over another as this is more profitable to them and will therefore seek customers to take out a specific contract, regardless of its suitability. We will address such misselling of loans in chapter 28.1.

On the other hand, customers might want to obtain a loan, even though they are not creditworthy, which could result in bak employees being bribed to grant a loan nevertheless. Such corruption when granting a loan and under which condition it is likely to emerge will be discussed in chapter 28.2. Any losses from loans, whether given to an initially creditworthy borrower or an uncreditworthy borrower, may threaten the stability or even survival of the bank. This gives rise to an incentive to defer realising such losses by extending loans, even though there are no prospects of the loan being repaid, and combine these losses with future profits from other loans and thereby avoid the failure of a bank. Chapter 28.3 considers the conditions under which banks may resort to such practices.

Finally, bank accounts are used to transfer money between different individuals, companies, and organisations, between accounts of the same entity at a different banks in the same and other countries, or exchanged for cash; this might include proceeds of criminal activity or funds that are moved to avoid taxation as well as other restrictions. Such activities are commonly referred to as money laundering. Banks are supposed to monitor the transactions of their customers and report any suspicious activities that indicate the involvement in money laundering. In chapter 28.4 the incentives for banks to conduct such monitoring and report suspicious activities in an efficient way are discussed.

28.1 Misselling of loans

Banks offer their customer a large variety of products and services, not all of which will be suitable to them. A short-term loan, for example, would in most cases not be suitable for a borrower seeking to buy a house, while a loan without the possibility of early repayment might not the suitable for a borrower expecting a large lump sum for which he has no other use. However, selling such products to borrowers might be more profitable to bank than selling another, more suitable product. We will here investigate the incentives for banks to sell borrowers loans that are not suitable for their needs, referred to as misselling.

The manager of bank i can exert effort at cost C and will succeed in developing a new loan specification with probability π. A new loan specification that has been developed does not have to be sold, thus it can be developed but never making any profits for the bank. The loan of amount L is fully financed by deposits at a desposit rate of r_D^i. This type of loan is suitable to a specific borrower with probability p, where this probability is only known when attempting to sell the loan to the borrower. Ex-ante, the value of p has a distribution $F(p)$. If the loan is suitable, the borrower obtains a high utility $\overline{U}$ and a low utility $\underline{U}$ if the loan is not suitable, where $\overline{U} > 0 > \underline{U}$; hence obtaining a suitable loan enhances the utility of the borrower and an unsuitable loan reduces his utility. A bank selling an unsuitable loan will face an additional loss $F \leq \left(1 + r_D^i\right) D$ which may consist of compensation to those who have been missold the loan, loss of market share due to a loss in reputation, as well as fines to the regulator. This fine is less than the costs of financing the loan. Furthermore, banks hold equity $E_i \leq F$, that is not lent out, implying that banks hold other assets such as cash reserves which we do not consider here; the equity a bank holds will not be sufficient to cover the fine from misselling the loan.

First we will consider the incentives of the manager in such a bank to sell a newly developed loan specification to borrowers.

Managerial decision If the manager does not sell a loan, either because he has not developed a loan with a new specification or he decides to not sell the loan he has developed as he deems it unsuitable for the customer, he receives a base salary w_i. If he sells the loan and it is not missold, he receives in addition a bonus $\overline{w}_i$, such that his remuneration is $w_i + \overline{w}_i$, obtained with probability p, i.e. the loan not missold.

If he missells the loan, we assume the manager receives no remuneration. Hence a developed loan will be sold if the remuneration from doing so exceeds that of not selling a loan, $p\left(w_i+\overline{w}_i\right)\geq w_i$. This transforms into

$$p\geq p_i^*=fracw_iw_i+\overline{w}_i. \tag{28.1}$$

The new loan specification will only be developed by the manager if it is profitable to do so, thus the expected remuneration less the costs of development have to exceed the remuneration of not developing the loan, w_i. Hence

$$w_i+\pi\int_{p_i^*}^1\left(p\overline{w}_i-(1-p)\,w_i\right)dF\left(p\right)-C\geq w_i. \tag{28.2}$$

The second term denotes the additional remuneration if the new loan development is successful, which happens with probability π and sold, which requires that $p>p_i^*$. The bonus is paid if the loan is not missold and if it is missold, the base salary w_i is lost.

As banks will keep their costs to a minimum, they will set the remuneration such that equation (28.2) is fulfilled with equality and similarly, equation (28.1) will be fulfilled with equality as the bonus will be kept at the lowest possible value, too. Inserting for the bonus from equation (28.1), we obtain from these two equalities that

$$\begin{aligned} w_i &= \frac{p_i^*C}{\pi\int_{p_i^*}^1\left(p-p_i^*\right)dF\left(p\right)}, \\ \overline{w}_i &= \frac{(1-p)\,C}{\pi\int_{p_I^*}^1\left(p-p_i^*\right)dF\left(p\right)}. \end{aligned} \tag{28.3}$$

We can easily derive how the base salary, w_i, and the bonus, $\overline{w}_i$, change with the ability of the manager to develop a new loan, π, and thne threshold at which it will be sold to customers, p_i^*. We obtain

$$\frac{\partial w_i}{\partial \pi} = -\frac{p_i^* C}{\pi_i^2 \int_{p_i^*}^1 \left(p - p_i^*\right) dF\left(p\right)} < 0, \tag{28.4}$$

$$\frac{\partial \overline{w}_i}{\partial \pi} = -\frac{\left(1 - p_i^*\right) C}{\pi_i^2 \int_{p_i^*}^1 \left(p - \pi_i^*\right) dF\left(p\right)} < 0,$$

$$\frac{\partial w_i}{\partial p_i^*} = \frac{C}{\pi \left(\int_{p_i^*}^1 \left(p - p_i^*\right) dF\left(p\right)\right)^2}$$

$$= \frac{\left(\int_{p_i^*} 1 \left(p - p_i^*\right) dF\left(p\right) + p_i^* \int_{p_i^*}^1 dF\left(p\right)\right) \int_{p_i^*}^1 p dF\left(p\right) C}{\pi \left(\int_{p_i^*}^1 \left(p - p_i^*\right) dF\left(p\right)\right)^2}$$

$$> 0,$$

$$\frac{\partial \overline{w}_i}{\partial p_i^*} > 0.$$

A higher threshold to sell the new loan specification, p_i^*, increases the base salary w_i and thus reduces the integrand in equation (28.2). Furthermore, the lower boundary of the integration is increased as we see from equation (28.1), reducing this expression even further. Therefore the value of $\overline{w}_i$ has to increase in equation (28.2)as the right-hand side is constant, giving the result in the last relationship.

We thus find that the base salary of more able managers, those with a higher probability p of successfully developing a new loan specification, is lower, but they will receive their bonus more likely as the product is more likely to be developed and hence their overall remuneration will be higher. If misselling is less likely to occur, as reflected in a higher threshold for selling the newly developed loan p_i^*, the remuneration is higher. As from solving equation (28.1) for $\frac{\overline{w}_i}{w_i}$ we can obtain

$$\frac{\partial \frac{\overline{w}_i}{w_i}}{\partial \pi_i^*} = -\frac{1}{\left(p_i^*\right)^2} < 0, \tag{28.5}$$

we see that the importance of the bonus declines as the threshold becomes higher. This is because the base salary is more often paid and hence the bonus can be reduced without affecting the total remuneration of the manager.

Loan pricing If the loan is sold, the bank makes profits, provided it is not missold and if it is missold the fine F exceeds its equity and the bank will fail, leaving the bank with no return. If we assume for simplicity that the loan provided will be repaid with certainty, the profits of the bank consist of the profits if the loan is not missold and the loss of equity if the loan is missold. Hence we have

$$\hat{\Pi}_B^i = p\left(\left(1 + r_L^i\right) L - \left(1 + r_D^i\right) L\right) - (1 - p) E_i. \tag{28.6}$$

These profits are only realized if the loan is actually developed, which happens with probability π, and sold to clients, wich occurs with probability $p > p_i^*$. Furthermore, the salary and bonus of the manager need to be considered. Hence the expected profits of the bank is given by

$$\Pi_B^i = \pi \int_{p_i^*}^1 \hat{\Pi}_B^i dF\left(p\right) - \left(w_i + C\right). \tag{28.7}$$

Here $w_i + C$ are the manager remuneration. The total remuneration is given by the first two terms in equation (28.2) and by using the equality of this relationship, the second term equals C, giving the above expression for the total remuneration of the manager.

Banks can now determine the optimal threshold for selling the newly developed loan by maximzing their profits, giving us the first-order condition

$$\frac{\partial \Pi_B^i}{\partial p_i^*} = -\pi \hat{\Pi}_B^i f\left(p_i^*\right) - \frac{\partial w_i}{\partial p_i^*} = 0. \tag{28.8}$$

Note that as $\frac{\partial w_i}{\partial p_i^*} >$ from equation (28.4), we see that choosing the optimal threshold as determined to be optimal for the manager in equation (28.1), p_i^* the bank makes a loss as the first order condition implies need $\hat{\Pi}_B^i < 0$. As banks would not operate if they are making losses, the optimal threshold for a bank will differ from that of the manager.

We can now easily obtain that

$$\frac{\partial^2 \Pi_B^2}{\partial \pi_i^* \partial E_i} = -\pi\left(1 - p_i^*\right) f\left(p_i^*\right) < 0. \tag{28.9}$$

With $\frac{\partial^2 \Pi_B}{\partial \left(p_i^*\right)^2} < 0$ as the second order condition, we have using the implicit function theorem that

$$\frac{\partial p_i^*}{\partial E_i} = -\frac{\frac{\partial^2 \Pi_B^i}{\partial p_i^* \partial E_i}}{\frac{\partial^2 \Pi_B}{\partial \left(p_i^*\right)^2}} > 0, \tag{28.10}$$

such that the threshold optimal for managers p_i^* could be implemented by capital requirements.

If the bank extracts any surplus from their borrowers, they will repay the utility they gained, if given the loan. Hence we have the loan rate determined from

$$\begin{aligned} \left(1 + r_L^i\right) L &= E\left[p\overline{U} + (1 - p)\underline{U} | p \geq p_i^*\right] \\ &= \frac{1}{1 - F\left(p_i^*\right)} \int_{p_i^*}^1 \left(p\overline{U} + (1 - p)\underline{U}\right) dF\left(p\right). \end{aligned} \tag{28.11}$$

We can easily verify that the loan rate increases the higher the threshold for selling a newly developed loan, p_i^*, as $\frac{\partial(1+r_L^i)}{\partial p_i^*} > 0$. A higher threshold p_i^* reduces the probability of borrowers receiving the low utility, given them a higher expected utility, which can then be extracted by the bank through a higher loan rate.

Banks and shadow banks Let us now assume that there are two type of financial institutions offering loans. One type is a conventional banks (B) which benefits from free deposit insurance such that depositors face no risk of making a loss. The other type of financial institution is a shadow bank (S) that is does not benefit from deposit insurance and hence if the shadow bank fails, its depositors will not be fully repaid.

The expected return to depositors from using banks and shadow banks would be identical in equilibrium. The return to depositors using the bank will be safe, but the deposits at the shadow bank may be partially lost. If the newly developed loan specification is sold to the borrower by the shadow bank, thus $p > p_S^*$, the deposit rate at shadow banks can be determined by

$$\begin{aligned}\left(1+r_D^B\right) D = {} & \frac{1}{F\left(p_S^*\right)} \int_{p_S^*}^{1} \left(p\left(1+r_D^S\right) D + (1-p)\right. \\ & \left.\times\left(\left(1+r_L^S\right) D + E_S - F\right)\right) dF(p)\end{aligned} \tag{28.12}$$

Here the first term on the right-hand side denotes that if no misselling occurs, the depositors are repaid. The second term denotes the resources available if the loan is missold. It consists of the loan repayment, the equity, reduced by the losses from the misselling. If the newly developed loan specification is not sold, the deposits are instantly returned, hence the expression $\frac{1}{1-F(p_S^*)}$ is accounts for this possibility.

Due to our assumption that the fine imposed for misselling loans exceeds the equity of the bank, $E_S \leq F$, we see that for the expression in the integral we have $\left(1+r_L^S\right) D + E_S - F < \left(1+r_D^S\right) D$ and hence the expression is smaller than $1+r_D^S$ overall, implying $1+r_D^S > 1+r_D^B$ and the deposit rate of shadow banks are higher than of conventional banks to compensate for the risk due to the absence of deposit insurance.

Assume now that the threshold for selling newly developed loans to borrowers of banks and shadow banks are identical, $p_S^* = p_B^*$, then from equation (28.11) we see that $r_L^S = r_L^B$ and hence due to $r_D^S > r_D^B$ that $\hat{\Pi}_B^B > \Pi_B^S$, using equation (28.6) for identical equity $E_S = E_B$. But then the first order condition in equation (28.8) cannot be fulfilled for both, banks and shadow banks. Suppose the first order condition in equation (28.8) holds for shadows banks, then for banks we have $\frac{\partial w_B}{\partial p_B^*} > -\pi\hat{\Pi}_B^B f\left(p_B^*\right)$ as $\hat{\Pi}_B^B > \hat{\Pi}_B^S$. This implies we need to reduce the threshold of banks, p_B^*, such that we find $p_B^* < p_S^*$.

We thus have established that banks are more willing to missell loans than shadow banks. This arises from the fact that if shadow banks are misselling they expose depositors to risk, which will be reflected in a higher deposit rate and hence lower profits for the shadow bank. It is therefore that shadow banks are more cautious about misseling loans to protect their own profits from higher deposit rates.

Social optimum The welfare from shadow banks consist of the benefits of the loan to borrowers, less the fines F and the costs of developing the loan, C, hence we have

$$\Pi_W^S = \pi \int_{p_S^*}^1 \left(\left(p\overline{U} + (1-p)\underline{U}\right) - (1-p)F\right) dF(p) - C. \qquad (28.13)$$

The interest paid by borrowers to the shadow bank is exactly offset by the interest received by the shadow bank and can therefore be neglected as it is only redistribution of funds. We do not account for the fine being received by the regulator by claiming that this will be used to cover the costs of any investigations into the misselling.

The optimal threshold for selling the newly developed loan is then given from the first order condition as

$$\frac{\partial \Pi_W^S}{\partial p_S^*} = -\left(\left(1 - p_S^*\overline{U} + \left(1-p_S^*\right)\underline{U}\right) - \left(1-p_S^*\right)L\right) f\left(p_S^*\right) = 0,$$

which solves for

$$p_S^{**} = \frac{F - \underline{U}}{\overline{U} - \underline{U} + F}. \qquad (28.14)$$

This social optimum can be implemented by an appropriate level of equity E_S^{**} in the same way that the optimal threshold of managers could be implemented for banks above.

For banks the costs of the deposit insurance need to be considered. The amount paid out by the deposit insurance is

$$T = \left(1 + r_D^B\right)D - \left(\left(1 + r_L^B\right)D + E_B - F\right), \qquad (28.15)$$

consisting of the amount due to depositors, reduced by the resources available to banks from the loan repayment, its equity and the additional losses due to the fine imposed on the bank. While these amounts benefit depositors, we assume a dead weight loss of λT to account for the financing of the deposit insurance and associated administrative costs. Hence the social welfare when using banks are given by

$$\Pi_W^B = \pi \int_{p_B^*}^1 \left(\left(p\overline{U} + (1-p)\underline{U}\right) - (1-p)(F + \lambda T)\right) dF(p) - C, \qquad (28.16)$$

and the first order condition for its maximum is given by

$$\frac{\partial \Pi_W^B}{\partial p_B^*} = -\left(\left(p_B^* \overline{U} + (1-p)\,\underline{U}\right) - \left(1-p_B^*\right)(F + \lambda T)\right) f\left(p_B^*\right) + \pi \int_{p_B^*}^1 (1-p)\,\lambda \frac{\partial\left(1+r_L^B\right)}{\partial p_B^*} D\,dF(p) = 0.$$

This condition can be solved for

$$p_B^* \overline{U} + \left(1-p_B^*\right)\underline{U} - \left(1-p_B^*\right)(F+\lambda T) = \frac{\lambda}{f\left(p_B^*\right)} \frac{\partial\left(1+r_B\right)}{\partial p_B^*} D \int_{p_B^*}^1 (1-p)\,dF(p) > 0, \tag{28.17}$$

where the inequality arises from $\frac{\partial\left(1+r_L^B\right)}{\partial p_B^*} > 0$. The solution of this condition for p_B^{**} can then again be implemented using capital requirements E_B^{**}.

Suppose that $p_B^{**} < p_S^{**}$, then

$$\begin{aligned} p_B^{**}\overline{U} + \left(1-p_B^{**}\right)\underline{U} - \left(1-p_B^{**}\right) &< p_S^{**}\overline{U} + \left(1-p_S^{**}\right)\underline{U} \\ &\quad - \left(1-p_S^{**}\right)F \\ &= 0, \end{aligned} \tag{28.18}$$

with the final equality arising from equation (28.14). To see the validity of equation (28.18) rewrite it as

$$\left(p_B^{**} - p_S^{**}\right)\left(\overline{U} - \underline{U}\right) < \left(p_S^{**} - p_B^{**}\right)F + \left(1-p_B^{**}\right)\lambda T.$$

With $p_B^{**} < p_S^{**}$ and $\overline{U} > \underline{U}$ the left-hand side is negative while the right-hand side is positive.

However, equation (28.17) requires the left-hand expression to be positive, a contradiction to our requirement that it is negative. Hence we need the threshold to sell the newly developed loan by banks in the social optimum to be higher than that of shadow banks, $p_B^{**} \geq p_S^{**}$; this result directly implies from equation (28.10) that banks will have higher capital requirements to implement the social optimum than shadow banks, $E_B^{**} > E_S^{**}$.

We have now established that it is socially optimal for banks to be less likely to missell than shadow banks and thus hold more equity to implement this solution. This in contrast to the optimal solution for banks, which were more likely than shadow banks to misell loans. Therefore, banks will require more restrictive capital requirements than shadow banks, if compared to the capital they would hold voluntarily to obtain their optimal threshold.

Allocation of managers Finally, let us assume that there are managers with a high ability to generate new loan specifications, π_H and managers with a low ability to generate new loan specifications $\pi_L < \pi_H$. However, banks cannot distinguish between these two types of managers, while managers are aware of their own type.

For the less able of the two managers, π_L, there will be no competition, hence remuneration will be set at the minimum as given in equation (28.3). This is because we have $p_B^{**} > p_S^{**}$, $\frac{\partial w_i}{\partial p_i^*} > 0$ and $\frac{\partial \overline{w}_i}{\partial p_i^*} > 0$, and therefore the bank will in any case offer the better conditions for this manager.

If we denote by w_i^H and $\overline{w}_i^H$ the base salary and bonus, respectively, of the high-ability manager and w_i^L and $\overline{w}_i^L$ the equivalent for the low-ability manager, we now need to ensure that the low-ability manager chooses w_i^L and $\overline{w}_i^L$, while the high-ability manager chooses w_i^H and $\overline{w}_i^H$. from equation (28.2) we therefore need

$$\begin{aligned} w_S^H + \pi_H \int_{p_S^{**}}^1 \left(p\overline{w}_S^H - (1-p)\, w_S^H\right) dF(p) - C \\ \geq w_B^L + \pi_H \int_{p_B^{**}}^1 \left(p\overline{w}_B^L - (1-p)\, w_B^L\right) dF(p) - C. \end{aligned} \tag{28.19}$$

We can neglect the contract for the low-ability manager at the shadow bank because this is dominated by the contract at the bank as outlined above. Hence we need the high-ability manager to select the contract at the shadow bank rather than the low-ability contract at the bank, which this condition implies.

At the same time, the the low-ability manager needs to prefer the bank over the shadow bank. This then implies that we require

$$\begin{aligned} \max\left\{w_S^H, w_S^H + \pi_L \int_{p_S^{**}}^1 \left(p\overline{w}_S^H - (1-p)\, w_S^H\right) dF(p) - C\right\} \\ \leq w_B^L + \pi_L \int_{p_B^{**}}^1 \left(p\overline{w}_B^L - (1-p)\, w_B^L\right) dF(p) - C. \end{aligned} \tag{28.20}$$

Here the first term accounts for the possibility that no new loan specification is developed by the manager. By virtue of equation (28.2), the final term in equation (28.20) becomes zero. If we now assume that $\pi_L \int_{p_S^{**}}^1 \left(p\overline{w}_S^H - (1-p)\, w_S^H\right) dF(p) - C \geq 0$, such that the low-ability manager would develop new loans even if in the working for the shadow bank, the condition in equation (28.20) becomes

$$w_B^L - w_S^H \geq \pi_L \int_{p_S^{**}}^1 \left(p\overline{w}_S^H + (1-p)\, w_S^H\right) dF(p) - C. \tag{28.21}$$

From equation (28.19) we then get that

$$w_B^L - w_S^L \leq \pi_H \left(\int_{p_S^{**}}^{1} \left(p_B S^H - (1-p)\, w_S^H \right) dF\,(p) \right. \tag{28.22}$$
$$\left. - \int_{p_B^{**}}^{1} \left(p\overline{w}_B^L - (1-p)\, w_B^L \right) dF\,(p) \right)$$
$$= \pi_H \int_{p_S^{**}}^{1} \left(p\overline{w}_S^H - (1-p)\, w_S^H \right) dF\,(p) - \frac{\pi_H}{\pi_L} C,$$

where we used equation (28.3) to obtain the final expression. Now using equation (28.1), we can write

$$\int_{p_S^{**}}^{1} \left(p\overline{w}_S^H - (1-p)\, w_S^H \right) dF\,(p) = w_S^H \int_{p_S^{**}}^{1} \frac{p - p_S^{**}}{p_S^{**}} dF((p)\,, \tag{28.23}$$

which is increasing in the base salary of the high-ability manager working for the shadow bank, w_S^H, as the integrand is positive in the relevant range.

This in turn implies that equation (28.21) becomes less binding as long as w_S^H increases; then the left-hand side decreases and the right-hand side increases. The opposite is true of equation (28.22). Hence as competition between banks and shadow banks increases, the base salary w_S^H, and indirectly thereby $\overline{w}_S^H$ through equation (28.1), equation (28.22) will eventually become binding. Thus we have

$$w_B^L - w_S^H = \frac{\pi_H}{\pi_L} \left(\pi_L \int_{p_S^{**}}^{1} \left(p_B \right) S^H - (1-p)\, w_S^H \right) dF\,(p) - C \right). \tag{28.24}$$

With our assumption that low-ability managers develop new loan specifications, equation (28.21) is fulfilled when inserting equation (28.24) due to our assumption that $\pi_H > \pi_L$.

We have thus established that more able managers will work for shadow banks and less able managers will work for banks. The higher propensity of banks to missell loans will make the efforts of highly-able managers less beneficial as more frequently their newly developed loan is missold and he therefore loses his base salary, while in shadow banks his new developed loans will be sold less frequent and therefore he will only lose out on his bonus. This makes shadow banks more attractive for highly able managers and banks more attractive to less able banks. The latter will lose their base salary less frequently with banks as they less often successfully develop new loan specifications.

Summary We have established that is socially desirable for banks to less frequently missell loans than for shadow banks, but banks themselves find it optimal to missell loans more often than shadow banks. This disparity between the social optimum and the optimal decision by banks can be addressed through a higher capital requirement by banks such that the social optimum is implemented; any such regulation would be much more stringent for banks than shadow banks. What remains, however, is the

willingness of managers to missell loans. In order to induce them to develop such new loan specification, they may well be willing to missell loans to easily, resulting in a conflict of interest between the bank and the managers selling these loans.

We have established that banks are less likely to missell loans than shadow banks if they are regulated, but due to lower deposits rates banks would prefer to missell more. Despite the lower risks from misselling, more able managers prefer to join shadow banks.

Readings Inderst & Ottaviani (2009), Song & Thakor (2022)

28.2 Corruption in the granting of loan

The decision to grant a loan to a borrower is made by designated employees of banks. They will make such decisions on the basis of an analysis of the creditworthiness of the borrower as well as having regard to the policy for granting policy the bank applies. The assessment of the creditworthiness of a borrower includes many subjective factors, either by giving different aspects a weighting or by the use of soft information which is not easily verifiable and might include the impression the employee has of the management of the company, or just his own experience from similar situations. Such components in the decision-making process can easily be interpreted differently by different employees such that the borrower might obtain a loan if being assessed by one employee, but the loan would not be granted by another employee. However, an employee could also easily distort his assessment of the borrower and grant a loan that if he would use his honest judgement would not be granted. With companies seeking loans to increase their profits or even requiring a loan to avert bankruptcy, the company might be tempted to bribe the employee making the loan decision. We will here assess under which conditions bank employees would grant loans based on the bribe they receive from a borrower. of course, rather than accepting a bribe to grant a loan, the employees might also accept a bribe to provide the borrower with better loan conditions than they would receive otherwise.

Let us assume that not all bank employees are susceptible to such bribes, but only a fraction $1 - \lambda$ will consider accepting a bribe B from a borrower to grant a loan that would otherwise be refused, while a fraction λ would never consider this. We furthermore assume that the original assessment of the employee whether to grant a loan or not is correct and employees would not withhold a loan they would normally grant only to extract a bribe from the borrower; hence we only consider the case where a loan should be refused, but the employee might be bribed to grant the loan. If a loan has been given and the loan is repaid, the employee receives a remuneration of w. He receives the same payment if no loan is granted as this is a signal that the loan was rightly refused as it would not be repaid.

We consider two types of companies. The first type, of which there is a proportion of p, are safe companies as regardless of the outcome of their investment, the loan can be repaid. With r_D denoting the interest on the deposits that finance the loan, we assume that with probability π the safe company manages a return of $\overline{R} > r_D$ and with probability $1-\pi$ it yields $\overline{R} > \underline{R} > r_D$. To ensure the safety of their company, it faces additional costs C from enhanced management. Loans to this companies always cover the funding costs, hence we observe that for the company profits we obtain

$$\Pi_C = \pi \left(1 + \overline{R}\right) L + (1-\pi)\left(1 + \underline{R}\right) L - C \geq (1 + r_D) L \tag{28.25}$$

The second type of company, with a proportion $1 - p$ in the economy, will yield $\underline{R} > r_D$ with probability π and zero otherwise. Therefore this company is risky as the loan can only be repaid if the investment succeeds. Furthermore, banks cannot distinguish between a successful risky company and an unsuccessful safe company, hence conditioning employee remuneration on the success or failure of the investment the companies makes is not possible, as only the employee will know the true type. This risky company is not creditworthy as we assume that it does not cover its funding costs, given that its profits are given by

$$\hat{\Pi}_C = \pi\left(1 + \underline{R}\right) < 1 + r_D. \tag{28.26}$$

We note that risky companies do not face the additional management costs C.

As safe companies are creditworthy and we assume that employees assessing them cannot reduce the status of a safe company, they would always obtain a loan and hence pay no bribe to the employee; as the loan is repaid, the employee would receive his remuneration of w. Only the risky company would bribe the employee who in granting this loan would obtain the bribe B and the bonus w in case the investment is successful and the loan repaid, thus he would obtain πB. Faced with the decision to grant a loan to a safe or a risky company, he would accept the bribe and grant the loan to the risky company, assuming he is susceptible to bribes, if it is more profitable for him to do so, $B + \pi w \geq w$. This can be solved for

$$B \geq \underline{B} = (1-\pi) w. \tag{28.27}$$

Thus a bribe would only be accepted if it is sufficiently large; such a bribe will only be paid by the company if it is profitable to do so. The risky company only repays the loan if its investment is successful, while the bribe comes out of general funds. Hence if the risky company approaches a susceptible employee, it will obtain the loan L such that the profits are

$$\hat{\Pi}_C = (1-\lambda)\left(\pi\left(\left(1+\underline{R}\right)L - (1+r_L)L\right) - B\right). \tag{28.28}$$

As not being granted the loan returns zero profits, the company will pay a bribe if $\hat{\Pi}_C \geq 0$. Therefore we require that

$$B \leq \overline{B} = \pi\left(\left(1+\underline{R}\right)L - (1+r_L)L\right). \tag{28.29}$$

For bribery to be feasible, it needs to be offered by companies and being accepted by the employee, therefore we require that both conditions in equations (28.27) and

(28.29) are fulfilled. This requires that a bribe can only be determined meeting both conditions if $\underline{S} \leq \overline{S}$, which solves for

$$w \leq w^{*} = \frac{\pi}{1-\pi}\left(\left(1+\underline{R}\right)L - \left(1+r_L\right)L\right). \tag{28.30}$$

If the wages of employees are sufficiently low, then risky companies and the employee could agree on a bribe that would induce the loan to be granted.

As risky companies would not be profitable, banks only grant loans to safe companies in the absence of bribery. As the repayment of the loan to a risky company is certain, the profits of the banks is given by

$$\Pi_B = p\left(\left(1+r_L\right)L - \left(1+r_D\right)L\right) - w, \tag{28.31}$$

noting that the remuneration to the employee is always payable in this case, given that the loan is repaid. A bank would be willing to lend if this generates profits, $\Pi_B \geq 0$, from which we can obtain the requirement that

$$w \leq w^{**} = p\left(\left(1+r_L\right) - \left(1+r_D\right)\right)L \tag{28.32}$$

Thus if the remuneration of employees is exceeds w^{**}, banks will not provide any loans as it is not profitable to do so.

If bribes are taken and accepted, the loan will only be repaid if it was either given to a safe company, p, or a risky company, $1-p$, that is successful, π, and been approved by a dishonest employee, $1-\lambda$. Loans are only refused if the company is risky, $1-p$ and an honest employee decided on the application, λ. The remuneration is always payable unless a risky loan, $1-p$, fails to repay, $1-\pi$, which had to be granted by a susceptible employee, $1-\lambda$. Hence the profits of the bank in this case are given by

$$\begin{aligned}\hat{\Pi}_B = {} & \left(p + (1-p)(1-\lambda)\pi\right)\left(1+r_L\right)L \\ & - \left(1-(1-p)\lambda\right)\left(1+r_D\right)L \\ & - \left(1-(1-p)(1-\pi)(1-\lambda)\right)w.\end{aligned} \tag{28.33}$$

Banks lend as long as it is profitable,$\hat{\Pi}_B \geq 0$, which requires

$$\begin{aligned} w \leq w^{***} = {} & \frac{1+(1-p)(1-\lambda)\pi}{1-(1-p)(1-\pi)(1-\lambda)}\left(1+r_L\right)L \\ & - \frac{1-(1-p)\lambda}{1-(1-p)(1-\pi)(1-\lambda)}\left(1+r_D\right)L. \end{aligned} \tag{28.34}$$

The profits of the safe company are given by

$$\Pi_C = \pi\left(1+\overline{R}\right)L + (1-\pi)\left(1+\underline{R}\right)L - C - \left(1+r_L\right)L. \tag{28.35}$$

It will ask for loans as long as it is profitable, $\Pi_C \geq 0$, which requires

$$1 + r_L \leq \pi \left(1 + \overline{R}\right) + (1 - \pi) \left(1 + \underline{R}\right) L - \frac{C}{L}. \tag{28.36}$$

A higher loan rate would ensure that only risky companies are demanding loans, and thus the market would collapse as banks would not lend because they would be fully aware that those demanding loans would be risky companies.

A bank maximizing profits would obviously chose the lowest possible remuneration for employees and the highest possible interest rate on the loan. If the bank would to avoid bribes becoming viable it would set be remuneration such that the condition in equation (28.30) for bribes to be feasible is fulfilled with equality, as well as setting the loan rate at the level of equation (28.36). These relationships can be inserted into equation (28.36) to obtain the highest possible profits in the absence of bribes.

The bank would prefer corruption if accepting such practices is more profitable than not accepting bribes, $\hat{\Pi}_B \geq \Pi_B$, due to the bank being able to pay lower remuneration to employees. This condition requires, after inserting equations (28.30) and (28.36) into equation (28.31) and comparing it with equation (28.33)

$$w \leq w^{****} = \frac{p + (1 - p)(1 - \lambda)\pi}{1 - (1 - p)(1 - \pi)(1 - \lambda)} - \gamma, \tag{28.37}$$

where

$$\gamma = \frac{\left\{ \begin{array}{l} (1 - p)(1 - \lambda)(1 + r_D) L + \frac{\pi(\pi + p(1-\pi))}{1-\pi} \left(1 + \overline{R}\right) L \\ \quad + \frac{(1-\pi)(\pi + p(1-\pi)) + \pi}{1-\pi} \left(1 + \underline{R}\right) L - \frac{\pi + p(1-\pi)}{1-\pi} C \end{array} \right\}}{1 - (1 - p)(1 - \pi)(1 - \lambda)}.$$

Figure 28.1 summarizes the different possible scenarios we have established. We clearly see that if the remuneration of employees is too high and the loan rate is too low, the bank would not be willing to provide any loans as it would make a loss. Above the line of w^* the remuneration is sufficiently high to prevent corruption, but above w^{**} the bank would make a loss due to the high remuneration and would thus not be active in the market. Above w^{***} corruption would occur, but given the low loan rate, the bank would make a loss and hence also not operate in the market. To the right of the line $\Pi_C = 0$ the loan rate is so high, that the risk-free borrower would not demand a loan and hence no lending would occur, given that it is then known that any loan demand would be from the risky borrower.

It is between w^* and w^{**} that banks are profitable and the high remuneration employees receive make them unwilling to accept a bribe; thus only risk-free loans are provided. In this region remuneration is sufficiently high that the risk of losing it due to a loan failing cannot be compensated for by the bribe as its size would be too high to be profitable to companies. Below w^*, the remuneration of employees is sufficiently low that they would be accepting a bribe, but the loan rate is also sufficiently high to allow banks to remain profitable. It is even that in the area below w^{****} the bank would be more profitable if employees accept bribes than if they do not accept bribes. This is because the bank could pay employees a very low remuneration,

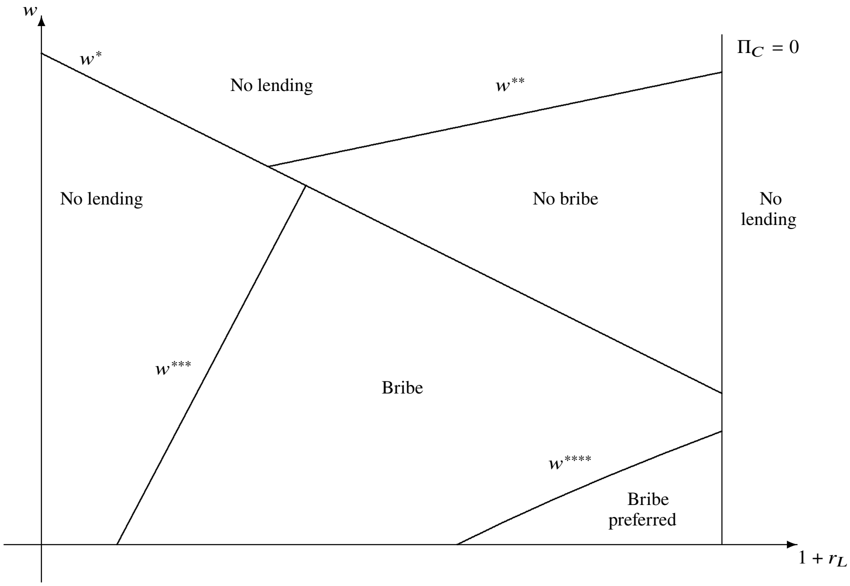


Fig. 28.1: Employees taking bribes to provide loans

increasing their profits, and employees supplement this low remuneration with the bribe they obtain from the company. The high loan rate will more than compensate banks for the risky loans that have been provided.

From inspecting the constant γ in equation (28.37) we see that if this term is large, we might find that $w^{****} < 0$ and hence not a viable solution, such that an equilibrium where banks would tolerate corruption does not occur. Such a situation emerges if the returns of investments, $\overline{R}$ and $\underline{R}$, are high, and in particular if there are only few risky companies, p, the success probability is high, π, or few employees are susceptible to bribes. Hence in such a favourable environment, corruption is less likely to occur as the bank ensures it is avoided, while in an environment that seems more susceptible to corruption with many susceptible employees, many risky companies, and lower probability of success, corruption is tolerated by banks.

We have thus seen that in cases where employees are not paid sufficiently well, they are willing to accept bribes to grant loans to companies that are not creditworthy. It is even that banks may be willing to tolerate such corruption as it allow them to pay low remuneration to their employees and for that are willing to accept losses from loans defaults. We have imposed no sanctions on employees accepting bribes and including such sanctions would make the their occurrence much less common, but would persists in cases where employees are receiving low remunerations. given the large amount of loans banks grant and the size of potential benefits to companies of obtaining such loans, it is likely that corruption will have a place where banks are not paying their employees a sufficiently high salary.

Reading Hwang, Jiang, & Wang (2007)

28.3 Deferred realisation of losses

The investments of companies may not succeed and therefore their loan could not be repaid due to a lack of funds by the company. Banks might be reluctant to recognise such losses as it might cause their own failing; we will here discuss the possibility of banks extending loans that a company cannot repay for another time period with the aim of generating more profits from other loans to cover the losses once they have to be recognised at a later point of time. There is no possibility of such loans recovering and being repaid at a later point by the company, but banks do this solely to avoid declaring themselves bankrupt.

Let us assume that a company can use a loan L to either invest for a single time period, a short-term investment, or for two time periods, a long-term investment. The bank cannot distinguish between companies that use either investment; they only know that a fraction λ of investments will be short-term and thus a fraction $1 - \lambda$ will be long-term investments. Either investment succeeds with probability π and as banks cannot distinguish between companies making either investment they will charge a common loan rate r_L and finance the loan fully by deposits on which interest r_D is payable. Given the fraction λ of short-term investments, the expected profits of the bank after one time period are given by

$$\Pi_B^1 = \lambda \left(\pi \left(1 + r_L\right) L - \left(1 + r_D\right) L\right). \tag{28.38}$$

We now assume that the investments are sufficiently risky for banks to make a loss, thus $\Pi_B^1 < 0$, and the bank would face bankruptcy. The bank making losses after one time period implies that the success rate of investments is below

$$\pi < \frac{1 + r_D}{1 + r_L}. \tag{28.39}$$

We can rewrite the profits of the bank after one time period in equation (28.38) as

$$\begin{aligned} \Pi_B^1 = {} & \lambda\pi \left(\left(1 + r_L\right) L - \left(1 + r_D\right) L\right) \\ & -\lambda \left(1 - \pi\right) \left(1 + r_D\right) L, \end{aligned} \tag{28.40}$$

where the first term denotes the profits of those companies that succeeded and the second term the losses from that failed.

The bank could now declare profits of $\lambda\pi \left(\left(1 + r_L\right) L - \left(1 + r_D\right) L\right)$ from those companies that succeeded and repaid the loan and extend the loan of those companies that failed. The bank would thus extend loan to the amount of $\lambda \left(1 - \pi\right) \left(1 + r_D\right) L$. This means that rather than declaring a loss and facing bankruptcy, the bank declares a profit and through not recognizing the losses of companies that have failed, can continue to operate.

In the second time period the profits of the bank are then given by

$$\Pi_B^2 = (1-\lambda)\left(\pi\,(1+r_L)^2 L - (1+r_D)^2 L\right) \\ -\lambda\,(1-\pi)\,(1+r_D)^2 L. \tag{28.41}$$

The first term denotes the profits from the loan to companies conducting long-term investments, where interests on the loan and the financing deposits are accumulated over two time periods. The second term represents the losses from the loan extended to the failing short-term loans. These loans are never repaid, but need to be financed by deposits by additional deposits.

The bank would want to avoid bankruptcy in period 1, and thus has to forego future profits from those companies that make long-term investments; this is possible if $\Pi_B^2 > 0$, which implies that we require

$$\frac{(1+r_D)^2}{\lambda\,(1+r_D)^2 + (1-\lambda)\,(1+r_L)^2} < \pi < \frac{1+r_D}{1+r_L} \tag{28.42}$$

with the final inequality arising from our assumption that the bank is not profitable in the first time period as shown by equation (28.39). A viable solution in which the probability of success of the investment, π, fulfills both conditions only emerges if

$$\lambda < \frac{1+r_L}{(1+r_L)+(1+r_D)}. \tag{28.43}$$

We therefore see that we need a sufficiently large fraction of companies seeking long-term investments to generate sufficient the profits in the second time period, compensating for the losses from failing short-term investments.

The effect of a bank employing this strategy is that profits in the second time period are reduced due to the losses from extending non-performing loans. Furthermore, the aggregate profits across both time periods are smaller due to the additional funding costs for the non-performing loans. As the bank in this way can capture future profits, albeit reduced, they will prefer this strategy to declaring themselves bankrupt after the first time period.

A regulator may also prefer banks to extend non-performing loans as this might avoid costs from requiring deposit insurance to refund depositors or a banking crisis arising from the failure of this bank, for example the emergence of bank runs or contagion effects. On the other hand, resources are bound within the bank for non-performing loans that otherwise could be used to provide more productive loans that support economic growth.

Not recognising the default on a loan by extending it for another time period to avoid the bankruptcy of the bank can be interpreted as misconduct as the bank should have declared itself bankrupt. However, it has to be recognised that such a situation will only be beneficial for the bank to ensure its survival if the success rate of investments is in the narrow band as indicated by equation (28.42), and thus is applicability would be limited to the case of risky, but not too risky loans.

Reading Niinimaki (2007)

28.4 Money laundering

Accounts held at banks are also used to transfer the proceeds of crimes or to evade financial restrictions individuals or companies might have imposed on them. It is for this reason that banks are required to report on any suspicious transaction that might aid any such criminal activities. Once a transaction has been reported as being suspicious, it will usually be stopped and then investigated by the relevant government authorities. If banks fail to report suspicious activities and a regulator or investigating government authority detect this failure, banks can face significant fines for their non-compliance with reporting requirements. Such fines are levied to incentivise banks to report such any suspicious transactions affecting their accounts and establish systems to monitor all transactions.

We will here analyse the optimal level of resources banks should devote to the monitoring and reporting of suspicious transaction, as well as the government investigating these. Based on this analysis we can then determine the size of the fine that allows any regulator to ensure the optimal level of compliance is achieved.

consider that the ex-ante probability of a transaction being related to money laundering is π. The bank can obtain a signal on the nature of the transaction, whether it is related to money laundering or not, which is correct with probability p; the acquisition of such a signal costs the bank monitoring costs of C_M. Using Bayesian rules we then can determine the probability of the transaction being the result of money laundering given a negative signal, π_0, and given a positive signal, π_1. We obtain

$$\begin{aligned}\pi_0 &= = \frac{\pi(1-p)}{\pi(1-p)+(1-\pi)p}, \\ \pi_1 &= = \frac{\pi p}{(1-\pi)(1-p)+\pi p},\end{aligned} \tag{28.44}$$

where it is easy to verify that $\pi_1 \geq \pi \geq \pi_0$.

If the bank reports its suspicion, it incurs additional reporting costs of C_R. Such costs may arising from the bank itself being subject to further investigation, losing customer confidence in case of reporting transactions that are found to be unrelated to money laundering, or a loss in confidence due to their (alleged) involvement in money laundering. The bank can decide to report only on positive signals, or on both positive and negative signals, the latter implying that they report any transaction. In this case the report to government authorities is not informative and it will assess that the likelihood of the transaction being laundering is π. If the bank only reports positive signals, then a report implies that a transaction is related to money laundering with probability π_1, and the absence of such a report implies a negative signal, such that the probability of the transaction being involve din money laundering is π_0.

Money laundering imposes social costs of C_S which can be alleviated if it is detected. We will now firstly determine the optimal level of monitoring and then use this result to determine fines for not reporting transactions that are found to be related to money laundering that allows us to implement this optimum.

Social optimum Let us assume that, depending on the report by the bank, a regulator puts effort in to detect money laundering itself based on the reported transactions. It does detect money laundering with probability γ_i and this detection costs C_G, which is increasing in the probability of detection, γ_i.

If the bank monitors for money laundering, the social welfare is given by

$$\begin{aligned} \Pi_W^1 &= \pi\gamma_1 C_S - C_G^1 - C_M - C_R - \pi C_S, \qquad (28.45) \\ \hat{\Pi}_W^1 &= ((1-\pi)(1-p) + \pi p)\left(\pi_1\hat{\gamma}_1 C_S - \hat{C}_G^1 - C_R\right) \\ &\quad + (1-(1-\pi)(1-p) - \pi p)\left(\pi_0\hat{\gamma}_0 C_S - \hat{C}_G^0\right) \\ &\quad - C_M - \pi C_S. \end{aligned}$$

for reporting all signals and only positive signals, respectively. C_G^1 refers to the costs of the regulator in identifying transactions related to money laundering with probability γ_1 if the bank reports all transactions, and $\hat{C}_G^1$ ($\hat{C}_G^0$) the costs leading to detection of $\hat{\gamma}_1$ ($\hat{\gamma}_0$) such transactions from reporting a positive (negative or no) signal. In the case that banks report all transactions as being suspicious, the first term denotes that money laundering occurs with probability π as the report is not informative, of these transactions that are involved in money laundering the regulator is able to identify a fraction γ, saving social costs C_S, but incurring monitoring, reporting costs, and the social costs of all transactions involved in money laundering. The second line for the case that the bank only reports positive signals achieves represents the same, although recognizing that only positive signals are reported in the first term and no report is obtained in the second term, implying a negative signal.

Without banks monitoring transactions, and hence no additional information to the regulator, we have social welfare given as

$$\Pi_W^0 = \pi\gamma_0 C_S - C_G^0 - \pi C_S, \qquad (28.46)$$

with government costs C_G^0 to identify γ_0 transactions correctly. There are no monitoring and reporting costs as no such costs are incurred without conducting monitoring. The optimal fraction of transactions identified as relating to money laundering for the regulator in this case, is given from the first order condition

$$\frac{\partial \Pi_W^0}{\partial \gamma_0} = \pi C_S - \frac{\partial C_G^0}{\partial \gamma_0} = 0. \qquad (28.47)$$

This solution is identical to the case that when monitoring happens, but all transactions are reported, which we see when differentiating the first line in equation (28.45). This emerges as the reports the bank submits convey no information. Hence it is easy to see that $\Pi_W^0 > \Pi_W^1$, due to the absence of monitoring and reporting costs.

From equation (28.45) the first order condition for a maximum of social welfare is given by

$$\frac{\partial \Pi_W^i}{\partial \hat{\gamma}_i} + \pi_i C_S - \frac{\partial \hat{C}_G^i}{\partial \hat{\gamma}_i} = 0. \tag{28.48}$$

If we assume that the marginal costs to achieve a higher detection rate are increasing, $\frac{\partial^2 C_G^i}{\partial \gamma^2} > 0$, then for positive signals only being reported, the regulator would use generate higher social welfare for their investigation. We see that the higher value of π_1 will increase the first term and hence the marginal costs have to be higher, thus implying higher costs but also higher detection rates, which will increase social welfare. Conversely, for those transactions not reported, repeating the same approach with the lower value of π_0, will reduce marginal costs and hence reduce the detection rate for these transactions. The regulator will therefore put more resources into transactions that have received a positive signal and are therefore more likely to be involved in money laundering and less resources into those transactions that have received a negative signal and are therefore no reports are being made; these transactions are less likely to be involved in money laundering. It is thus that more transactions involved in money laundering are detected. While it will be dependent on the excat paramter constellations, let us assume that reporting only positive signals increases welfare, thus $\hat{\Pi}_W^1 > \Pi_W^0$.

We will thus now seek to determine fines for banks not reporting transactions that regulators identify as related to money laundering such that banks will monitor transactions and only report positive signals.

Optimal fines If banks report only positive signals, their costs are

$$\begin{aligned} \hat{C}_1 = {} & ((1-\pi)(1-p) + \pi p) C_R \\ & + (1 - ((1-\pi)(1-p) + \pi p)) \hat{\gamma}_0 \pi_0 F + C_M, \end{aligned} \tag{28.49}$$

where the first term denotes the probability of a positive signal with the associated reporting costs, the second term denotes the unreported signals that were found relating to money laundering on investigation by the regulator and for which the bank is fined F as they were not reported. The final term are the monitoring costs.

The costs to banks of reporting all transactions are given by

$$C_1 = C_R + C_M, \tag{28.50}$$

as all are transactions will be reported reporting and monitoring costs, but as they have reported all transactions, they cannot be fined. If banks do not monitor, they incur no monitoring or reporting costs, but will face fines for all those transactions that are identified by the regulator to be related to money laundering, hence we have

$$C_0 = \pi \gamma_0 F. \tag{28.51}$$

To achieve that banks monitor and report only positive signals, the costs to banks need to be smaller than when reporting only positive signals than reporting all signals, $\hat{C}_1 < C_1$, and also lower than when not monitoring transactions, $\hat{C}_1 < C_0$. These two conditions solve for

$$\frac{((1-\pi)(1-p)+\pi p)\,C_R + C_M}{\pi\gamma_0 - (1-((1-\pi)(1-p)=\pi p))\,\hat{\gamma}_1\pi_0} < F < \frac{C_R}{\pi_0\hat{\gamma}_1}. \tag{28.52}$$

Hence, in order to ensure that banks report only positive signals and thus reduce the costs, fines need to be sufficiently large to ensure monitoring happens, but not too high to deter reporting only positive signals. This latter approach ensures that regulators are not overwhelmed with reports, from which they then have to identify transactions related to money laundering.

If the reporting costs are too small, no viable solution exists as the incentives to report each transaction are dominating. Equation (28.52) is only viable if

$$C_R > \frac{\pi_0\hat{\gamma}_1}{\pi_0\hat{\gamma}_1 - \pi\gamma_0} C_M \tag{28.53}$$

and hence the reporting costs are sufficiently higher than the monitoring costs. It is therefore not optimal to make reporting suspected transactions related to money laundering a low-cost option for banks.

Summary We have seen that banks can be enticed to monitor transactions on their accounts for signs of money laundering and report any suspicions they have through appropriate fines that are levied on any transactions that have not been reported to a regulator, but are found by this regulator to involve money laundering. The fine must be sufficiently high to enure that monitoring is conducted, but must not be so high that banks become overly cautious and report any transaction as having the potential to be involve din money laundering. This approach allows regulators to focus their resources on investigating those transactions that have been identified by banks, while paying much less attention to those that have not been rased as a concern. With that strategy, money laundering can be detected more efficiently and social welfare will increase. Such an approach is only feasible, however, if the costs of banks to report suspicious transactions are sufficiently high; this will reduce the instances where banks become overcautious and in order to avoid a possible fine report each transaction as suspicious. Thus regulators should make it sufficiently costly to report suspicious transactions, but of course not prohibitively expensive relative to the fine they impose as to deter any reporting. For example, regulators might raise costs for banks of reporting suspicious transactions by increasing the costs of them conducting business and ensure less money laundering is attempted in the future, this might be achieved by putting restrictions on their business or enhanced monitoring and auditing of processes.

Reading Takáts (2009)

Conclusions

We have seen that there are several incentives for banks or their employees to engage in practices that are not in the interest of their customers, such as misselling of loans, might not be in the interest of the bank but only the employees, for example the acceptance of bribes to grant loans; other practices might be in the interest of the bank, but not always in the public interest. It would be in the interest of the bank to defer realising any losses from loans if this allows the bank to survive and generate future profits, and while this might also be in the public interest to avoid a contagion of any bank failures, the lack of transparency on the position of the bank will generally not be looked upon favourably. Depositors might be inferring that the bank is stable and retain their deposits, while it would be optimal for them to withdraw these if they knew the true position of the bank. Regulators might also want to be informed about the true position of a bank in order to effectively supervise the bank and initiate measures to mitigate any consequences of a future failure or the potential for contagion of the failure. The incentives to effectively monitor money laundering by banks have been established to consist of fines for not identifying such transactions, where these fines have to be high enough to ensure monitoring by banks takes place, but not so high that banks become overly cautious and report every transaction to regulators and overwhelm their ability to properly investigate possible cases.

The incentives of banks and employees in these models are purely driven by monetary considerations and moral concerns employees might have not been considered. We would expect that the willingness of employees to engage in activities that are seen as malpractice or even misconduct to be low, even if monetary gains can be made. However, as the monetary incentives attached to such decisions can be substantial, there will always be a temptation for employees, at junior as well as senior level, to 'follow the money'. The frequently reported 'sharp practices' by banks and the resulting public condemnation of their actions show that the incentives discussed in this chapter are very much present in banks.

Chapter 29

Bank strategy

Banks have to develop a strategy how they seek to conduct their business. In very simple terms, they could focus on existing customers and improve their experience, which will in turn result in customer retention and customers using more of the products and services the bank offers. Following this strategy, banks will increase profits by obtaining higher revenue from existing customers and saving costs by retaining existing customers. Alternatively, they could seek to expand their business by attracting new customers and increase profits from a larger customer basis. Of course, both strategies can be followed simultaneously, however, given that resources overall are limited, banks need to decide on the optimal combination of these two strategies. In chapter 29.1 we will see how banks decide between attracting new borrowers and working with existing borrowers to reduce risks.

When analysing banks, the focus is usually on the profits of banks from conducting their business. Many companies, including banks, are also engaged in non-commercial activities, such as charity work and the sponsoring of cultural or sporting events at global, national, or local level. Even taking into account the marketing effect of such activities, they are often not commercially justifiable. How the use of such non-commercial activities may nevertheless be beneficial to banks is discussed in chapter 29.2.

29.1 Bank growth

The ability to bank employees to conduct their work is limited by the amount of time available to them. They might focus their efforts either on acquiring new customers or they might work to reduce the risk of default in loans they have already granted. Thus banks have to make a choice between making the bank safer, thus reducing the risk of loan defaults, or grow their business by attracting new loans.

We know that a fraction p of loans in the economy that will be repaid and are therefore risk-free, while a fraction $1-p$ of loans will default with certainty. Managers can now exert effort in identifying the qualities of these loans; such identification is not perfect, however, and a defaulting loan is identified correctly only with probability e_S. Risk-free loans are always identified correctly. The probability of granting loan that will not be repaid is given by $(1-p)(1-e_S)$, reflecting that the bank has provided a defaulting loan, $1-p$, wrongly identified as a risk-free loan, $1-e_S$. With risk-free loans given with probability p, the probability of granting any loan is $p+(1-p)(1-e_S)$ as obviously loans that will default are and are identified as such will never be granted. The probability of granting a risk-free loan, given that a loan is granted is given by

$$\hat{p} = \frac{p}{p+(1-p)(1-e_S)}. \tag{29.1}$$

The bank does not provide a loan with probability $(1-p)\,e_S$, consisting of defaulting loans that are identified as such. No loan might also be granted if no suitable other borrowers have been identified, or the manager has not even sought to identify additional borrowers. We assume that the reason for nor granting a loan, whether it is due to all loans being identified as going to default or no new borrowers have been identified, cannot be distinguished from each other.

Let us propose that the manager receives a salary w_0 if no loan is given, a salary $\overline{w}$ if the loan is repaid, and $\underline{w}$ if the loan is not repaid. For simplicity we assume that $\underline{w}=0$ and thus we have

$$\overline{w} > w_0 > \underline{w} = 0. \tag{29.2}$$

Apart from exerting effort to establish whether a loan is repaid, e_S, effort also needs to be exerted to acquire new borrowers and grow the bank, e_G, where the likelihood that a new loan is found is given by e_G. The amount of effort the manager overall can exert is limited such that $e_S+e_G=1$. We thus will focus our attention on the allocation of effort on reducing risk, e_S, and growing the banking business, e_G, while taking the overall effort of managers as given. Exerting effort imposes a cost C on managers and it is not observable if in which combination of e_S and e_G effort is being exerted.

We furthermore assume that banks can only repay their deposits D, on which they rely to finance their loan L, with interest r_D, if the loan they grant is repaid. An abundance of deposits implies perfect competition between depositors and thus $\Pi_D = \hat{p}\,(1+r_D)\,L - L = 0$. The loan is repaid with probability p, that is if the loan granted is the risk-free loan. We can thus determine the deposit rate as

$$1+r_D = \frac{p+(1-p)(1-e_S)}{p}. \tag{29.3}$$

The remuneration of the manager is given by

$$\begin{aligned}\Pi_M = {} & e_G\left(p\overline{w} + (1-p)\,e_S w_0 + (1-p)(1-e_S)\,\underline{w}\right) \\ & + (1-e_G)\,w_0 - C.\end{aligned} \tag{29.4}$$

The last term denotes the costs of effort and the penultimate term the wage if despite effort no new loan can found. The first term encompasses the case where a new loan is granted and a risk-free loan ($p\overline{w}$), or no loans are given as only defaulting loans have been found, $(1-p)\,e_S w_0$, or when loans are granted but they default, $(1-p)\,((1-e_S)\,\underline{w}$; As we assumed that $\underline{w}=0$, this final expression will be ignored.

The manager would allocated optimally by maximizing the effort to reduce the risk of loans, giving us the first order condition

$$\frac{\partial \Pi_M}{\partial e_S} = -\left(p\overline{w} + (1-p)\,e_S w_0\right) + (1-e_S)\,(1-p)\,w_0 + w_0 = 0,$$

where we took into account that $e_S + e_G = 1$. The optimal effort levels to reduce the risk of the bank and grow it, respectively, are given by

$$\begin{aligned} e_S &= \frac{2-p}{2\,(1-p)} - \frac{p}{2\,(1-p)}\frac{\overline{w}}{w_0}, \\ e_G &= \frac{p}{2\,(1-p)}\left(\frac{\overline{w}}{w_0} - 1\right). \end{aligned} \tag{29.5}$$

We see that a larger remuneration in the case of a repaid loan, $\overline{w}$, allows the manager to put more emphasis on growing the business through new loans. This higher remuneration then compensates him for the lower probability of granting risk-free loans, which are given in higher volumes.

In order for managers to exert any effort, the profits they make, need to exceed that of making no effort at all by not seeking to acquire any loan and hence not facing any defaults; in this case they receive a salary of w_0 for sure. A competitive market for managers implies that the remuneration they receive will be the same as if making no effort and collecting the given remuneration, thus $\Pi_M = w_0$, which solves for

$$\frac{\overline{w}}{w_0} = 1 + \frac{4C\,(1-p)}{p^2\,(\overline{w} - w_0)}. \tag{29.6}$$

We see that the bank needs to maintain a sufficiently high remuneration for managers whose loans are repaid such that they are willing to exert additional effort to acquire new borrowers.

The profits of the bank lending at an interest rate r_L are then easily given by

$$\begin{aligned} \Pi_B = {} & e_G\,\big(p\,((1+r_L)\,L - (1+r_D)\,L - \overline{w}) \\ & - (1-p)\,e_S w_0 - (1-p)\,(1-e_S)\,\underline{w}\big) \\ & - (1-e_G)\,w_0. \end{aligned} \tag{29.7}$$

The first term covers the case that a loan is found, where it can be successful and be repaid, no loan is given as it is identified as defaulting and only the salary w_0 paid, or the loan is unsuccessful; as $\underline{w} = 0$, this last expression will not be considered further. The second term covers the case that no loan is found and they obtain the base salary. We know that managers will exert effort as the bank ensures salary is paid in accordance with equation (29.6).

We can now establish the optimal solutions for the remuneration of managers and use this to determine the distribution of the optimal effort level between ensuring the safety of the bank and its growth.

Banks ensure effort and direct effort allocation As a benchmark let us assume that the bank can determine and contractually enforce that effort is exerted and how it is allocated. In this case the bank can pay a fixed salary of C to the manager to compensate them for their costs, we have the bank profits given by

$$\Pi_B = e_G \left(p \left(1 + r_L\right) L - \left(1 + r_D\right) L\right) - C, \tag{29.8}$$

such that the optimum effort level after inserting from equation (29.3) for the deposit rate needs to fulfill the first order condition

$$\frac{\partial \Pi_B}{\partial e_G} = pr_L - 2\left(1 - p\right) e_G = 0, \tag{29.9}$$

which solves for

$$e_G^* = \frac{p}{2\left(1 - p\right)} r_L. \tag{29.10}$$

If the effort cannot be directly controlled by the bank, it needs to ensure to use the salaries $\overline{w}$ and w_0 to achieve its optimum. Equating equations (29.5) and (29.10), we easily obtain that

$$\frac{\overline{w}}{w_0} = 1 + \frac{2\left(1 - p\right)}{p} e_G^* = 1 + r_L. \tag{29.11}$$

To ensure the manager participates we need $\Pi_M = 0$; as managers cannot refuse to exert effort we onlyneed to ensure that their remuneration covers their costs, but not the reward theyw ould receive from not exerting any effort. From equation (29.4), using equations (29.10) and (29.11), this condition then becomes

$$w_0^* = \frac{4\left(1 - p\right)^2}{4\left(1 - p\right) - \left(2 - p\right)^2 pr_L} C. \tag{29.12}$$

We have thereby established the optimal level of effort put into growing the company and the remuneration managers need to obtain in order to implement this optimum.

Banks direct effort allocation If the bank, however, cannot ensure that managers exert effort, then to ensure managers do so, the bank needs to set the remuneration such that managers are willing to do so, thus we require $\Pi_M = w_0$. The incentives for the bank and the manager need to be aligned, thus equations (29.6) and (29.11) need to be equal, which solves for the remunerations to be set at

$$\overline{w} = w_0 + \frac{2C}{pe_G}, \tag{29.13}$$
$$w_0 = \frac{C}{(1-p)\,e_G^2}$$

Noting the expression for Π_M in equation (29.4) and that $\Pi_M = w_0$ we can rewrite equation (29.7) as

$$\Pi_B = pe_G\left((1+r_L)\,L - (1+r_D)\,L\right) - (w_0 + C)\,. \tag{29.14}$$

Inserting for w_0 from equation (29.13) and the deposit rate $1 + r_D$ from equation (29.3), the first order condition for maximizing the bank profits becomes

$$\frac{\partial \Pi_B}{\partial e_G} = pr_L - 2\,(1-p)\,e_G + \frac{2C}{(1-p)\,e_G^3} = 0, \tag{29.15}$$

which is identical to the expression in the first order condition of equation (29.9), apart from the last term, which makes this expression positive for any $e_G^* > 0$. As $\frac{\partial^2 \Pi_B}{\partial e_G^2} < 0$, this implies that for if banks cannot compel managers to exert effort, the optimal effort put on growth e_G will be larger than in the case where effort by managers is ensured. Thus by having to provide managers with incentives to exert effort, they will put a higher emphasis on acquiring new borrowers and less effort on ensuring only risk-free loans are given. From equation (29.11) we also see that $\frac{\overline{w}}{w_0}$ is higher here, thus, success is rewarded more than in the case where effort can be contracted for.

Competition between banks Thus far we had implicitly assumed that the bank is a monopolist and exerting effort to attract new borrowers will be instantly successful. However, if banks compete with each other this will affect their ability to identify new borrowers. With two banks, i and j, we define $\widetilde{e}_G^i = e_G^i - \lambda e_G^j$, where $\lambda \in [0;1]$ indicates the degree of competition between banks. In the absence of competition, $\lambda = 0$, the previous case of a single bank emerges, and as λ increases competition increases and any effort by banks to attract new borrowers will be reduced by the effort of the other bank doing the same.

For managers, their remuneration becomes

$$\Pi_M^i = \widetilde{e}_G^i\left(p\overline{w} + (1-p)\,e_S^i w_0\right) + \left(1 - \widetilde{e}_G^i\right) w_0 - C, \tag{29.16}$$

having replaced e_G by $\widetilde{e}_G^i$ in equation (29.4) and used that $\underline{w} = 0$. This gives us the first order condition for the optimal allocation of effort for both banks as

$$\begin{aligned}\frac{\partial \Pi_M^i}{\partial e_G^i} &= p\overline{w} + (1-p)\left(1-e_G^i\right)w_0 \\ &\quad - \left(e_G^i - \lambda e_G^j\right)(1-p)\,w_0 - w_0 = 0, \\ \frac{\partial \Pi_M^j}{\partial e_G^j} &= p\overline{w} + (1-p)\left(1-e_G^j\right)w_0 \\ &\quad - \left(e_G^j - \lambda e_G^i\right)(1-p)\,w_0 - w_0 = 0.\end{aligned} \tag{29.17}$$

which solve for

$$\begin{aligned}e_G^i &= \frac{p}{2(1-p)}\frac{\overline{w}}{w_0} - \frac{p}{2(1-p)} + \frac{\lambda}{2}e_G^j, \\ e_G^j &= \frac{p}{2(1-p)}\frac{\overline{w}}{w_0} - \frac{p}{2(1-p)} + \frac{\lambda}{2}e_G^i.\end{aligned} \tag{29.18}$$

These two equations can be rewritten as

$$e_G^i - \frac{\lambda}{2}e_G^j = \frac{p}{2(1-p)}\left(\frac{\overline{w}}{w_0} - 1\right). \tag{29.19}$$

As the banks are ex-ante identical, we would expect them to exert the same level of effort in equilibrium, $e_G^i = e_G^j$, such that

$$e_G^i = e_G^j = \frac{2}{2-\lambda}\frac{p}{2(1-p)}\left(\frac{\overline{w}}{w_0} - 1\right), \tag{29.20}$$

which implies from comparing with equation (29.5) that the effort level for growth, e_G, is increasing in competition, λ. Competition to attract borrowers leads to more efforts into this area as some of the effort is off-set by the equivalent effort of the other bank. This higher effort put into growing the bank, will come at the cost of evaluating those borrowers and hence the safety of the bank.

The participation constraint for managers, $\Pi_M = w_0$, implies that when using $\widetilde{e}_G^i$ instead of e_G and inserting from equation (29.20), we get

$$\frac{\overline{w}}{w_0} = 1 + 2\frac{2-\lambda}{1-\lambda}\frac{1-p}{p^2}\frac{C}{\overline{w} - w_0}. \tag{29.21}$$

A comparison with equation (29.6) shows that as competition increases, the reward for success, $\overline{w}$, is increasing in competition. This is done as compensation for attracting fewer new borrowers despite putting more effort in, while more loan will default as less effort is put into evaluating borrowers. In order to attract managers to the bank, the remuneration in case of the loan being repaid needs to be increased, relative to the base salary, w_0. Looking back at equation (29.20), we see that this increase in remuneration for repaid loans further increases the incentives to attract new borrowers, e_G^i.

We derive from equation (29.15) that

$$\frac{\partial^2 \Pi_B}{\partial e_G^2} = -2(1-p) - \frac{6C}{(1-p)\, e_G^4} < 0,$$
$$\frac{\partial^2 \Pi_B}{\partial e_G \partial p} = r_L + 2e_G + \frac{2C}{(1-p)^2\, e_G^3} > 0,$$

and hence using the implicit function theorem, we can derive that

$$\frac{\partial e_G}{\partial p} = -\frac{\frac{\partial^2 \Pi_B}{\partial e_G \partial p}}{\frac{\partial^2 \Pi_B}{\partial e_G^2}} > 0. \tag{29.22}$$

Thus, if banks that are more optimistic by assuming that risk-free loans are more likely, they will put more emphasis on attracting new customers as the threat of losses is reduced. This relationship is maintained in the presence of competition between banks.

Matching banks and managers Rather assuming that banks and managers agree in their assessment of the prevalence of risk-free loans, p, we can now assume that they may have different opinions on the market by assuming that $p_M \neq p_B$. Thus the fraction of risk-free loans in the market is evaluated differently by managers, p_M, and banks, p_B.

In this case equation (29.5) becomes

$$e_G = \frac{p_M}{2(1-p_M)} \left(\frac{\overline{w}}{w_0} - 1 \right) \tag{29.23}$$

as the effort managers put into growing the bank is only determined by the manager's assessment. Similarly equation (29.13) becomes

$$w_0 = \frac{C}{(1-p_M)\, e_G^2}. \tag{29.24}$$

We now have with depositors and banks agreeing that the likelihood of a risk-free loan being p_B the profits of the bank given by

$$\begin{aligned}\Pi_B &= e_G \left(p_B \left((1+r_L) L - (1+r_D) L - \overline{w}\right)\right. \\ &\quad \left. - (1-p_B)(1-e_G) w_0\right) - (1-e_G) w_0 \\ &= e_G \left(p_B (1+r_L) L - (p_B + (1-p_B) e_G) L\right) \\ &\quad - \left(e_G (p_B - p_M) \overline{w} + e_G (p_M - p_B) w_0 (1-e_G)\right. \\ &\quad \left. + \Pi_M + C\right) \\ &= e_G \left(p_B r_L L - (1-p_B) e_G L\right) - w_0 - C \\ &\quad - \frac{(2-p_M)(p_M - p_B)}{p_M (1-p_M)} C,\end{aligned} \tag{29.25}$$

when inserting from equations (29.23) and (29.24). After inserting for w_0 from equation (29.24), we can get the first order condition of the bank maximizing its profits as

$$\frac{\partial \Pi_B}{\partial e_G} = p_B r_L L - 2(1-p_B) e_G L + \frac{2C}{(1-p_M) e_G^3} = 0. \tag{29.26}$$

If p_M increases, the last term decreases and hence the effort put into attracting new borrowers, e_G, needs to increase; we thus find that $\frac{\partial e_G}{\partial p_M} > 0$.

From equation (29.26) we furthermore have

$$\begin{aligned}\frac{\partial^2 \Pi_B}{\partial e_G \partial p_B} &= r_L L + e_G L > 0, \\ \frac{\partial^2 \Pi_B}{\partial e_G^2} &= -(1-p_B) L - \frac{6C}{(1-p_M) e_G^4} > 0.\end{aligned}$$

and thus by the implicit function theorem

$$\frac{\partial e_G}{\partial p_B} = -\frac{\frac{\partial^2 \Pi_B}{\partial e_G \partial p_B}}{\frac{\partial^2 \Pi_B}{\partial e_G^2}} > 0, \tag{29.27}$$

such as above, the effort to attract new borrowers is increasing as the bank becomes more optimistic.

From equation (29.24) it is now obvious that the basic salary w_0 is decreasing in the probability of obtaining a risk-free loan, p_B; this effect originates in the influence on e_G. As the remuneration of the manager is given by this base salary as we assumed that managers are rewarded only to exert effort,$\Pi_M = w_0$, they prefer banks that assume a low fraction of risk-free loans, p_B, as this increases their remuneration.

Let now e_G denote the solution to the optimization for the given values of p_M and p_B in equation (29.26). If $\hat{e}_G$ is the solution for $\hat{p}_M$ and p_B, where $\hat{p}_M > p_M$, then as $\hat{e}_G$ is not optimal for p_M and p_B, the profits to the bank would be lower. From equation (29.25) we thus get

$$\begin{aligned}
\Pi_B &= e_G \left(p_B r_L L - (1-p_B)\, e_G L\right) - \frac{C}{(1-p_M)\, e_G^2} - C - \\
&\quad \frac{(2-p_M)(p_M-p_B)}{p_M(1-p_M)} C \\
&\geq \hat{e}_G \left(p_B r_L L - (1-p_B)\, \hat{e}_G L\right) - \frac{C}{(1-p_M)\, \hat{e}_G^2} - C \\
&\quad - \frac{(2-p_M)(p_M-p_B)}{p_M(1-p_M)} C \\
&\geq \hat{e}_G \left(p_B r_L L - (1-p_B)\, \hat{e}_G L\right) - \frac{C}{(1-\hat{p}_M)\, \hat{e}_G^2} - C \\
&\quad - \frac{(2-\hat{p}_M)(\hat{p}_M-p_B)}{\hat{p}_M(1-\hat{p}_M)} C \\
&= \hat{\Pi}_B
\end{aligned} \tag{29.28}$$

The last inequality arises from the third term increasing when increasing p_M and for the last term we find

$$\frac{\partial \frac{(2-p_M)(p_M-p_B)}{p_M(1-p_M)}}{\partial p_M} = \frac{(2p_B+1)\left(\left(p_M - 2\frac{p_B}{2p_B+1}\right)^2 + \frac{2p_B}{2p_B+1}\right)}{p_M^2\,(1-p_M)^2} > 0.$$

Hence the last term is increasing in p_M. This leads to the conclusion that banks prefer managers that expect risk-free loans to be less common, a low p_M.

With both the banks and the managers wanting to secure partners with the lowest possible opinions on the availability of risk-free loans, p_i, it will be that the lowest in each group are matched, then the second lowest, and so on. Assuming that and such ordered pair the respective probabilities are matched, $p_M = p_B = p$, we see that we match banks and managers with the same outlook.

Competition between heterogeneous banks We have thus established that banks and managers are matched in their assessment of the availability of risk-free loan. We can now extend the competition between banks to the kase of two banks with different assessments on the availability of risky loans competing. Let us therefore assume we have two banks, i and j, with different outlooks on the fraction of successful loans, $p_i > p_j$, on which the bank and their manager agree.

We can now rewrite the profits of the bank from equation (29.16) as

$$\begin{aligned}
\Pi_B^i &= \left(e_G^i - \lambda e_G^j\right)\left(p_i\,(1+r_L)\, L - \left(p_i + (1-p_i)\, e_G^i\right) L\right) \\
&\quad - \left(w_0^i + C\right),
\end{aligned} \tag{29.29}$$

the last term emerging from collecting all salary terms and noting that $\Pi_M^i = w_0^i$, where w_0^i is given in equation (29.13). The first order condition when maximizing bank profits over the optimal effort in attracting new borrowers is then given by

$$\frac{\partial \Pi_B^i}{\partial e_G^i} = p_i \left((1 + r_L) L - (1 + r_D) L \right) - 2 \left(e_G^i - \lambda e_G^j \right) (1 - p_i) \quad (29.30)$$
$$+ \frac{2C}{(1 - p_i) \left(e_G^i - \lambda e G^j \right)^3} = 0,$$
$$\frac{\partial \Pi_B^l}{\partial e_G^l} = p_j \left((1 + r_L) L - (1 + r_D) L \right) - 2 \left(e_G^j - \lambda e_G^i \right) \left(1 - p_j\right)$$
$$+ \frac{2C}{\left(1 - p_j\right) \left(e_G^j - \lambda e G^i \right)^3} = 0.$$

If we assume that $p_i > p_j$, the first term in the first condition is larger than in the second condition. Let us further assume $e_G^i < e_G^j$, then the second term is smaller and the last term larger again, hence overall the first condition would be larger than the second condition. Therefore both condition cannot be fulfilled simultaneously with these assumptions. Hence, for a solution we require that

$$e_G^i > e_G^j \quad (29.31)$$

under our assumption that $p_i > p_j$. Thus, the bank which has a more optimistic assessment about the risk of loans will seek to exert more effort into attracting new borrowers.

Let us compare this result with a situation with both banks are equally optimistic, i.e. $p_i = p_j$ and hence $e_G^i = e_G^j$. Thus for bank j we get equation (29.30) as

$$\frac{\partial \Pi_B^j}{\partial e_G^j} = p_j \left((1 + r_L) L - (1 + r_D) L \right) - (2 - \lambda) \left(1 - p_j\right) e_G^j L \quad (29.32)$$
$$+ \frac{1}{(1 - \lambda)^3} \frac{2C}{1 - p_j} \frac{1}{\left(e_G^j \right)^3} = 0$$

and for banks with p_i if they can direct the allocation of effort, it is from equation (29.15) that

$$\frac{\partial \Pi_B^i}{\partial e_G^i} = p_i \left((1 + r_L) L - (1 + r_D) L \right) - 2 (1 - p_i) e_G^i L \quad (29.33)$$
$$+ \frac{2C}{(1 - p_i) \left(e_G^i \right)^3} = 0.$$

We immediately see that the first term in equation (29.33) is larger than in equation (29.32) and for the second term this relationship is reversed when using $e_G^i > e)G^j$ from equation (29.31); thus for the first two terms the expression in equation (29.33) is larger. As λ now increases, the first equation becomes larger, necessitating e_G^j to increase. There will be a $\lambda = \lambda^*$ such that e_G^j becomes larger than the solution to equation (29.33). Therefore, competition between banks can cause even the more pessimistic bank to become more growth focused (a high allocation of effort towards e_G) than an optimistic bank without competition. From equation (29.31) we see that the more optimistic bank will become even more growth focused.

Summary We have established that banks who are assessing loans as being less risky, will focus more of their efforts on growing their loan portfolio. This is the result of incentives given to managers in allocating their effort between reducing risks and attracting new borrowers. If loans are seen as less risky, the value of reducing risks further is small and hence such activities would not be rewarded much by the bank. Instead, the bank would prefer managers to focus on attracting additional borrowers and thereby increase profits. It is thus that banks which believe to have a low-risk loan portfolio will focus on the growth of this loan portfolio rather than risk management, while a bank which assesses the risk of their loan portfolio as being higher, will put a larger emphasis on risk management. Competition between banks makes attracting new borrowers more difficult as the efforts of banks offset each other partially and this property of banks seeing themselves as low-risk focussing on growth will become even more pronounced, to an extend that even a bank with high risks may focus mainly on growth.

Such incentives can lead to a much more risk-based culture at the cost of safety in lending. Banks will divert resources from managing the risk of lending towards gaining market share and increasing competition makes this process even more pronounced. Thus, the riskiness of loans a bank grants will affect its strategy; in times of low risks it will seek to expand its business, often enhanced by increasing competition from banks focussing more on market expansion, building up large loan portfolios without assessing the risks of such loans well. With risk being seen as low in times of high economic growth, loans will become ever more readily available and the economic expansion will continue. If the economic conditions change and risks increase, banks will become more cautious and this might lead to a contraction in lending, making any economic downturn more severe. In addition, banks will also be exposed to a large loan portfolio with risks larger than anticipated, increasing the risk of a bank failing.

Reading Song & Thakor (2019)

29.2 Non-commercial activities

We usually consider banks that are seeking to maximize their profits from lending and accepting deposits. However, many banks engage in charitable activities, sponsor museums, exhibitions, music concerts, or sports events. While some of these sponsorships can be interpreted as marketing and thus be done to maximize the profits of the bank, the returns are likely to be lower than the costs on many occasions. We will analyse here under which conditions such sponsorships, which are non-commercial activities, are beneficial to banks.

We have two types of banks, one type has a higher ability to manage loan defaults and has a basic repayment rate of π_H, while the other type of banks has a lower ability and their basic repayment rate on loans is $\pi_L < \pi_H$. The actual repayment rate $\hat{\pi}_i$ also depends on the level of effort by the managers of the bank, e_i, such that $\hat{\pi}_i = \pi_i e_i$, where $0 \leq e_i \leq 1$. If the loan is repaid, the manager receives his wages w_i and effort is costly with costs c such that the benefits to a manager is

$$\Pi_M^i = \hat{\pi}_i w_i - \frac{1}{2} c e_i^2 = \pi_i e_i W_i - \frac{1}{2} c e_i^2, \tag{29.34}$$

if we assume that the manager knows the type of bank he is working for. Hence the optimal effort level is obtained from the first order condition

$$\frac{\partial \Pi_M^i}{\partial e_i} = \pi_i w_i - c e_i = 0 \tag{29.35}$$

from which we obtain the optimal effort level as

$$e_i = \frac{\pi_i}{c} w_i. \tag{29.36}$$

Banks provide loans L with a loan rate of r_L and finance this with deposits D requiring a deposit rate of r_D, where we assume that deposits are insured and hence the deposit rate is not dependent on the bank type. Similarly, the loan rate does not depend on the bank type as we assume that the different repayment rates are the result of the ability of banks to support their borrowers.

Known bank type If banks are repaid the loan, they divert a fraction α_i of their profits to non-commercial causes, from which they only gain a smaller benefit $\lambda < 1$. This smaller return represents the benefits of increased publicity, for example. Thus the profits of the bank are given by

$$\begin{aligned} \Pi_B^i &= \hat{\pi}_i \left((1 - \alpha_i) \left((1 + r_L) L - (1 + r_D) D - w_i \right) \right. \\ &\quad \left. + \alpha_i \lambda \left((1 + r_L) L - (1 + r_D) D - w_i \right) \right) \\ &= \hat{\pi}_i \left(1 - \alpha_i (1 - \lambda) \right) \left((1 + r_L) L - (1 + r_D) D - w_i \right) \\ &= \frac{\pi_i^2}{c} w_i \left(1 - \alpha_i (1 - \lambda) \right) \left((1 + r_L) L - (1 + r_D) D - w_i \right), \end{aligned} \tag{29.37}$$

where we have inserted the optimal effort level of managers from equation (29.36). The first term denotes the profits of the bank retained and the second term the return on the non-commercial activities. The optimal salary is thenobtained from the first order condition

$$\frac{\partial \Pi_B^i}{\partial w_i} = \frac{\pi_i^2}{c} \left(1 - \alpha_i (1 - \lambda) \right) \left((1 + r_L) L - (1 + r_D) D - 2 w_i \right) = 0, \tag{29.38}$$

which can be solved for

$$w_i = \frac{1}{2}\left(\left(1 + r_L\right) L - \left(1 + r_D\right) D\right). \tag{29.39}$$

We see that the salary is determined independent of the type of the bank, π_i, as well as the degree of non-commercial activities, α_i.

Inserting this into the profit function of equation (29.37), we obtain the opitmal profits as

$$\Pi_B^i = \frac{\pi_i^2}{4c}\left(1 - \alpha_i\left(1 - \lambda\right)\right)\left(\left(1 + r_L\right) L - \left(1 + r_D\right) D\right)^2. \tag{29.40}$$

As $\lambda < 1$, it is obvious that the profits are highest if no non-commercial activities are conducted, thus $\alpha_i = 0$. In this case profits are

$$\Pi_B^i = \frac{\pi_i^2}{4c}\left(\left(1 + r_L\right) L - \left(1 + r_D\right) D\right)^2. \tag{29.41}$$

Unknown bank type Thus far we assumed that the type of bank was known, but we now turn to the case where the type of bank is now known. Suppose that the bank with a high repayment rate of loans, π_H occurs with probability p and the repayment rate is low, π_L with probability $1 - p$. If the manager does not know the type of bank, then his profits are given by

$$\begin{aligned} \Pi_M &= \left(p\hat{\pi}_H + (1 - p)\,\hat{\pi}_L\right) w_i - \frac{1}{2}ce^2 \\ &= \left(p\pi_H + (1 - p)\,\pi_L\right) e w_i - \frac{1}{2}ce^2 \end{aligned} \tag{29.42}$$

and the first order condition for the optimal effort level is given by

$$\frac{\partial \Pi_M}{\partial e} = \left(p\pi_H + (1 - p)\,\pi_L\right) w_i - ce = 0, \tag{29.43}$$

which solves for

$$e = \frac{p\pi_H + (1 - p)\,\pi_L}{c} w_i. \tag{29.44}$$

Using the optimal effort level from equation (29.44) in the expression $\hat{\pi}_i = \pi_i e$ in the first line of the bank profits in equation (29.37), we have

$$\Pi_B^i = \pi_i \frac{p\pi_H + (1 - p)\,\pi_L}{c} w\left(1 - \alpha_i\left(1 - \lambda\right)\right)\left(\left(1 + r_L\right) L - \left(1 + r_D\right) D - w_i\right). \tag{29.45}$$

We assume that the bank knows its own type, the optimal salary can be determined from the first order condition

$$\frac{\partial \Pi_B^i}{\partial w} = \pi_i \frac{p\pi_H + (1-p)\pi_L}{c} (1 - \alpha_i (1-\lambda)) ((1+r_L) - (1+rD) D - 2w) = 0 \tag{29.46}$$

which becomes

$$w_i = \frac{1}{2} ((1+r_L) L - (1+r_D) D), \tag{29.47}$$

which is identical to the previous case where managers knew the type of bank.

The optimal level of non-commercial causes is again $\alpha_i = 0$ and this gives us together with equation (29.47) from equation (29.45) the optimal profits as

$$\Pi_B^i = \pi_i \frac{p\pi_H + (1-p)\pi_L}{4c} ((1+r_L) L - (1+r_d) D)^2. \tag{29.48}$$

Let us now assume that the bank with the low success rate pi_L tries to mimic the behaviour of the bank with the high success rate, π_H, by setting engaging in some non-commercial activities, $\alpha_L = \alpha_H > 0$, and extracting the same level of effort from managers, $e_L = e_H = \frac{\pi_H}{c} w_H$. This gives us the success rate of this bank as

$$\hat{\pi}_L = \pi_L e_H = \frac{\pi_L \pi_H ((1+r_L) L - (1+r_D) D)}{2c} \tag{29.49}$$

and hence bank profits of

$$\begin{aligned} \Pi_B^L &= \hat{\pi}_L (1 - \alpha_H (1-\lambda_L)) ((1+r_L) L - (1+r_D) D - W_H) \\ &= \frac{\pi_L \pi_H}{4c} (1 - \alpha_H (1-\lambda_L)) ((1+r_L) L - (1+r_D) D)^2. \end{aligned} \tag{29.50}$$

If the bank with the low success rate, π_L does not mimic the other bank and sets $\alpha_L = 0$, then the bank type is known due to it using a different level of commercial activities, and we get from equation (29.41) the bank profits as

$$\Pi_B^L = \frac{\pi_L^2}{4c} ((1+r_L) L - (1+r_d) D)^2. \tag{29.51}$$

Hence if the expression in equation (29.51) is larger than the expression in equation (29.50), the bank with the low success rate would not mimic the behaviour of the other bank; this is the case if

$$\alpha_H \geq \alpha_H^* = \frac{\pi_H - \pi_L}{\pi_H (1-\lambda_L)}. \tag{29.52}$$

If the non-commercial activities are sufficiently high, the bank with a low success rate will not mimic its rival.

From equation (29.40) the profits of the bank with the higher success rate, π_H is then given by

$$\Pi_B^H = \frac{\pi_H^2}{4c} (1 - \alpha_H (1-\lambda_H)) ((1+r_L) L - (1+r_D) D)^2. \tag{29.53}$$

If this were to not set $\alpha_H > 0$, but set $\alpha_H = 0$, the two types of banks cannot be distinguished. The profits of this bank is then given by equation (29.48) and we have

$$\Pi_B^H = \pi_H \frac{p\pi_H + (1-p)\,\pi_L}{4c}\left((1+r_L)\,L - (1+r_D)\,D\right)^2 . \tag{29.54}$$

Thus the banks with the high success rate will set $\alpha_H > 0$ if the expression in equation (29.53) exceeds that in equation (29.54). This condition requires $p\,(1-\lambda) \geq 0$, which is always fulfilled.

Hence we can obtain a separating equilibrium if the banks with a high success rate engages in non-commercial activities to the extent that $\alpha_H = \alpha^*$, given that the bank would not engage more than necessary to distinguish itself from other banks because non-commercial activities reduce their profits. In this case the non-commercial activity can be seen as a sign that the bank is highly profitable and this information induces managers to exert higher effort than they would do without this knowledge.

Summary We have seen that highly profitable banks can use non-commercial activities to signal that their success, which is otherwise not observable. The high profit margin of these type of banks is used to engage in loss-making non-commercial activities, which less well performing banks cannot afford to. Managers can use this information to learn the type of the bank they are working for and as a consequence will exert more effort, more than compensating the bank for the losses from non-commercial activities.

It is therefore that we might observe highly profitable banks engaging in the sponsoring of cultural or sporting events, but less well performing banks will not do so. Such activities might not have its origin in any charitable goals of the bank, but due to its signalling effects, will increase the profits of the bank. Thus their engagement, although on first sight a loss-making activity, will benefit the bank and increase its profits.

Reading Bunderson & Thakor (2021)

Conclusions

We have seen that in a more competitive environment banks will focus their resources more on expanding their loan portfolio than managing existing risks. This is done in part to attract additional profits from new customers, but resources are also expended to prevent customers from being attracted by other banks. As all banks seek to attract new customers, a lot of the effort to grow the loan portfolio is offset by the same activities of other banks. This leaves less than optimal amount of effort available to manage the risks of the loan the bank has granted. Thus, competition between banks induces an overly strong focus on growth and too little emphasis on the management of risks from existing customers. Such behaviour will increase the risks of banks and the banking system as a whole.

Despite the emphasis on profits by banks, they engage in non-commercial activities that are loss-making to banks. Nevertheless, banks may obtain a benefit if it allows them to signal to the market that their profitability is high and they therefore can afford to engage in such activities, while banks with a lower profitability would not be able to do so. The value to the bank of making it known that they are generating high profits might be in an increased stock price, but also in managers exerting higher efforts as they know they will participate in the profits the bank generates. If the profitability of banks is low, these benefits will now outweigh the costs. It can therefore be rational for banks to sponsor music festivals, concerts, exhibitions, or engage in charitable activities.

Chapter 30

Ownership structure of banks

It is common to assume that banks are organised as companies owned by individuals which are unconnected to the provision of loans or the accepting of deposits. It is then that maximizing the profits of such banks can be seen as maximizing the profits of the stakeholders in the bank, their owners. However, alternative forms of banks exist, the most common is the mutual bank, or cooperative bank. Such banks are owned not by outside owners but by those individuals who seek loans and provide the bank with deposits. It is therefore that maximizing bank profits would give a distorted result, the bank profits should be included in the outcomes of borrowers and depositors. We will investigate the implications for the riskiness of loans a mutual bank is willing to provide in chapter 30.1, along with a discussion who would want to borrow from mutual banks rather than conventional banks.

It is often also that banks are at least partially owned by government and politicians yield significant influence on loan decisions. We will explore in chapter 30.2 how such political interference in lending decisions can lead to the provisions of loans that are politically desirable but not sustainable for banks seeking to maximize their profits.

30.1 Mutual banks

Mutual banks are not owned by shareholders or other outside owners, but by their customers themselves. A characteristic of mutual banks is that customers are receiving a dividend from the bank based on the profits the bank makes; such an arrangement would not be found in conventional banks where dividends are paid to the outside owners of the bank. Such mutual banks co-exist in many countries alongside conventional banks. We will explore in chapter 30.1.1 which type of customers prefer to use mutual banks over conventional banks and chapter 30.1.2 investigate the behaviour of mutual banks in comparison to conventional banks in terms of the risks they are willing to take.

30.1.1 Beneficiaries of mutual banks

Mutual banks are mainly used by individual and small companies and most customers have incomes that make them financially comfortable and can thus be classified as seeking low-risk loans. We will explore why such customers benefit from mutual banks, while more risky customers are preferring conventional banks.

There are two types of borrowers who differ in their ability to repay their loan L_i, given at a given loan rate r_L^i. For type repays their loans with probability π_H^j and the other type with a lower probability $\pi_L^j < \pi_H^j$, given the state of the economy, j. There are two possible states such that for borrower type i we have $\pi_i^H > \pi_i^L$, where state H, a well performing economy, occurs with probability p and state L, where the economy performs less well occurs with probability $1 - p$. Borrowers know their own types, but banks do not know the borrowers' types and the state of the economy is unknown to both, banks and borrowers.

Having lent L_i to borrower i in state j, the bank makes profits from lending of

$$\hat{\Pi}_B^{ij} = \left(\pi_i^j \left(1 + r_L^i\right) - \left(1 + r_D\right)\right) L_i. \tag{30.1}$$

A mutual bank will pay out a fraction α_i^j of their lending profits to the borrower, provided the loan was repaid; the bank retains the profits that have not been paid out to their customers. The expected profits of the bank, taking into account the different states of the economy are then

$$\Pi_B^i = p\left(1 - \pi_i^H \alpha_i^H\right) \hat{\Pi}_B^{iH} + (1 - p)\left(1 - \pi_i^L \alpha_i^L\right) \hat{\Pi}_B^{iL}. \tag{30.2}$$

We now assume that the borrower has an existing investment that cannot be liquidated and hence seeks a loan to finance his consumption. In time period 1 he can consume his loan L_i and in time period 2 the proceeds of his investment $(1 + R) I$, less the repayment of the loan and added to by the profits-sharing of the bank. If the investment is not successful, the borrower receives nothing. Hence

$$\begin{aligned} C_1^i &= L_i, \\ C_2^{ij} &= \begin{cases} (1 + R) I - \left(1 + r_L^i\right) L_i + \alpha_i^j \hat{\Pi}_B^{ij} & \text{w.p.} \quad \pi_i^j \\ 0 & \text{w.p.} \ 1 - \pi_i^j \end{cases}. \end{aligned} \tag{30.3}$$

Borrowers discount future consumption by ρ such that their utility is given by

$$U^i = C_i^i + \rho E\left[C_2^{ij}\right]. \tag{30.4}$$

A low-risk (hight-risk) borrower H (L) will choose the loan designed for his type if the utility of doing so exceeds that of choosing the loan designed for the other type; this would be possible because banks do not know the type of their borrowers. Using $E\left[C_2^{ij}\right] = p\pi_i^H C_2^{iH} + (1 - p)\pi_i^L C_2^{iL}$, where C_2^{iL} denotes the first line in equation (30.3), we get

$$C_1^H + \rho\left(p\pi_H^H C_2^{HH} + (1-p)\,\pi_H^L C_2^{HL}\right) \tag{30.5}$$
$$\geq C_1^L + \rho\left(p\pi_H^H C_2^{LH} + (1-p)\,\pi_H^L C_2^{LL}\right),$$

$$C_1^L + \rho\left(p\pi_L^H C_2^{LH} + (1-p)\,\pi_L^L C_2^{LL}\right) \tag{30.6}$$
$$\geq C_1^H + \rho\left(p\pi_L^H C_2^{HH} + (1-p)\,\pi_L^L C_2^{HL}\right).$$

Equations (30.1) and (30.3) we can rewrite as

$$\left(1-\pi_i^j\alpha_i^j\right)\hat{\Pi}_B^{ij} = \pi_i^j\left((1+R)\,I - C_2^{ij}\right) \tag{30.7}$$
$$-(1+r_D)\,C_1^i.$$

Using this expression in equation (30.2), we then get

$$\Pi_B^i = p\pi_i^H\left((1+R)\,I - C_e^{iH}\right) + (1-p)\,\pi_i^L\left((1+R)\,I - C_2^{iL}\right) - (1+r_D)\,C_1^i. \tag{30.8}$$

In a competitive market, bank profits would be eroded such that $\Pi_B^i = 0$. We can now investigate the optimal profit sharing between borrowers and the bank by determining the optimal value for α_i^j.

High-risk borrowers We will initially consider high-risk borrowers only. These borrowers seek to maximize their utility U^L subject to the constraint that banks make no profits from lending to them, $\Pi_B^L = 0$. The latter constraint implies from equation (30.8) that

$$p\pi_L^H C_2^{LH} + (1-p)\,\pi_L^L C_2^{LL} \tag{30.9}$$
$$= -(1+r_D)\,C_1^L + \left(p\pi_L^H + (1-p)\,\pi_L^L\right)(1+R)\,I,$$

where the right-hand side equals the expected consumption in time period 2, $E\left[C_2^{Lj}\right]$. Thus we find the utility of the high-risk borrower to be given by

$$U^L = (1-\rho\,(1+r_D))\,C_1^L + \left(p\pi_L^H + (1-p)\,\pi_l^L\right)(1+R)\,I. \tag{30.10}$$

Assuming $\rho < \frac{1}{1+r_D}$ we see that this expression is maximized if the first period consumptionC_i^L is maximized. From equation (30.9) we see that the value for the first period consumption C_1^L will be largest if $C_2^{LH} = C_2^{LL} = 0$ as in this case the full adjustment will fall on C_1^L. This then gives us directly that

$$C_1^L = \frac{p\pi_L^H + (1-p)\,\pi_L^L}{1+r_D}\,(1+R)\,I = L_L. \tag{30.11}$$

As the borrower has no funds available, it will have to fund this consumption through a loan, L_L.

Using that $C_2^{Lj} = 0$, we get from equation (30.3) that

$$(1+R)\,I - \left(1+r_L^L\right) L_L + \alpha_L^j \hat{\Pi}_B^{Lj} = 0, \tag{30.12}$$

and hence

$$\alpha_L^L \hat{\Pi}_B^{LL} = \alpha_L^H \hat{\Pi}_B^{LH}. \tag{30.13}$$

Suppose now that $\alpha_L^H > 0$ and $\alpha_L^L = 0$. From equation (30.13) we then need $\hat{\Pi}_B^{LH} = 0$, which when using equation (30.1) gives us $1 + r_L^L = \frac{1+r_D}{\pi_L^H}$. Using $C_2^{LH} = 0$, we get from equation (30.3) applied to equation (30.11) that

$$(1+R)\,I - \frac{p\pi_L^H + (1-p)\,\pi_L^L}{\pi_L^H}\,(1+R)\,I + \alpha_L^H \hat{\Pi}_B^{LH} = 0, \tag{30.14}$$

with the final term being zero from equation (30.13) such that we need

$$1 - \frac{p\pi_L^H + (1-p)\,\pi_L^L}{\pi_L^H} = 0 \tag{30.15}$$

or $\pi_L^H = \pi_L^L$, in contradiction to our assumption that $\pi_L^H > \pi_L^L$. We therefore can rule out that $\alpha_L^H > 0$ and $\alpha_L^L = 0$.

Secondly, suppose that $\alpha_L^H = 0$ and $\alpha_L^L > 0$. Then from equation (30.13) we get that $\hat{\Pi}_B^{LL} = 0$ and hence $1 + r_L^L = \frac{1+r_D}{\pi_L^L}$. Using now that $C_2^{LL} = 0$, we get again with equations (30.3) and (30.11) that

$$(1+R)\,I - \frac{p\pi_L^H + (1-p)\,\pi_L^L}{\pi_L^L}\,(1+R)\,I + \alpha_L^L \hat{\Pi}_B^{LL} = 0 \tag{30.16}$$

with the final term equal to zero from equation (30.13) and hence

$$1 - \frac{1 - p\pi_L^H\,(1-p)\,\pi_L^L}{\pi_L^L} = 0, \tag{30.17}$$

or $\pi_L^H = \pi_L^L$, again a contradiction to our assumption that $\pi_L^H > \pi_L^L$, ruling out that $\alpha_L^H = 0$ and $\alpha_L^L > 0$.

Finally, assume that $\alpha_L^H > 0$ and $\alpha_L^L > 0$. A solution to equation (30.13) would be that $\hat{\Pi}_B^{LH} = \hat{\Pi}_B^{LL} = 0$, which would then require $1 + r_L^L = \frac{1+r_D}{\pi_L^H} = \frac{1+r_D}{\pi_L^L}$, implying $\pi_L^H = \pi_L^L$, which can be ruled out again as we assumed that $\pi_L^H > \pi_L^L$. Thus we require $\hat{\Pi}_B^{Lj}$ to be non-negative, but it requires the same sign. Requiring $\Pi_B^L = 0$, we see from equation (30.2) that with $\hat{\Pi}_B^{Lj}$ having the same sign, this is only possible if

$$1 - \pi_L^H \alpha_L^H = 1 - \pi_L^L \alpha_L^L = 0 \tag{30.18}$$

or

$$\alpha_L^j = \frac{1}{\pi_L^j}. \tag{30.19}$$

Inserting into this result into equation (30.13) and using equation (30.1), we then get

$$\left(\left(1 + r_L^L\right) - \frac{1 + r_D}{\pi_L^L}\right) L_L = \left(\left(1 + r_L^H\right) - \frac{1 + r_D}{\pi_L^H}\right) L_L, \tag{30.20}$$

implying again the contradiction to our assumption that $\pi_L^L < \pi_L^H$ and we can rule out that $\alpha_L^H > 0$ and $\alpha_L^L > 0$.

Having ruled out these three possibilities, the only feasible solution is $\alpha_L^L = \alpha_L^H = 0$ and thus high-risk borrowers do not use mutual banks as the bank they prefer would not share any profits.

We can now shift our focus to low risk borrowers and their preferences for the optimal amount of profits that banks should share.

Low-risk borrowers Again we need assume that banks are in perfect competition and make no profits from lending to low-risk borrowers,$\Pi_B^H = 0$. Following the same steps as in deriving equation(30.10) we easily obtain

$$U^H = (1 - \rho (1 + r_D)) C_i^H + \left(p\pi_H^H + (1 - p) \pi_H^L\right) (1 + R) I, \tag{30.21}$$

which again is maximized using the highest possible value for C_1^H. We get from our assumption that $\Pi_B^H = 0$ the consumption in the first time period as

$$C_1^H = \frac{\left(p\pi_H^H + (1 - p) \pi_H^L\right) (1 + R) I - \left(p\pi_H^H C_2^{HH} + (1 - p) \pi_H^L C_2^{HL}\right)}{1 + r_D}, \tag{30.22}$$

which is the amount that the consumer needs to borrow in order to finance their consumption.

We know that $C_2^{LH} = C_2^{LL} = 0$ and if we suppose that $C_2^{HH} = C_2^{HL} = 0$, then from equations (30.5) and (30.6) we get that $C_1^H \geq C_1^L$ and $C_1^L \geq C_1^H$, requiring $C_1^H = C_1^L$. Comparing equations (30.11) and (30.22), then requires that

$$p\left(\pi_H^H - \pi_L^H\right) + (1 - p)\left(\pi_H^L - \pi_L^L\right) = 0.$$

As $\pi_H^H > \pi_L^H$ and $\pi_H^L > \pi_L^L$, this can never be fulfilled. Hence at least one of the $C_2^{Hj} > 0$. Assume now that $C_2^{HL} = 0$. Constraint (30.6) will be fulfilled with equality as the bank will extract as much surplus as possible from borrowers. Hence with our assumption that $C_2^{HL} = 0$ we obtain the consumption in the first time period as

$$\begin{aligned} C_1^L &= C_i^H + \rho\left(p\pi_L^H C_2^{HH} + (1 - p) \pi_H^L C_2^{HL}\right) \\ &= C_1^H + \rho\pi_L^H C_2^{HH}. \end{aligned} \tag{30.23}$$

Inserting from equaations (30.11) and (30.22), we get solving for C_2^{HH} that

$$C_2^{HH} = \frac{p\left(\pi_H^H - \pi_L^H\right) + (1-p)\left(\pi_H^L - \pi_L^L\right)}{\pi_H^H - \pi_L^H \rho\,(1+r_D)} > 0, \tag{30.24}$$

which is positive as $\rho\,(1+r_D) < 1$ and $\pi_H^j > \pi_L^j$.

Using equation (30.3) we then get

$$\begin{aligned} C_2^{HL} &= (1+R)\,I - \left(1 + r_L^H\right) L_H + \alpha_H^L \hat{\Pi}_B^{HL} = 0, \\ C_2^{HH} &= (1+R)\,I - \left(1 + r_L^H\right) L_H + \alpha_H^H \hat{\Pi}_B^{HH}. \end{aligned} \tag{30.25}$$

Solving the first line for $(1+R)\,I - \left(1 + r_L^H\right) L_H$, this expression becomes

$$C_2^{HH} = \alpha_H^H \hat{\Pi}_B^{HH} - \alpha_H^L \hat{\Pi}_B^{HL} > 0. \tag{30.26}$$

If we set $\alpha_H^H = \alpha_H^L = 0$, then $C_2^{HH} = 0$, in contradiction to equation (30.24). Thus at least one of the α_H^j must be positive, giving scope for the emergence of mutual banks among low-risk borrowers as the would find profit-sharing with the bank optimal.

Summary We have established that low-risk borrowers would benefit from using a mutual bank for their borrowing, while high-risk borrowers would prefer to not have a profit sharing agreement and therefore preferring conventional banks. The low risk of failure induces these borrowers to forego current consumption, and thus smaller loans, for future participation in the bank's profits and therefore higher future consumption; for high risk borrowers the possibility of not being able to repay the loan and hence having no consumption in the second time period makes such an arrangement less attractive and high-risk borrowers prefer conventional banks.

We thus have confirmed the observation that customers at mutual banks are mainly low-risk borrowers in solid financial circumstances and more risky borrowers, such as those in a more precarious financial situation as well as companies with higher risks, prefer to use conventional banks.

Reading Smith & Stutzer (1990)

30.1.2 Risk-taking by mutual banks

Mutual banks, also called cooperative banks, differ from conventional banks in that their primary objective is not to generate profits for their shareholders, but to provide benefits to their members, who are also customers of the bank. We will here investigate how the different objectives of these two types of banks affect the provision of loans, focussing on the ability of companies to obtain a loan.

The profits of a purely profit-driven bank are given by

$$\Pi_B^i = \pi \left(1 + r_L\right) L - \left(1 + r_D\right) L, \tag{30.27}$$

assuming the bank knows the probability π with which a loan L is repaid. The loan rate is set to be r_L and fully financed by deposits on which a deposit rate r_D is payable. The bank will lend as long as as it is profitable, $\Pi_B^i \geq 0$, which can be solved for

$$\pi \geq \pi_B^* = \frac{1 + r_D}{1 + r_L}. \tag{30.28}$$

A mutual bank has not only the objective of profits, but also that of providing surplus to its members, especially its borrowers. The surplus of borrowers is determined as

$$\Pi_C^i = \pi \left(\left(1 + R\right) L - \left(1 + r_L\right) L\right), \tag{30.29}$$

where R denotes the return on the investment the company conducts. Giving a weight α to the surplus of its members, and a weight $1 - \alpha$ to the profits of the mutual bank, we get the objective of the mutual to be

$$\begin{aligned} \Pi_M^i = {} & \alpha\pi \left(\left(1 + R\right) L - \left(1 + r_L\right) L\right) \\ & + \left(1 - \alpha\right) \left(\pi \left(1 + r_L\right) L - \left(1 + r_D\right) L\right). \end{aligned} \tag{30.30}$$

Lending occurs as long as it is profitable to do so, $\Pi_M^i \geq 0$, which implies that we require

$$\pi \geq \pi_M^* = \frac{\left(1 - \alpha\right)\left(1 + r_D\right)}{\alpha \left(1 + R\right) + \left(1 - 2\alpha\right)\left(1 + r_L\right)}. \tag{30.31}$$

We see that as long as the return on the successful investment, R, exceeds the loans rate, r_L, thus $R > r_L$, we have $\pi_M^* < \pi_B^*$. In other words, as long as it is profitable for the company to pursue the investment, the threshold in term of the likelihood that the loan can be repaid, is lower for mutual banks. Therefore, mutual banks will provide more risky loans than conventional banks as they will take into account the benefits that accrue to companies from such loans.

We can also investigate the impact an increase in competition has on the riskiness of loans that banks are willing to provide. If competition between banks increases, traditional economic theory suggests that loans rates will reduce and we will thus investigate this situation. We consider how the threshold for the likelihood of repaying the loan, π_B^* and π_M^*, are affected by such a change in the loan rate. For conventional banks we easily obtain that

$$\frac{\partial \pi_B^*}{\partial \left(1 + r_L\right)} = -\frac{1 + r_D}{\left(1 + r_L\right)^2} < 0 \tag{30.32}$$

and hence an increase in the loan rate, a reduction in competition, would induce an increase of the risks the bank is willing to take due to the higher profit opportunities. Thus an increase in competition leads conventional banks to provide more risky loans and thus they become more risky overall.

For mutual banks we become similarly that

$$\frac{\partial \pi_M^*}{\partial (1 + r_L)} = -\frac{(1 - 2\alpha)(1 - \alpha)(1 + r_D)}{(\alpha(1 + R) + (1 - 2\alpha)(1 + r_L))^2}. \tag{30.33}$$

This expression is positive if $1 - 2\alpha < 0$, or $\alpha > \frac{1}{2}$. Hence if the concerns for the surplus of their members dominate, mutual banks reduces the risks as the loan rate increases. In contrast to conventional banks, the risks of mutual banks would decrease as competition between banks increases. While increased competition erodes the profits of banks, it also increases the profits of the companies obtaining loans due to the lower loan rates; if the concerns of the company dominate the overall benefits to the mutual bank are increasing in competition, allowing them to reduce lending to more risky companies while maintaining their profitability.

We observe that firstly mutual banks are willing to give more risky loans than conventional profit-driven banks, but as competition between banks increases, the risks willing to be taken by mutual banks reduce while that of conventional banks increases. While mutual banks will always be take more risks than conventional banks, these differences will reduce as competition increases.

Reading Amendola, Barra, Boccia, & Papaccio (2021)

Résumé

We have seen that mutual banks are preferred by borrowers that show low risks of defaulting on their loans. Such borrowers are willing to forego consumption early on by obtaining a smaller loan in order to increase consumption at a later stage when obtaining parts of the profits the bank generated at the earlier stage. The risk of default from high risk borrowers makes consumption at a later stage less likely and they prefer to obtain a larger loan from a conventional bank for immediate consumption and accept lower future consumption. It is thus that the customers of mutual banks are mainly low-risk borrowers.

While the borrowers of mutual banks might well be of lower risk than those of conventional banks, mutual banks are generally willing to provide more risky loans if the concerns for the welfare of their members dominate over the concern of making a profit for the bank, but with competition between banks increasing, the differences in the riskiness of loans are narrowing. Thus mutual banks are willing to provide more risky loans than conventional banks in the same situation, allowing some of those to finance larger investments, for example, than they would be able to achieve when using conventional banks.

30.2 Government ownership

While most banks are owned by private individuals and banks therefore are purely profit oriented. In many countries it is common, however, for banks to be at least partly controlled by government, either through direct ownership of the capital of a banks or through other forms of control. Governments do not act purely on a commercial basis but might pursue political motives when deciding on granting loans. We will investigate how such political influence affects decision-making in banks.

Let us assume that private shareholders own a fraction α of the bank, while the government holds the remaining fraction $1 - \alpha$ of the shares and that governments may force banks to provide unproductive loans to companies that benefit the aims of their government, but that are never repaid as the recipients are not creditworthy. Such loans may be granted to companies that are struggling to avoid default but due to the loss of employment politicians may like to keep them operating.

Banks finance their loans L entirely from deposits at interest rate r_D. In order to advance loans desired by politicians, banks obtain a transfer T from the government. Such transfers might take the form of subsidies or might be non-monetary, such as being to less stringent regulation. We assume here that the transfer the bank obtains does not cover the full costs of the loan such that $T(L) \leq (1 + r_D) L$.

The net costs of such loans to the government are then given by

$$\Pi_G = T - (1 - \alpha)(T - (1 + r_D) L) = \alpha T + (1 - \alpha)(1 + r_D) L \tag{30.34}$$

The transfer is balanced against the loss of $(1 + r_D))L$ from the loans which are not repaid, including deposit interest; as the government owns a fraction $1 - \alpha$ of the bank, this share of the losses accrue to them.

From providing these loans, politicians derive benefits $B(L)$ such that if no loans are given no benefits are obtained, $B(0) = 0$, but are increasing in the size of the loan, $\frac{\partial B}{\partial L} > 0$, but the benefits are diminishing as loans become larger, $\frac{\partial^2 B}{\partial L^2} < 0$. Benefits might be higher employment in the region, with politicians benefitting through higher chances of being re-elected. There are also costs for raising the net funds Π_G, $C(\Pi_G)$, with no costs if no net transfers are made, $C(0) = 0$, which in increasing in the size of the loan needed, $\frac{\partial C}{\partial L} > 0$, and these costs are accelerating as loans become larger, $\frac{\partial^2 C}{\partial L^2} > 0$. Such costs might be the financing costs of the transfer T the financing costs of these transfers, but also the costs from delayed regeneration of the region or distorted incentives to companies. The benefits to politicians are overall given by the differences between these benefits and costs, hence

$$\Pi_P = B - C. \tag{30.35}$$

In addition to the granting of the politically motivated loans, the bank has profits Π_B^0 from other sources and hence the profits accruing to the private owners of the bank are given by

$$\Pi_B = \alpha \left(\Pi_B^0 + T - (1 + r_D) L \right) = \alpha \Pi_B^0 + \Pi_G - (1 + r_D) L, \tag{30.36}$$

acknowledging that private shareholders gain a fraction α of the total profits of the bank.

We can now determine the equilibrium amount of political loans the bank provides.

Non-cooperative equilibria Let us at first assume that the provision of the loan is controlled by the private shareholders of the bank. Hence for any given transfer T they receive, the bank would determine the optimal loan amount by maximizing profits using equation (30.36), which gives us the first order condition

$$\frac{\partial \Pi_B}{\partial L} = \frac{\partial T}{\partial L} - (1 + r_D) < 0, \tag{30.37}$$

With the transfer obtained not covering the cost of the loan, we know that $\frac{\partial T}{\partial L} < 1 + r_D$ and hence the expression above is negative. The bank would therefore choose the smallest amount of loans, $L = 0$. This in turn implies $T = 0$ and as without loans no benefits or costs accrue to politicians, they also make no gains, $\Pi_P = B - C = 0$. Hence political loans are not given and banks obtain no transfers from the government.

We easily see that the profits to the private shareholders then consist of their share of the profits the bank accumulates from their other business, thus

$$\hat{\Pi}_B = \alpha \Pi_B^0 \tag{30.38}$$

If politicians determine the provision of such loans rather than the private owners of the bank, they seek to maximize their profits Π_P, but need to consider that banks need to be profitable in order to remain active in the markets, $\Pi_B \geq 0$. Politicians will seek to extract the all surplus from banks, hence we will find that $\Pi_B = 0$, which implies from equation (30.36) that the net surplus to government is given by

$$\Pi_G = (1 + r_D) L - \alpha \Pi_B^0. \tag{30.39}$$

Seeking to maximize their own benefits, politicians will solve the first order condition

$$\frac{\partial \Pi_P}{\partial L} = \frac{\partial B}{\partial L} - \frac{\partial C}{\partial \Pi_G} \frac{\partial \Pi_G}{\partial L} = \frac{\partial B}{\partial L} - (1 + r_D) \frac{\partial C}{\partial \Pi_G} = 0. \tag{30.40}$$

Denoting the solution to equation (30.40) by B^* and C^*, we get the profits of the bank and the politicians as

$$\begin{aligned} \hat{\hat{\Pi}}_B &= 0, \\ \hat{\hat{\Pi}}_P &= B^* - C^*. \end{aligned} \tag{30.41}$$

Rather seeking an antagonistic approach between the private bank owners and politicians, we will now explore the optimal solution if banks and politicians cooperate by negotiating the loan amount the bank provides.

Nash bargaining without bribes We consider the bank and politicians bargaining about the loan amount L and the next costs Π_G. First we consider the case that if the negotiation breaks down, the private bank owners decide the level of loans. Thus the outside option of for banks are given by equation (30.38) and for politicians this would be $\hat{\Pi}_P = 0$.

The objective function is when using Nash bargaining is then given by

$$\mathcal{L} = \left(\Pi_P - \hat{\Pi}_P\right)\left(\Pi_B - \hat{\Pi}_B\right) = (B - C)\left(\Pi_G - (1 + r_D) L\right) L, \tag{30.42}$$

such that the first order conditions become

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial L} &= \frac{\partial B}{\partial L}\left(\Pi_G - (1 + r_D) L\right) - (B - C)(1 + r_D) = 0, \\
\frac{\partial \mathcal{L}}{\partial \Pi_G} &= -\frac{\partial C}{\partial \Pi_G}\left(\Pi_G - (1 + r_D) L\right) + (B - C) = 0.
\end{aligned} \tag{30.43}$$

We note that the equilibrium does not depend on the private ownership of the bank, α, but only on the costs and benefits of the politicians. We can simplify equation (30.43) to

$$\frac{\partial B}{\partial L} - (1 + r_D)\frac{\partial C}{\partial \Pi_G} = 0, \tag{30.44}$$

the same condition as in equation (30.40) for the optimum where politicians decide on the amount of lending. Thus the lending in this case will be optimal for politicians.

In the case that politicians can determine the amount of the political loan if negotiations break down, the outside options are given by equation (30.41) and hence the objective function of the Nash bargaining solution becomes

$$\begin{aligned}
\mathcal{L} &= \left(\Pi_P - \hat{\hat{\Pi}}_P\right)\left(\Pi_B - \hat{\hat{\Pi}}_B\right) \\
&= \left((B - C) - (B^* - C^*)\right)\left(\alpha \Pi_B^0 + \Pi_G - (1 + r_D) L\right).
\end{aligned} \tag{30.45}$$

The first order conditions are then given by

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial L} &= \frac{\partial B}{\partial L}\left(\alpha \Pi_B^0 + \Pi_G - (1 + r_D) L\right) \\
&\quad - \left((B - C) - (B^* - C^*)\right)(1 + r_D) = 0, \\
\frac{\partial \mathcal{L}}{\partial \Pi_G} &= -\frac{\partial C}{\partial \Pi_G}\left(\alpha \Pi_G^0 + \Pi_G - (1 + r_D) L\right) \\
&\quad + \left((B - C) - (B^* - C^*)\right) = 0.
\end{aligned} \tag{30.46}$$

The results here depend on the fraction α that private owners hold in the bank. If this private ownership increases, the in the first expression the first term increases, hence we need the loan amount L to increase as that reduces the first term in brackets and $\frac{\partial B}{\partial L}$ will also decrease as $\frac{\partial^2 B}{\partial L^2} < 0$. From the second line, we see that the net government benefits Π_G need to decrease as then the more negative first term becomes smaller, also as $\frac{\partial^2 C}{\partial \Pi_G^2} > 0$. Hence an increase in private ownership will decrease the transfer government net benefits Π_G and increase the loan provided. This is because the politicians threaten to confiscate a larger fraction of the bank, giving them a stronger bargaining position, which will result in them extracting more surplus in through a larger loan.

Combining the equations from equations (30.46) we again obtain

$$\frac{\partial B}{\partial L} = (1 + r_D) \frac{\partial C}{\partial \Pi_G}, \tag{30.47}$$

implying the loan amount provided is as optimal as it would be if politicians would decide the loan amount directly.

Nash bargaining with bribes Rather than politicians making transfers to the bank through government resources, politicians could resort to paying private shareholders directly to provide incentives for the granting of such political loans. Alternatively, the private shareholders could provide funds to the politicians directly with the aim of reducing the amount of loans they provide. Such payments are made directly between politicians and private shareholders and bypass government and bank funds; we will therefore call such payments 'bribes', b.

Let us assume that a bribe b can be paid as part of the negotiation with $b > 0$ corresponding to private shareholders paying politicians and $b < 0$ to politicians paying private shareholders.

Adding this term to the profits of politicians and private shareholders, Π_P and Π_B, we then get, the equivalent of the Nash bargaining objective function in equation (30.42) as

$$\mathcal{L} = (B - C + b) (\Pi_G - (1 + r_D) L - b), \tag{30.48}$$

where assumed that the bank retains control of the lending decision if negotiations break down. The first order conditions of the Nash bargaining solution then become

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial L} &= \frac{\partial B}{\partial L} (\Pi_G - (1 + r_D) L - b) - (1 + r_D) (B - C + b) \\
&= 0, \\
\frac{\partial \mathcal{L}}{\partial \Pi_G} &= -\frac{\partial C}{\partial \Pi_G} (\Pi_G - (1 + r_D) L - b) + (B - C - b), \\
\frac{\partial \mathcal{L}}{\partial b} &= (\Pi_G - (1 + r_D) L - b) - (B - C + b) = 0,
\end{aligned} \tag{30.49}$$

which solve for

$$\begin{aligned}
\frac{\partial B}{\partial L} &= 1 + r_D, \\
\frac{\partial C}{\partial \Pi_G} &= 1, \\
b^* &= \frac{1}{2} (\Pi_G - (1 + r_D) L - (B - C)).
\end{aligned} \tag{30.50}$$

Firstly we note that all results are independent of the ownership structure and we again find $\frac{\partial B}{\partial L} = (1 + r_D) \frac{\partial C}{\partial \Pi_G}$ and hence the loan amount is optimal for politicians. If we insert for the net government benefits Π_G from equation (30.34), the solution for the optimal bribe becomes $b^* = \frac{1}{2}\left(\alpha\left(T-\left(1+r_D\right)L\right)-\left(B-C\right)\right)$. As we assumed that the transfers are less than the costs of the loan, thus $T < \left(1+r_D\right)L$, the first term is negative and if the direct benefits to politicians are exceeding their costs, $B > C$, we have that $b^* < 0$ and politicians would pay private shareholders to provide such loans. The size of the bribe is independent of the ownership structure of the bank.

If, on the other hand, politicians determine the loan amount if negotiations break down, we have the equivalent of equation (30.45) for the Nash bargaining given as

$$\mathcal{L} = \left(\left(B-C+b\right)-\left(B^*-C^*\right)\right)\left(\alpha\Pi_B^0+\Pi_G-\left(1+r_D\right)L-b\right) \tag{30.51}$$

and hence the first order conditions are given by

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial L} &= \frac{\partial B}{\partial L}\left(\alpha\Pi_B^0+\Pi_G-\left(1+r_D\right)L-b\right) \\
&\quad -\left(1+r_D\right)\left(\left(B-C\right)+b-\left(B^*-C^*\right)\right)=0, \\
\frac{\partial\mathcal{L}}{\partial\Pi_G} &= -\frac{\partial C}{\partial\Pi_G}\left(\alpha\Pi_B^0+\Pi_G-\left(1+r_D\right)L+b\right) \\
&\quad +\left(\left(B-C\right)+b-\left(B^*-C^*\right)\right)=0, \\
\frac{\mathcal{L}}{\partial b} &= \left(\alpha\Pi_B^0+\Pi_G-\left(1+r_D\right)L-b\right) \\
&\quad -\left(\left(B-C\right)+b-\left(B^*-C^*\right)\right)=0,
\end{aligned} \tag{30.52}$$

which solves for

$$\begin{aligned}
\frac{\partial B}{\partial L} &= 1+r_D, \\
\frac{\partial C}{\partial\Pi_G} &= 1, \\
b^{**} &= \frac{1}{2}\left(\alpha\Pi_B^0+\Pi_G-\left(1+r_D\right)L-\left(B-C\right)\right. \\
&\qquad \left.+\left(B^*-C^*\right)\right) \\
&= b^*+\frac{1}{2}\left(\alpha\Pi_B^0+\left(B^*-C^*\right)\right) > b^*.
\end{aligned} \tag{30.53}$$

We firstly observe that as $B^* \geq C^*$, given $\Pi_P \geq 0$, the bribe is higher if politicians can take control, while the resulting amount of loans is again optimal for politicians. Secondly, the size of the bribe is now affected by the ownership structure of the bank as α is included in this solution; the bribe is increasing in the private ownership of the bank. This is because the private owners of the bank have a large exposure to being expropriated by politicians and seek to limit this through bribes. If we assume that the bank is sufficiently profitable in their other business areas, thus Π_B^0 is large, then the expression for the bribe b^{**} will become positive and the private owners of the bank will pay politicians to reduce the amount of loans they have to provide.

If the bribe is positive, thus private bank owners pay politicians, the first condition in equation (30.52) shows that the loan amount is reduced compared to the case where no bribes are paid. This is the result of the bribes paid, which compensates politicians for the lower loan amount.

Summary We have thus seen that if banks are partially owned by the government, they might be coerced into providing loans that are politically expedient but economically not viable and the higher the government ownership is, the larger such loans become. The benefits to the politicians are reduced by the costs imposed on the bank for such loans, which is partially borne by the government through lower profits on their share of the bank profits. With private bank owners seeking to avoid the provision of such loans, banks are vulnerable to corruption, the most importantly if politicians have the power to enforce the provision of political loans; private bank owners may make payments to politicians to reduce the size of these loans.

Political influence on banks could, of course, also happen if the bank is not government owned, but with government ownership the influence will be larger; for example government representatives are likely to be on the supervisory board of the bank and the executive board may even have members that are inclined to meet the demands of governments and politicians. This makes the scenario where the breakdown of negotiations allows politicians to determine the loan amount more realistic, especially if combined with the ability of politicians to influence regulators or regulatory requirements.

Reading Shleifer & Vishny (1994)

Conclusions

Chapter 31

Price complexity

It is common to focus on the price of a product, such as the loan rate, deposit rate, or fees, and make the best choice by comparing these conditions. At times such prices need to be adjusted for a small number of other factors, such as the risk of deposits not being repaid or the requirement to provide collateral for a loan. However, banks often make contracts much more complex through a number of terms and conditions. There might be differences between banks offering services associated with deposit accounts; there might fees to be paid for the account, which however might be reduced or even waived under certain conditions, but on the other hand penalties for breaching even minor conditions, such as exceeding an agreed credit limit might be imposed, while other banks might be more lenient. Loan contract might also vary considerably with penalties for making payments late, but also the ability to repay loans early and in some cases withdraw any overpayment again in the future. Having to take into account all these different aspect of a contract makes it difficult to compare the prices directly and choose the bank offering the best value. We will here see that having such complex terms and conditions might be a deliberate strategy of banks to reduce competition between them.

We consider N banks that do not only compete on the price of a product or service, but we also take into account the complexity of providing information about the price and value of the product due to complex terms and conditions. Banks can, at no cost, assign a complexity $0 \leq \xi \leq 1$ to their products such that only a fraction ξ customers can asses them properly and are thus informed; the remaining customers are not able to assess the price the bank charges. It is thus that the more complex a bank makes its products the fewer customers are able to select the best bank for their needs; other customers will have to make uninformed decisions.

Normal price competition with customers being informed about the value of the product or service would imply that the price of the service is reduced to its marginal costs. We here assume that a bank charging the lowest price commands the entire market of the ξ informed customers, or if M multiple banks charge the same lowest price, a fraction of $\frac{1}{M}$ of the customers. In addition, the fraction of uninformed customers, $1 - \xi$, are equally shared between all N banks. Hence charging marginal costs would lead to zero profits, but when increasing the price, the bank would make a profit from those customers that are uninformed and remain with the bank. Therefore the price may not be set uniformly at marginal costs for all banks, but will be above such costs.

Such a price set uniformly above the marginal costs would, however, allow banks to undercut each other and gain the market of the informed customers fully, increasing their own profits. Thus this cannot be an equilibrium either. The only possible equilibrium is one in which the price is set as a mixed strategy with some distribution $F(P)$, for $\underline{P} \leq P_i \leq \overline{P}$, with $\underline{P}$ denoting the marginal costs of the bank and $\overline{P}$ the value of the service to the customer; P will denote the price charged.

The bank will gain the market for the informed customers if its price is lowest, that is all other banks have higher prices. This happens with probability $1-F(P_i)$ for each bank, and thus for each of the other $N-1$ banks with probability $(1-F(P))^{N-1}$. Neglecting the costs of the bank for convenience and thus marginal costs of zero, we have bank profits given as

$$\Pi_B = P\left(\xi(1-F(P))^{N-1} + (1-\xi)\frac{1}{N}\right), \tag{31.1}$$

where the final term denotes the market share of the uninformed customers. The optimal complexity is then given by maximizing the bank profits, which gives the first order condition as

$$\frac{\partial \Pi_B^i}{\partial \xi_i} = P\left((1-F(P))^{N-1} - \frac{1}{N}\right). \tag{31.2}$$

If $(1-F(P_i))^{N-1} - \frac{1}{N} > 0$, the derivative is positive and hence $\xi_i = 0$ is chosen; banks choose the highest possible complexity for their products and services as then no customer can assess them correctly. This condition can be solved for

$$P < P^* = F^{-1}\left(1 - \frac{1}{N^{\frac{1}{N-1}}}\right). \tag{31.3}$$

Similarly, if $P > P^*$, then $\xi = 1$ as the derivative in equation (31.2) would be negative and hence the lowest possible complexity $xi = 1$ would be chosen; for $P = P^*$ any complexity could be chosen as the expression is zero for any value of ξ. Thus we find that the optimal complexity is given by

$$\xi = \begin{cases} 0 & \text{if } P_i < P_i^* \\ [0;1] & \text{if } P_i = P_i^* \\ 1 & \text{if } P_i > P_i^* \end{cases}. \tag{31.4}$$

We thus see that for products and services that attract low prices, and hence low profits margins, banks will optimally choose a highly complex pricing structure that is very intransparent with the aim of reducing competition and thereby retaining their market share despite charging a high price for the product.

We can now show that

$$\frac{\partial P^*}{\partial N} = \frac{N(\ln N - 1) + 1}{f(P^*)(N-1)^2 N^{\frac{1}{N-1}}} > 0. \tag{31.5}$$

Thus as the number of banks increases, the threshold P_i^* increases and therefore $F(P_i^*)$ increases, implying that the likelihood of highly complex pricing structures is increasing as competition between banks increases and will thus reduce the profit margin if competing for customers. For perfect competition, $N \to \infty$, P^* grows beyond bounds and thus complexity ia always maximal.

We see that banks offer complex pricing structures for products and services with low profit margins in an attempt to limit competition by denying a customers the ability to compare the offerings and thereby maintain the profitability of such products and services. This strategy allows banks to retain customers as no price competition between banks emerges and thus customers select banks ar randomly; for products with a higher product margin, banks will seek to compete by offering clear and easy-to-understand pricing policies with the aim to undercut their competitors through the well-understood prices and gain market share from their competitors.

Reading Carlin (2009)

Review

The last few chapters have shown that behind the often simplifying assumptions of banks having a given balance sheet structure, employees making decisions in the interest of the bank, and banks competing on the price of their products and services are either not fulfilled or the result of an elaborate system of 'checks and balances' that provide employees with incentives through their remuneration to make decisions that are maximizing the profits of the bank they are working for. Banks themselves might also make decisions that limit price competition through intransparent and complex pricing strategies, or engage in activities that they know are detrimental to their customers or society, but beneficial to them.

Banks have developed systems that seeks to align the incentives of their employees and the owners of the bank, which often feature the use of bonus payments for successful employees or the reduction of payment for employees leading to undesired outcomes to the bank. These results show how the decisions in banks are formed, and distorted, but often conflicting interests, showing that the objectives of decision-making in banks are much less clear than most of the models employed would suggest. While the arrangements banks have made will often align the interests of banks and their employees, any decision taken may be distorted by this not being achieved fully. It is therefore that comparing the results of models as presented here with actual decisions made in banks will be made more difficult.

It is not only that employees are rewarded such that their incentives align with that of the bank they work for, but it also affects the bank an employee chooses to work for. We have seen in several model that the best-performing banks employ the best-performing employees and it is tempting to attribute this fact to highly-performing employees generating a high performance for the bank that employs them. Thus the implication is that banks are high-performing because they employ high-performing employees. However, the causality is reversed; high-performing banks attract high-performing employees as the abnk is able to pay them a higher remuneration due to the large profits they generate. Of course, employing highly-performing individuals will then reinforce that the bank is highly-performing.

We have also seen that is matters who owns a bank as the objective the bank itself pursues might change and with mutual banks the interest of their customer should be taken into account, while for privately owned banks only the interests of the owners should matter. We have seen how with such a different ownership structure, decisions on the provision of loans might be different in some instances.

Political influence in (partly) state-owned banks can lead to the provision of loans that a privately owned banks would never agree to, opening the banking sector up to corruption in order to prevent banks having to give such loans. But corruption can also occur if employees are given bribes such that loans to uncreditworthy borrowers are provided. We have seen how banks, and their employees, can be susceptible to morally questionable or even illegal behaviour. They might knowingly sell products to customers that are unsuitable to their needs as long as this is not detected too easily, they might not pursue money laundering as vigorously as would be desirable and need to be incentivised to cooperate with regulators through fines.

With banks perceiving risks from lending to be low, they will pursue a strategy to grow at the expense of providing enhanced services to their existing customers, including reducing risk to these customers as well as themselves, while those perceiving risks to be high will be more cautious about growing their business. It will thus be that banks who underestimate risks are increasing their balance sheet and accumulate additional risks, while banks who overestimate risks will seek to reduce their risk exposure by taking measures on their existing risks. this has implications for the riskiness of the banking system as a whole as it is banks that believe to take low risks and therefore can afford to expand, will find themselves facing higher risks than expected, causing the banking system to be potentially unstable.

Part VI
Systemic risk

Thus far we have mainly considered the decisions of a single bank and what implications such decisions had on the bank itself as well as its customers, borrowers as well as depositors; we included the effect on other banks only as far as they were competitors for borrowers, depositors or employees. However, banks are closely interconnected through interbank loans, providing loans to similar borrowers, and the value of their loans may depend on their ability to sell these in the market, where other bank might sell similar loans. Through such connections between banks, the failure of one bank, even though specific to the circumstances of that bank, may induce the failure of other banks. The possibility of the failure of one bank causing the failure of other banks is referred to as systemic risk. It is not a systemic risk if several banks fail due to the same reason, for example a higher than expected default rate on loans, the failure of banks can be attributed to different, unconnected reasons. In order to have a systemic banking crisis it must be that the failure of one, or a small number of banks, causes the failure of other banks or the failure of different banks is connected by a decision of banks to be exposed to the same risk as other banks.

In chapter 32 we will explore the different ways a bank failure can spread, which is also referred to as contagion, and thereby trigger a systemic banking crisis. If faced with a systemic banking crisis, any government or regulator needs to consider whether the affected banks should be saved through the injection of capital or liquidity; such bail-outs are discussed in chapter 33. This decision is different from the decision to prevent the failure of a single bank as the resources required will be significantly higher if multiple banks fail. However, such bail-outs, if instigated at an early stage of a systemic banking crisis, might also help to prevent failures to spread further and thus prevent or limit a systemic banking crisis.

Chapter 32

Contagion mechanism

The losses of one bank can spread to other banks mainly through three different mechanisms; if such losses are sufficiently large, the banks will fail. Firstly, banks connected to each other through interbank loans, where banks provide each other with loans as discussed in chapter 16. If a bank is not able to repay its interbank loan, this will impose a loss on the lender, which in turn might fail as the losses exceed its equity. We might also have a situation where a bank withdraws its interbank loan, which can in most cases be done without having to give notice, for example by not extending an existing loan beyond the agreed maturity date. This withdrawal of an interbank loan will affect the borrowing bank in that it will have to repay the interbank loan and thus will have to either hold sufficient cash reserves, or it will have to raise such cash reserves if the existing cash reserves are not sufficient. A bank which is not able to raise enough cash reserves will fail due to illiquidity. In chapter 32.1 we will investigate how the failure of one bank can spread to other banks through such interbank loans.

Another way the failure of one bank is an indication that other banks might also be subject to losses and potential failures, is the case that banks have a common exposure to risks as discussed in chapter 32.2. If banks provide loans and invest other assets whose losses are highly correlated across banks, then the failure of one bank can lead to a reasonable conclusion that another bank will have suffered similar losses. We will look at the incentives of banks to choose loans and assets that have a high correlation and how information on the losses of one bank can cause another bank to fail.

Finally, if a bank needs to raise cash reserves as interbank loans or deposits are withdrawn, it will have to sell assets they are holding, for example loans they have provided. The same would be observed if a bank has experienced losses and has failed such that their assets are now liquidated. The sale of these assets, often under pressure for a quick realisation of the proceeds, will increase the supply of such assets, while at the same time there are less banks willing to purchase additional assets. With arguments of demand and supply, this would reduce the price assets can be sold for; with market prices being the basis of the valuation of assets, other banks holding these and similar assets will experience a loss as their value has reduced. Such a loss might now cause these banks to fail, further increasing the sale of assets and reducing market prices further. We will investigate in chapter 32.3 how such fire sales of assets caused by the failure of one bank can spread and cause other banks to fail.

32.1 Interbank markets

In chapter 16 we have discussed that banks may lend and borrow from each other through what is known as interbank loans. Such borrowing and lending gives rise to interconnections between banks that can lead to the losses of one banks spreading to other banks as interbank loans are not repaid. Similarly, banks not renewing their interbank lending can lead to a liquidity shortage at other banks, which can also spread as these banks cannot extend the interbank loans they have provided. Thus, interbank loans have the potential to give rise to systemic risk. We will look here at the way such systemic risk can emerge from the interbank market.

In chapter 32.1.1 we will investigate how the interbank market may cease to exists, or freeze, as banks are no longer willing to provide interbank loans and thus cause liquidity shortages among banks. While the interbank market might not fully freeze, it might transmit any shocks that one bank has experienced, for example the unexpected withdrawal of deposits that have been transferred to another bank, and while sufficient liquidity is available in the market, the interbank loans might not be effective in directing this existing liquidity towards banks that are short of cash reserves as we will see in chapter 32.1.2.

We will also have to consider how a liquidity shock affecting one bank, such as a bank run on that bank, will spread to other banks as interbank loans will be withdrawn; chapter 32.1.3 will discuss this issue which can give rise to systemic risk. Finally, we consider in chapter 32.1.4 that banks may fail and are therefore not able to repay their interbank loans; this is also known as counterparty risk. Such a failure by one bank will impose losses on other banks and we will investigate how banks are managing such risks.

32.1.1 Contagion through fundamental shocks

Banks will face losses if loans are not repaid as expected, this might especially be the case if bank face an economic downturn with higher default rates on loans. We refer to this as a fundamental shock to a bank. Depositors observing such losses by banks may withdraw their funds, causing a bank run. If other bank have provided interbank loans to a bank affected by such a bank run, they will also suffer losses and this in turn could cause the lending bank to experience a bank run, even if unaffected by higher default rates itself. We will here investigate under which condition such bank runs can transmit through interbank loans from one bank to another bank.

Let us consider an economy with two banks, each taking deposits D from the general public. Banks use their deposits to provide loans L_i to companies, charging interest r_L and these loans being repaid with probability π. While deposits can be withdrawn after a single time period, loans are maturing only after two time periods. Instead of recalling loans, bank can sell some of these loans, $\hat{L}$, at a fraction $\lambda < 1$ of the final expected value $\pi(1 + r_L)\hat{L}$, where $\hat{L} \leq L_i$.

We furthermore assume that bank 1 uses an amount M of their deposits to provide an interbank loan to bank 2, who uses these proceeds to provide additional loans. Hence the amount of loans the two banks provide is given by

$$\begin{aligned} L_1 &= D - M, \\ L_2 &= D + M. \end{aligned} \tag{32.1}$$

With r_M denoting the interest on interbank loans and r_D on deposits, we have the profits of the two banks given by

$$\begin{aligned} \Pi_B^1 &= \pi(1 + r_L)L_1 + (1 + r_M)M - (1 + r_D)D \\ &= (\pi(1 + r_L) - (1 + r_D))D + ((1 + r_M) - \pi(1 + r_L))M, \\ \Pi_B^2 &= \pi - (1 + r_L)L_2 - (1 + r_M)M - (1 + r_D)D \\ &= (\pi(1 + r_L) - (1 + r_D))D + (\pi(1 + r_L) - (1 + r_M))M. \end{aligned} \tag{32.2}$$

We assume that $\Pi_B^i > 0$ and hence deposits and interbank loans are risk free. This implies that for both banks deposits and interbank loans are perfect substitutes, implying in turn that the respective interest rates have to be equal, $r_D = r_M$, simplifying equation (32.2) to become

$$\Pi_B^i = (\pi(1 + r_L) - (1 + r_D))L_i. \tag{32.3}$$

Suppose now that bank 2, the bank obtaining the interbank loan from bank 1, faces an unanticipated shock in time period 1 that causes the probability of loans being repaid to fall from π to $\hat{\pi} < \pi$. At this point, due to the information on this shock, depositors can withdraw their funds and a bank run may emerge; similarly, the lending bank can recall its interbank loans and we assume that they withdraw a fraction $0 \leq \gamma \leq 1$.

We can now assess the position of the borrowing bank, assuming initially that the decision to withdraw interbank loans is given. We will later consider the optimal decision whether to withdraw interbank loans by the lending bank.

The borrowing bank Let us for now assume that the lending bank's decision to recall a fraction γ of the interbank loan is given, hence the bank requires funds of γM to meet these recalls, assuming that in recalled interbank loans no interest is payable. These funds need to be raised by selling loans obtaining $\lambda\hat{\pi}(1 + r_L)\hat{L}$. Setting these two expressions equal, we get the amount of loans that need to be liquidated as

$$\hat{L} = \frac{\gamma}{\lambda\hat{\pi}(1 + r_L)}M \tag{32.4}$$

The profits of the borrowing bank after two time periods are then given by

$$\Pi_B^2 = \hat{\pi}\left(1 + r_L\right)\left(L_2 - \hat{L}\right) - \left(1 + r_D\right) D - \left(1 - \gamma\right)\left(1 + r_D\right) M \tag{32.5}$$

where the first term denotes the proceeds of the non-recalled loans on which repayments are received, the second the repayment of deposits and the final term the non-recalled interbank loans. We easily see that from equation (32.4) we have

$$\frac{\partial \hat{L}}{\partial \gamma} = \frac{M}{\lambda \hat{p} i \left(1 + r_L\right)} \tag{32.6}$$

and from equation (32.5), using this result, that

$$\frac{\partial \Pi_B^2}{\partial \gamma} = -\hat{\pi}\left(1 + r_L\right)\frac{\partial \hat{L}}{\partial \gamma} + \left(1 + r_D\right) M = \left(1 + r_D - \frac{1}{\lambda}\right) M < 0, \tag{32.7}$$

with the inequality arising if we assume λ to be sufficiently small. The relationship in equation (32.4) can only hold if $\hat{L} \leq L_2$, otherwise not all interbank loans can be repaid, but only the amount of $\lambda\hat{\pi}\left(1 + r_L\right) L_2$. In this case

$$\Pi_B^2 = -\left(1 + r_D\right) D - \left(1 + r_D\right)\left(M - \lambda\hat{\pi}\left(1 + r_L\right) L_2\right) < 0; \tag{32.8}$$

this is negative as no loans are retained and thus no revenue obtained from them and we have $\lambda\hat{\pi}\left(1 + r_L\right) L_2 < M$ by assumption.

Let us now assume that if no interbank loans are recalled, $\gamma = 0$, we have the borrowing bank's profits in time period 2 given as

$$\begin{aligned} \Pi_B^2 &= \hat{\pi}\left(1 + r_L\right) L_2 - \left(1 + r_D\right)\left(D_M\right) \\ &= \left(\hat{\pi}\left(1 + r_L\right) - \left(1 + r_D\right)\right) L_2 > 0 \end{aligned} \tag{32.9}$$

and if all interbank loans are recalled, $\gamma = 1$, that with equation (32.1) these profits become

$$\begin{aligned} \Pi_B^2 &= \hat{\pi}\left(1 + r_L\right)\left(L_2 - \frac{M}{\lambda\hat{\pi}\left(1 + r_L\right)}\right) - \left(1 + r_D\right) D \\ &= \left(\hat{\pi}\left(1 + r_L\right) - \left(1 + r_D\right)\right) D + \left(1 + r_D - \frac{1}{\lambda}\right) M < 0. \end{aligned} \tag{32.10}$$

We note that the second term is negative with the assumption in equation (32.7) and hence profuts over all will be negative.

If $\Pi_B^2 < 0$, then depositors cannot be fully paid and therefore they would cause a bank run. Due to equation (32.7) showing that the bank's profits are reducing with the fraction of interbank loans withdrawn, γ, and we found that the profits at $\gamma = 0$ and $\gamma = 1$ have different signs, there will be an $\underline{\gamma}$ such that $\Pi_B^2 = 0$. If $\gamma < \underline{\gamma}$ the borrowing bank is solvent and hence no bank run should be observed, while for $\gamma > \underline{\gamma}$ a bank run will occur.

Having investigated the implications of the withdrawal of interbank loans, we can now continue to investigate the optimal withdrawal of interbank loans by the lending bank.

The lending bank We now consider the optimal recall of interbank loans. The most the borrowing bank can repay is $\lambda\hat{\pi}\left(1+r_L\right)L_2$ by selling all loans. These funds are then paid to the lending bank and the depositors. The funds required for this are $\left(1+r_D\right)D$ to depositors and $\gamma\left(1+r_D\right)M$ for the called interbank loans, noting our result that $r_M=r_D$. Setting $\lambda\hat{\pi}\left(1+r_L\right)L_2=\overline{\gamma}\left(1+r_D\right)M+\left(1+r_D\right)D$, with $\overline{\gamma}$ denoting the highest interbank withdrawal that would allow interbank loans to be repaid fully, we get with equation (32.1) that

$$\overline{\gamma}=\frac{\left(\lambda\hat{\pi}\left(1+r_L\right)-\left(1+r_D\right)\right)D+\hat{\pi}\left(1+r_L\right)M}{\left(1+r_D\right)M}. \tag{32.11}$$

If $\gamma<\overline{\gamma}$, the borrowing bank can repay all interbank loans recalled, and for $\gamma>\overline{\gamma}$, this is not possible.

In the case that $\gamma\leq\underline{\gamma}$, the borrowing bank does not face a bank run and can meet its obligations in full, hence its profits are given by

$$\begin{aligned}\Pi_B^1&=\pi\left(1+r_L\right)L_1-\left(1+r_D\right)D-\gamma M-\left(1-\gamma\right)\left(1+r_D\right)M \quad (32.12)\\&=\left(\pi\left(1+r_L\right)-\left(1+r_D\right)\right)D\\&=\left(\pi\left(1+r_L\right)-\gamma-\left(1-\gamma\right)\left(1+r_D\right)\right)M,\end{aligned}$$

where we note that the lending bank does not face a shock to the default rate π, this is only affecting the bank borrowing in the interbank market; furthermore, recalled interbank loans do not attract interest as before.

For $\underline{\gamma}<\gamma\leq\overline{\gamma}$ the recalled loans can be fully repaid, but as the borrowing bank faces a bank run, those interbank loans not recalled will only be repaid partially. The amount recalled needs to cover the interbank loans and the deposits that are withdrawn in the bank run, hence $\lambda\hat{\pi}\left(1+r_L\right)\hat{L}=\gamma\left(1+r_D\right)M+\left(1+r_D\right)D$ as in the derivation of (32.11), hence the amount of loans that need to be repaid are given by

$$\hat{L}=\frac{1+r_D}{\lambda\hat{\pi}\left(1+r_L\right)}\left(\gamma M+D\right). \tag{32.13}$$

The lending bank will seize the remaining assets of the borrowing bank to recover parts of their interbank loan, the amount that can be recovered being $\hat{\pi}\left(1+r_L\right)\left(L_2-\hat{L}\right)$. Hence their profits will be

$$\Pi_B^1=\pi\left(1+r_L\right)L_1-\left(1+r_D\right)D+\gamma M+\hat{\pi}\left(1+r_L\right)\left(L_2-\hat{L}\right). \tag{32.14}$$

With $\frac{\partial\hat{L}}{\partial\gamma}=\frac{1+r_D}{\lambda\hat{\pi}(1+r_L)}M$ from equation (32.13), we get

$$\frac{\partial\Pi_B^1}{\partial\gamma}=M-\hat{\pi}\left(1+r_L\right)\frac{\partial\hat{L}}{\partial\gamma}=\left(1-\frac{1+r_D}{\lambda}\right)M<0. \tag{32.15}$$

From equation (32.12) we also get $\frac{\partial \Pi_B^1}{\partial \gamma} < 0$ and we can show that at $\gamma = \underline{\gamma}$ the expression in equation (32.14) is less than in equation (32.12).

If $\gamma > \overline{\gamma}$, not all recalled interbank loans can be repaid and those not recalled will not be paid at all. The lending bank is repaid its share of the deposits and recalled loan that are due, assuming they are of equal seniority, to be paid from the proceeds of selling all loans. Hence

$$\Pi_B^1 = \pi (1 + r_L) L_1 - (1 + r_D) + \frac{\gamma M}{\gamma M + D} \lambda \hat{\pi} (1 + r_L) L_2. \tag{32.16}$$

We then get that

$$\frac{\partial \Pi_B^1}{\partial \gamma} = \frac{MD}{(\gamma M + D)^2} \lambda \hat{\pi} (1 + r_L) L_2 > 0. \tag{32.17}$$

We can summarize the relationship between the amount of interbank loans being recalled, γ, and the profits of the lending bank in figure 32.1. We see that the only possible optimum is either $\gamma = 0$ or $\gamma = 1$, depending on the value of the other parameters. The next step now is to assess the conditions under which banks will be subjected to a bank run.

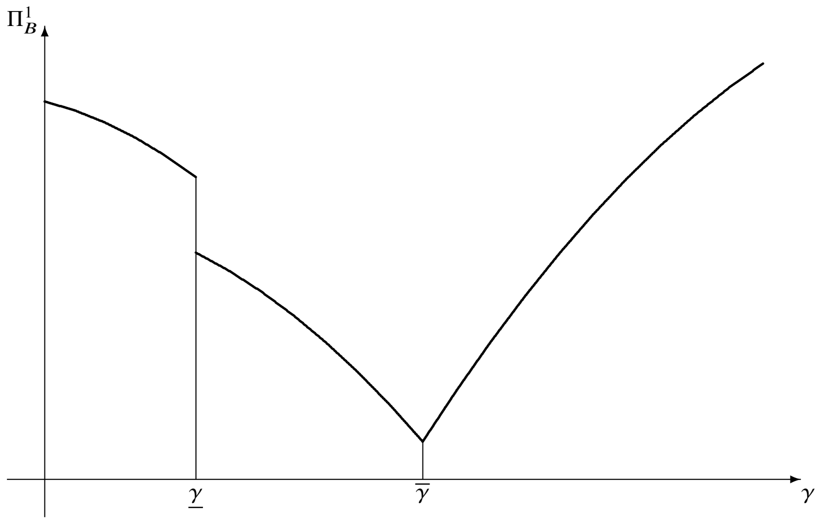


Fig. 32.1: Interbank loan recalls and profits of the lending bank

Inducing bank runs in the borrowing bank We first consider the possibility of a bk run int he bank borrowing from the interbank market. If depositors will not be able to obtain their initial investment, D, from the bank after the second time period, they will seek to withdraw their funds early and thus cause a bank run. Thus a bank run can be avoided if the resources they have from the repayment of loans at the lower repayment rate $\hat{\pi}$, depositors will have no reason to withdraw their funds, thus we require $D < \hat{\pi}(1 + r_L) L_2$ such that no bank run will occur. The lending bank also has no reason to call in loans in this situation and hence $\gamma = 0$, as it will also be repaid. Hence if

$$\hat{\pi} > \pi^* = \frac{1}{1 + r_L} \frac{D}{D + M} \tag{32.18}$$

no bank run on the borrowing bank occurs, while a lower $\hat{\pi}$ would cause such a bank run. Thus if the risk the borrowing bank faces increases sufficiently, a bank run will be observed. We can now continue to analyse the situation in which the borrowing bank is subject to a bank run further by looking at the incentives of the lending bank to withdraw interbank loans and whether this bank will also be subjected to a bank run.

Interbank loan recalls If at $\hat{\pi} < \pi^*$ the borrowing bank faces a bank run, the lending bank will have to decide whether to recall their interbank loans, $\gamma = 1$, or retain interbank loans with the bank, $\gamma = 0$. In the case that all interbank loans being retained, $\gamma = 0$, the lending bank obtains all of the proceeds from the loans not sold to meet the requirements of the depositors. Of course the borrowing bank cannot sell more loan than they are actually holding, hence the profits if the lending bank are given by

$$\Pi_B^1 = \pi(1 + r_L) L_1 - (1 + r_D) D + \max\left\{0; \hat{\pi}(1 + r_L)\left(L_2 - \hat{L}\right)\right\}. \tag{32.19}$$

Using equation (32.13) with $\gamma = 0$, we obtain that $\frac{\partial \hat{L}}{\partial \hat{\pi}} = -\frac{\hat{L}}{\hat{\pi}}$, such that we find that

$$\frac{\partial \Pi_B^1}{\partial \hat{\pi}} = (1 + r_L)\left(L_2 - \hat{L}\right) - \hat{\pi}(1 + r_L)\frac{\partial \hat{L}}{\partial \hat{\pi}} = (1 + r_L) L_2 > 0. \tag{32.20}$$

On the other hand, if all interbank loans are recalled, $\gamma = 1$, we get from equation (32.16) that

$$\Pi_B^1 = \frac{M}{D + M}\lambda\hat{\pi}(1 + r_L) L_2 = \lambda\hat{\pi}(1 + r_L) M \tag{32.21}$$

and hence

$$\frac{\partial \Pi_B^1}{\partial \hat{\pi}} = \lambda(1 + r_L) M > 0. \tag{32.22}$$

We had established above that the lending bank will either recall all interbank loans, $\gamma = 1$, or no interbank loans are recalled, $\gamma = 0$, such that we can limit our analysis to these two cases.

Let us denote the difference in the profits to the lending bank from retaining all interbank loans, equation (32.19), and recalling all interbank loans, equation (32.21), by $\Delta\Pi_B^1$. Then we can subtract equation(32.19) from equation(32.21) to obtain

$$\Delta\Pi_B^1 = \frac{D}{\lambda} - \hat{\pi}(1 + r_L)((1 - \lambda) M + D), \tag{32.23}$$

from which we easily get that

$$\frac{\partial \Delta \Pi_B^1}{\partial \hat{\pi}} = -\left(1 + r_L\right)\left(\left(1 - \lambda\right) M + D\right) < 0. \tag{32.24}$$

Let us now assume that at $\hat{\pi} = \pi^*$ we find $\Delta \Pi_B^1 > 0$ and hence recalling interbank loans is more beneficial. Because $\Delta \Pi_B^1$ is decreasing in $\hat{\pi}$ due to (32.24), the lending bank would always call in their interbank loans. If, on the other hand $\Delta \Pi_B^1 < 0$ at $\hat{\pi} = \pi^*$, then there will exist a π^{**} such that $\Delta \Pi_B^1 = 0$, where $\pi^{**} < \pi^*$. If $\hat{\pi} < \pi^{**}$, the lending bank will recall their interbank loans.

In order to ensure that interbank loans are always recalled, we need $\Delta \Pi_B^1 > 0$, which from equation (32.23) implies that

$$M < M^{**} = \left(\frac{1}{\lambda \pi^{**}\left(1 + r_L\right)} - 1\right) \frac{D}{1 - \lambda}. \tag{32.25}$$

meaning that if the interbank loan is sufficiently small, we have $\pi^{**} = \pi^*$ and despite the borrowing bank experiencing a bank run, interbank loans are not recalled from this bank.

We also easily find that from equation (32.23) we get

$$\frac{\partial \Delta \Pi_B^1}{\partial M} = -\hat{\pi}\left(1 + r_L\right)\left(1 - \lambda\right) < 0. \tag{32.26}$$

Using equation (32.24) and the implicit function theorem, we get

$$\frac{\partial \pi^{**}}{\partial M} = -\frac{\frac{\partial \Delta \Pi_N^1}{\partial M}}{\frac{\partial \Delta \Pi_B^1}{\partial \pi^{**}}} < 0 \tag{32.27}$$

and from equation (32.18) we have

$$\frac{\partial \pi^*}{\partial M} = -\frac{D}{1 + r_L} \frac{1}{\left(D + L\right)^2} < 0. \tag{32.28}$$

We have thus established that interbank loans are withdrawn from a bank facing a bank run due to an increase the loan risk if this increase is sufficiently large and the interbank loans are not too large. The losses the borrowing bank would make from having to raise additional funds to meet the requirements of the withdrawn interbank loans will be so large, that the lending bank would prefer to wait for the remaining loans to be repaid; this repayment for large interbank loans will be higher than forcing the bank to sell more loans at a loss. Thus for large interbank loans, these are nor called in, even if the borrowing faces a bank run.

We will now consider the impact the losses interbank loans impose on the lending bank, in particular whether these losses lead to a bank run on the lending bank.

Inducing bank runs in the lending bank We now focus on the depositors of the lending bank and their decision whether to withdraw their deposits. The profits of the lending bank are given by

$$\Pi_B^1 = \pi (1 + r_L) L_1 - (1 + r_D) D + \Pi_B^2 \tag{32.29}$$

where the last term is the profits of the borrowing bank. As this bank is failing, all its value will be distributed to the lending bank. If $\Pi_B^1 > 0$, the depositors have been paid in full and there is no need to instigate a bank run.

For $\gamma = 0$, we get $\Pi_B^2 = \max\left\{0; \hat{\pi} (1 + r_L) \left(L_2 - \hat{L}\right)\right\}$, as these represent the resources available to repay interbank loans as discussed for equation (32.19), and for $\gamma = 1$ we have $\Pi_B^2 = \lambda\hat{\pi} (1 + r_L) M$ as in equation (32.21). In both cases we easily see that $\frac{\partial \Pi_B^2}{\partial \hat{\pi}} > 0$ and hence from equation (32.29) that $\frac{\partial \Pi_B^1}{\partial \hat{\pi}} > 0$. Therefore there will exist a π^{***} such that $\Pi_B^1 = 0$ and the lending bank experiences a bank run if $\hat{\pi} < \pi^{***}$. This inference is made under the assumption that for $\hat{pi} = \pi$ the bank profits are positive and for $\hat{\pi} = 0$ these are negative. As the repayment of the interbank loan, Π_B^2 is always non-negative, in order for Π_B^1 to become negative, we need the first two terms in equation (32.29) to be negative, which will only happen if L_1 is sufficiently small as we assume that $\pi (1 + r_L) > 1 + r_D$. As $L_1 = D - M$, this implies a sufficiently large interbank loan, $M > M^{***} = \frac{\pi(1+r_L)-(1+r_D)}{\pi(1+r_L)} D$, for any bank run on the lending bank to be possible. Hence bank runs on the lending bank only occur if the exposure of the bank to interbank loans is sufficiently large such that losses from interbank lending exceed the profits from the other lending of the bank providing the interbank loan.

We now get easily using $L_1 = D - M$ and equation (32.26)that for $\gamma = 1$ we have

$$\frac{\partial \Pi_B^2}{\partial M} = \frac{\partial \Pi_B^1}{\partial M} - \pi (1 + r_L) = (\lambda\hat{\pi} - \pi) (1 + r_L) < 0, \tag{32.30}$$

where the sign of this derivative emerges using that $\hat{\pi} < \pi$ and $\lambda < 1$. Similarly, we get that from equation (32.22) that

$$\frac{\partial \Pi_B^1}{\partial \hat{\pi}} = \frac{\partial \Pi_B^2}{\partial \hat{\pi}} = \lambda (1 + r_L) M > 0. \tag{32.31}$$

The implicit function theorem then gives us

$$\frac{\partial \pi^{***}}{\partial M} = -\frac{\frac{\partial \Pi_B^1}{\partial M}}{\frac{\partial \Pi_B^1}{\partial \pi^{***}}} = \frac{\pi - \lambda\hat{\pi}}{\lambda M} > 0. \tag{32.32}$$

If $\gamma = 0$, then $\Pi_B^2 = \max\left\{0; \hat{\pi} (1 + r_L) \left(L_2 - \hat{L}\right)\right\}$. Inserting for L_2 and $\hat{L}$ from equations (32.1) and (32.13), we get

$$\Pi_B^2 = \max\left\{0; \hat{\pi} (1 + r_L) M + \left(\hat{\pi} (1 + r_L) - \frac{1 + r_D}{\lambda}\right) D\right\}. \tag{32.33}$$

This then gives us when using equation (32.31) that

$$\frac{\partial \Pi_B^1}{\partial M} = (\hat{\pi} - \pi)(1 + r_L) < 0, \tag{32.34}$$
$$\frac{\partial \Pi_B^1}{\partial \hat{\pi}} = (1 + r_L) L_2 > 0,$$

hence similar to equation (32.32) we have

$$\frac{\partial \pi^{***}}{\partial M} = \frac{\pi - \hat{\pi}}{L_2} > 0. \tag{32.35}$$

It is easy to show that the slope in equation (32.32) is larger than in equation (32.35). The change in slope happens as the lending bank switches from not recalling interbank loans to recalling interbank loans, which happens at the point at which $\pi^{***} = \pi^{**}$.

We can now combine all elements of our analysis to provide a comprehensive analysis of the results we have obtained from this model.

Graphical analysis Figure 32.2 summarizes the results of our analysis. If the reduction in the repayment rate of the borrowing bank is sufficiently small such that $\hat{\pi} > \frac{1+r_D}{1+r_L}$, then from equation (32.2) we see that the borrowing bank is profitable and hence no bank run will occur. Once the repayment rate at the borrowing bank reduces further, the bank will make a loss and depositors in the affected bank withdraw their funds such that a bank run emerges. The lending bank does not withdraw their interbank loan because the losses from selling the loans outweigh the losses from retaining the interbank loan and as the lending bank remains profitable, no bank run will occur at that bank. The larger the interbank loan is, the more important these losses from withdrawing interbank loans become due to the borrowing bank having to sell more and more loans to meet the demands from the withdrawal of interbank loans; therefore the threshold in the repayment rate decreases the larger the interbank loan is.

An even larger reduction of the repayment rate at the borrowing bank will increase losses to the lending bank and hence with small interbank loans these are called in to limit the losses. For larger interbank loans, these are not recalled despite the borrowing bank being insolvent and not repaying all interbank loans. This is because recalling the interbank loan results in losses from the loan sale that outweigh the losses from the losses with the interbank loans remaining with the borrowing bank.

The same distinction on calling in interbank loans continues if the repayment rate reduces even further. However, here the losses to the lending bank from interbank loans that are not fully repaid are so larger that the lending bank also incurs an overall loss and consequently faces a bank run. The larger the interbank exposure is, the smaller the reduction in the repayment rate needs to be and for sufficiently small interbank loans, the losses to the lending bank are never sufficient to cause it to fail and a bank run to occur.

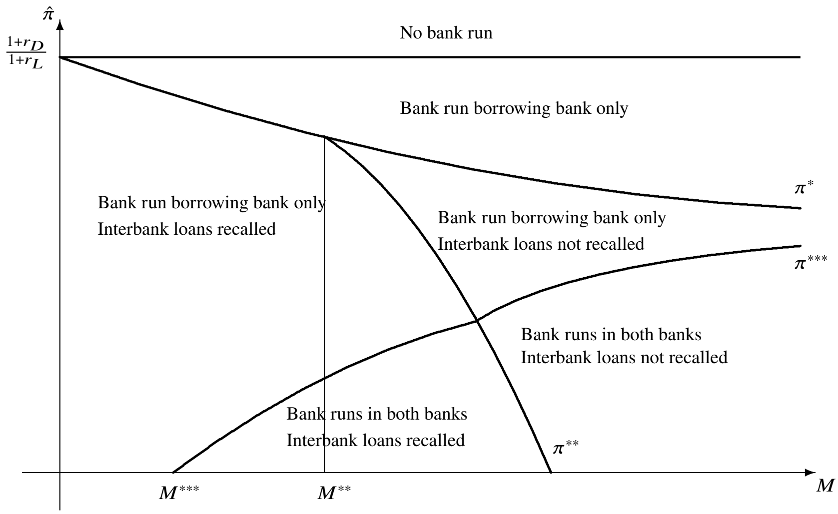


Fig. 32.2: Interbank lending and bank runs

We thus observe that for banks with a small exposure to the interbank market, their interbank loans are not withdrawn, even if they face a bank run, while with a a larger exposure to the interbank markets, these interbank loans will be withdrawn. This is due to the additional losses the borrowing bank will experience if more funds need to be raised to repay interbank loans, which reduces the payment the lending bank receives more than if they were to retain their interbank loans in the bank. We would thus expect banks with a few interbank loans to have their funding withdrawn in case there are doubt about the quality of their assets, while banks with a large exposure to interbank markets would continue to receive such funding.

Summary We have seen how in the event of adverse developments it is mainly banks with small exposures to the interbank market experience these to be withdrawn, while those with larger exposures retain acess to their interbank loans. but we have also seen that the losses imposed on banks providing interbank loans can lead to the spread of bank runs to an otherwise unaffected bank. the cause of the losses was the increase in the risk to a bank receiving interbank loans; this increased risk can cause a bank run in the affected bank. Such a bank run would be based on information and is an therefore efficient bank run. The losses arising from such a bank run can impose losses on banks providing interbank loans as these will also not be repaid fully. If these loans are sufficiently large for the lending bank, then this can induce a bank run at this bank; this bank run is also rational in that it is based on the losses the bank experiences, but it reflects systemic risk due as the failure of the borrowing bank due to its bank run has spread to the lending bank. This contagion is transmitted through interbank loans. These interbank loans may be retained with the borrowing bank in order to minimise losses, but the losses on such loans will induce a bank run on the lending bank.

Reading Li, Milne, & Qiu (2016)

32.1.2 Coordination failure

Banks are providing each other with interbank loans, often such loans are the result of depositors transferring funds between accounts at different banks; rather than clearing such payments, banks may retain any imbalances as interbank loans. Having thus an exposure to each other, it is intuitive that the failure of one bank will impose losses on other banks, who in turn might fail. However, here we are mainly concerned whether banks retain their interbank loans if a single bank becomes insolvent. Banks might retain interbank loans even with an insolvent bank in order to minimise their losses or they might decide to withdraw interbank loans from otherwise solvent banks fearing losses from their exposure to the failing bank. This withdrawal of interbank loans could then impose additional costs on banks whose funding is withdrawn, causing them to become subject to insolvency. We will here explore under which conditions interbank loans to insolvent banks are retained as well as interbank loans to solvent banks withdrawn.

We consider an economy with N banks that invest their deposits D into long-term loans L_i, where deposits can be withdrawn at any time. If deposits are withdrawing early, the bank has to reduce the amount of loans they hold, for example by selling them if they cannot be recalled from their borrowers. In addition to withdrawing deposits from the banking system as a whole, such as cash withdrawals or transfer into other countries, depositors will also transfer funds between banks. Such transfer we assume are not settled but as temporary imbalances the receiving bank gives an interbank loan to the originating bank, at interest rate r_M^i.

As a normal use of their deposits, depositors transfer funds to another bank with probability p and thus with probability $1 - p$ all funds stay within the bank; they might be transferred between accounts owned by different depositors at the same bank. Those transferring funds at bank i, will transfer them to bank j with probability p_{ij}. With $\mathbf{P}$ denoting the matrix of p_{ij}, we can define

$$\hat{\mathbf{P}} = (1 - p)\,\mathbf{I} + p\mathbf{P} \tag{32.36}$$

as the matrix that determines the destination of funds for each bank. As we are not interested in balances between banks to impose distortions, we assume that $\hat{\mathbf{P}}\iota = \hat{\mathbf{P}}^T\iota = \iota$, ι representing a vector with all entries equal to 1, and hence the inflows and outflows of each bank are balanced. The first term represents those transfers that remain within the bank and $\mathbf{I}$ represents the identity matrix.

We now assume that of the funds transferred from bank i to bank j a fraction λ_{ij} is retained at that bank and the remainder withdrawn, necessitating loans to be liquidated. Thus depositors are withdrawing their funds from the banking system, such as a cash withdrawals or transfers to a different, foreign, banking system. A bank run can then be identified through $\lambda_{ij} = 1$ as in this case all funds at bank i are withdrawn. Hence after such withdrawals, the amount of loans the bank holds is given by

$$L_i = D - \sum_{j=1}^{N} \hat{p}_{ij} \left(1 - \lambda_{ij}\right) D, \tag{32.37}$$

where $\hat{p}_{ij}$ denotes the elements of $\hat{\mathbf{P}}$ as defined in equation (32.36). Here the second term denotes all the amounts withdrawn from the banking system that have been transferred from all the other banks, including with bank i itself.

Loans are repaid with probability π, including interest r_L. Hence at maturity of the loan, it will be worth $\pi\left(1 + r_L\right) L_i$. The other assets of the bank are the interbank loans given by other banks, provided these received deposits are not withdrawn, including any interest r_M^j. In terms of liabilities, there are the loans obtained from other banks, on which interest r_M^i is payable. In addition the not-withdrawn deposits need to be repaid, with interest r_D. With assets and liabilities equaling, we get

$$\begin{aligned} \pi\left(1 + r_L\right) L_i + \sum_{j=1}^{N} p p_{ij} \lambda_{ij} \left(1 + r_M^j\right) D \\ = \sum_{j=1}^{N} p p_{ij} \lambda_{ij} \left(1 + r_M^i\right) D + \left(1 + r_D^i\right) D. \end{aligned} \tag{32.38}$$

As interbank loans and deposits are clearly perfect substitutes for banks, we have $r_D^i = r_M^i$. We further assume that banks are solvent and the repayment from loans exceeds the funding costs, $\pi\left(1 + r_L\right) > 1 + r_D^i = 1 + r_M^i$. As cash does not pay any interest, interbank loans are given as long as $r_M^i > 0$ and depositors also retain their deposits if $r_D^i = r_M^i > 0$. In this case all deposits are retained, no bank run occurs, and no funds are withdrawn such that $\lambda_{ij} = 1$. These considerations reduce equation (32.37) to become $L_i = D$ and equation (32.38) then reduces to

$$\pi\left(1 + r_L\right) + p \sum_{j=1}^{N} p_{ij} \left(1 + r_M^j\right) = p\left(1 + r_M^i\right) \sum_{j=1}^{N} p_{ij} + \left(1 + r_M^i\right). \tag{32.39}$$

Noting that as $\hat{\mathbf{P}}\iota = \iota$ due to inflows and outflows in banks being equal, we also have $\mathbf{P}\iota = \iota$ as $\lambda_{ij} = 1$, and hence $\sum_{j=1}^{N} p_{ij} = 1$. Denoting by $\iota + \mathbf{r_M}$ the vector of $1 + r_M^i$, we can rewrite equation (32.39) in matrix form as

$$\pi\left(1 + r_L\right)\iota + p\mathbf{P}^T\left(\iota + r_M\right) = (1 + p)\left(\iota + \mathbf{r_M}\right). \tag{32.40}$$

Solving equation (32.36) for $p\mathbf{P}^T = \hat{\mathbf{P}}^T - (1 - p)\,\mathbf{I}$, we can rewrite equation (32.40) as

$$\iota + \mathbf{r_M} = \pi\left(1 + r_L\right)\left(2\mathbf{I} - \hat{\mathbf{P}}^T\right)^{-1}\iota \tag{32.41}$$

Let us now define

$$\boldsymbol{\Theta} = \left(2\mathbf{I} - \hat{\mathbf{P}}^T\right)^{-1} = \frac{1}{1+p}\left(\mathbf{I} - \frac{p}{1+p}\mathbf{P}^T\right)^{-1}, \tag{32.42}$$

using equation (32.36) for the final expression. Using matrix theory, we can show that all elements of $\boldsymbol{\Theta}$ are non-negative and $\boldsymbol{\Theta}\iota = \iota$. Hence as $\pi(1+r_L) > 1$, we see from equation (32.41) that $\iota + \mathbf{r_M} = \pi(1+r_L)\,\iota$ and thus all positive interbank loan rates are positive. With interbank lon rates and deposit rates being equal, no deposits are withdrawn and interbank loans extended.

Insolvent bank We now assume that bank 1 is insolvent unexpectedly as loans at this bank are no longer repaid, $\pi_1 = 0$; all other banks remain solvent and the repayment rate of their loans remains unchanged. With $\hat{\iota} = [0 \quad \iota]^T$, equation (32.41) becomes

$$\iota + \mathbf{r_M} = \pi(1+r_L)\left(2\mathbf{I} - \hat{\mathbf{P}}^T\right)^{-1}\hat{\iota}. \tag{32.43}$$

We are now interested in the condition that interbank loans are extended to all banks, including the insolvent bank 1, thus we need again $1 + r_M^i > 1$, including $1 + r_M^1 > 1$. This condition will in general depend on the detailed structure of interbank loans as represented in the matrix of these loans, $\mathbf{P}$. Therefore we consider a special case where loans are perfectly evenly distributed between banks, such that

$$\mathbf{P} = \frac{1}{N-1}\begin{pmatrix} 0 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \tag{32.44}$$

which is symmetric. From matrix theory we know that for any matrix $\mathbf{A}$ we have $(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{+\infty}\mathbf{A}^k$. Hence from equation (32.42) we obtain

$$\begin{aligned}\boldsymbol{\Theta} &= \frac{1}{1+p}\sum_{k=0}^{+\infty}\left(\frac{p}{1+p}\mathbf{P}\right)^k \\ &= \frac{1}{1+p}\sum_{k=1}^{N}\left(\frac{p}{1+p}\right)^k\left(\beta_k\mathbf{I} + (1-\beta_k)\,\mathbf{P}\right)\end{aligned} \tag{32.45}$$

as $\mathbf{P}^k = \beta_k\mathbf{I} + (1-\beta_k)\,\mathbf{P}$ and $\beta_k = \frac{1}{N}\left(1 - \left(1 - \frac{1}{N-1}\right)^{k-1}\right)$. The derivation of β_k is based on the assumption that all interbank loans are identical.

As the solvent banks will obtain interbank loans if the insolvent bank 1 does, we focus on bank 1 only and can then conclude that if the insolvent bank obtains interbank loans, all banks can obtain such loans; this is because the position of the other bank is better than of bank 1 and hence there is no reason to extend an interbank loan to the insolvent bank 1 but not the other solvent banks. Using equation (32.43) in connection with equation (32.45), we can easily calculate the first row of $\iota + \mathbf{r_M}$ as

$$1 + r_M^1 = \frac{\pi (1 + r_L)}{1 + p} \sum_{k=0}^{+\infty} \left(\frac{p}{1 + p} \right)^k (1 - \beta_k) \frac{N - 1}{N - 1} = \frac{\pi (1 + r_L)}{\frac{1+p}{p} + \frac{1}{N-1}}, \tag{32.46}$$

where the first equality emerges from noting $\mathbf{I}\hat{\iota} = 0$ and all entries in $\mathbf{P}$ are $\frac{1}{N-1}$ and the second equality applies the theory of infinite sums. The interbank loan being retained requires $1 + r_M^1 > 1$, which easily solves for

$$p > p^* = \frac{1}{\pi (1 + r_L) - \frac{N}{N-1}}. \tag{32.47}$$

This result implies that, provided depositors transfer funds at a sufficient rate p and thus interbank loans are plentiful, an insolvent bank will be able to survive and not be forced to close as interbank loans are extended. This is because any losses of this bank are widely distributed and losses to each bank providing an interbank loan to this bank are therefore small, allowing them to retain interbank loans. This loss is smaller than the losses that would arise if interbank loans are withdrawn by depositors to make cash transfer instead, as this reduces the loans they can give. Hence the insolvent bank is not liquidated, but remains active due to interbank loans covering their losses and spreading them widely between banks and depositors, avoiding a bank run. Thus, although the bank is insolvent, it will not be liquidated as it the same of loans to finance any withdrawal of funds would impose higher losses.

Bank failure Assume now that the insolvent bank is actually closed. In this case all funds in the bank are withdrawn and we assume they are lost to the banking system. This is equivalent to a fraction $\lambda_{1j} = 0$ being retained after transferring deposits and equation (32.37) becomes

$$L_i = D - pp_{i1}D - \sum_{j=2}^{N} \hat{p}_{ij} \left(1 - \lambda_{ij}\right) D \tag{32.48}$$

Looking again at the case that for the other banks no deposits are withdrawn from the banking system, i.e. $\lambda_{ij} = 1$, we get when inserting this into the above relationship that

$$L_i = (1 - pp_{i1}) D \tag{32.49}$$

The balance of assets and liabilities in equation (32.38) is retained, but we have to exclude bank 1 as this has failed. If we note that $p_{ij} = \frac{1}{N-1}$ and $\sum_{j=2}^{N} p_{ij} = 1 - \frac{1}{N-1}$, we thus get

$$\pi (1 + r_L) \left(1 - pp_{ij}\right) + p \sum_{j=2}^{N} p_{ji} \left(1 + r_M^j\right) \tag{32.50}$$
$$= p \left(1 - \frac{1}{N - 1} \left(1 + r_M^i\right) + \left(1 + r_M^i\right) (1 - pp_{i1})\right).$$

denoting $\mathbf{P}_1$ as the matrix $\mathbf{P}$ from which the first row and the first column have been removed, and similarly for $\hat{\mathbf{P}}$, then equation (32.50) becomes

$$\begin{aligned}\left(1 + p - \frac{2p}{N-1}\right)(1 + \mathbf{r_M}) = \pi\,(1 + r_L)\left(1 - \frac{p}{N-1}\right)\iota \\ + p\mathbf{P}_1\,(1 + \mathbf{r_M}) =\end{aligned} \tag{32.51}$$

Following similar steps to deriving equation (32.41), we get

$$\begin{aligned}\boldsymbol{\Theta} &= \left(2\mathbf{I} - \frac{N-1}{N-1-p}\hat{\mathbf{P}}_1\right)^{-1} \\ &= \frac{1}{1+p}\left(\mathbf{I} - \frac{p}{1+p}\frac{N-1}{N-1-p}\mathbf{P}_1\right)^{-1} \\ &= \frac{1}{1+p}\sum_{k=0}^{+\infty}\left(\frac{p}{1+p}\frac{N-1}{N-1-p}\mathbf{P}_1\right)^{k}\end{aligned} \tag{32.52}$$

with $\mathbf{P}_1^l = \beta_k M + (1 - \beta_k)\,\mathbf{P}_1$ and $\beta_k = \frac{1}{N-1}\left(1 - \left(1 - \frac{1}{N-2}\right)^{k-1}\right)$, as N is replaced by $N - 1$ due to bank 1 being eliminated. As now all remaining banks are equal, we get similar to equation (32.46) that

$$1 + r_M^i = \frac{\pi\,(1 + r_L)}{\frac{1+p}{p}\frac{N-1-p}{N-1} + \frac{1}{N-2}}. \tag{32.53}$$

Comparing the expression in equation (32.53) with that in equation (32.46) it is easy to show that for the same transfer rate of deposits between banks, p, the expression here is larger. If we define p^{**} as the value of p for which $1 + r_M^i = 1$ in equation (32.53), then $p^{**} < p^*$. To see this note that

$$\frac{\partial\left(1 + r_L^i\right)}{\partial p} = \frac{\left(\frac{N-1}{p} + p\right)\pi\,(1 + r_L)}{p\,(N-1)\left(\frac{1+p}{p}\frac{N-1-p}{N-1} + \frac{1}{N-2}\right)^2} > 0, \tag{32.54}$$

hence a lower interest rate corresponds to a lower value for p.

If $p > p^{**}$ and hence sufficiently large amounts of interbank lending then the surviving solvent banks will obtain interbank loans and deposits are retained, avoiding a bank run for any solvent bank. For lower levels of interbank lending, $p < p^{**}$, even solvent banks are liquidated because the costs of selling loans are less than the costs from the losses of the failing bank on interbank loans; this is because these losses are distributed between fewer interbank loans. Hence all banks fail in a bank run and contagion has occurred as no other bank, except bank 1, has experienced any losses from their normal business. The losses that occur only arise because banks are not willing to provide interbank loans, necessitating the sale of loans at a loss. Thus if interbank loans are sufficiently rare, then the losses of bank 1 will cause interbank loans between all banks to be withdrawn, causing all banks to fail. The liquidation of the insolvent bank becomes self-fulfilling as no interbank loans would be advanced that would allow this bank to survive.

Summary Combining the results we have obtained, we see that if $p > p^*$, insolvent banks retain their deposits and interbank loans. As interbank loans are extended to the insolvent bank, it is not failing despite being insolvent. If, however, $p < p^{**}$, even solvent banks are failing as they cannot obtain interbank loans. Only if $p^{**} < p < p^*$ are solvent banks surviving and insolvent banks failing. The wrong decision stems from a failure of coordination between banks and depositors. In $p > p^*$ they do not manage to agree that the insolvent bank should fail as this increases their losses and if $p < p^{**}$ they cannot coordinate to keep the solvent banks open for the same reason. Thus banks coordinate their responses to an insolvent bank poorly, withdrawing interbank loans in cases of few such loans will lead to the spread of the failure by a single bank, while in cases with a high level of interbank lending, even the insolvenyt bank survives. Banks are only able to coordinate their withdrawal of interbank loans optimally in an intermediate range of interbank lending where solvent banks are retaining their interbank loans, while they are withdrawn from the insolvent bank.

We thus see that contagion affecting solvent banks through interbank loans occurs if the interbank loans are not very common, while if interbank loans are very common, banks will never fail, even if they are insolvent. It is only for an intermediate range of interbank loans that insolvent banks are failing while solvent banks are continuing to operate. Systemic risk would here be the result of banks sharing risks insufficiently through interbank loans; on the other hand, excessive risk-sharing will cause inefficiencies in that insolvent banks are not failing.

Reading Freixas, Parigi, & Rochet (2000)

32.1.3 Spread of liquidity shocks

Banks account for the potential withdrawal of deposits by holding cash reserves and if the withdrawal of deposits is higher than they have accounted for are normally able to borrow additional cash reserves in the interbank market. This mechanism works if the cash reserves in the banking system overall are sufficient to meet the deposit withdrawals, thus if a bank faces a higher withdrawal rate, other banks will have to face a lower withdrawal rate. We will investigate here how an unexpected rise in the withdrawal rate at a single bank can impose losses on other banks and lead to their failure.

Banks hold short-term deposits D that can be withdrawn at any time at face value or kept long-term to give interest r_D. Depositors may want to withdraw their monies early for reasons exogenous to the model with probability p_i; depositors might be motivated by consumption needs or the investment of their deposits into other assets. Here we assume that the withdrawal rate is either, p_H, or low, $p_L < p_H$, with equal probability. Which withdrawal rate the bank faces is not known to the bank, but is only revealed at the time of early withdrawal as banks observe their withdrawals. We define $p = \frac{p_H + p_L}{2} = p_H - \frac{1}{2}\Delta p$, where $\Delta p = p_H - p_L$, as the average withdrawal rate.

Banks invest deposits into a long-term loan, that is repaid with probability π, including interest r_L. Cash reserves C are held to meet early withdrawals of deposits as loans can only be liquidated at a fraction λ of their value if the bank required additional cash reserves to meet the withdrawal of deposits. In order to meet the withdrawals, a bank needs on average cash reserve of

$$C = pD. \tag{32.55}$$

When holding such a cash reserve, banks with a high withdrawal rate, p_H, could borrow from banks with a low withdrawal rate, p_L, to meet their demand for cash reserves. This would also allow banks with the low withdrawal rate to invest their excess cash. At maturity of the loans they have provided, banks repay those deposits that have not been withdrawn with interest from the proceeds of the loan such that we have

$$(1 - p)(1 + r_D) D = \pi (1 + r_L) L. \tag{32.56}$$

With the deposit rate determined such that this equation is fulfilled, they extract any surplus from the bank and thus we assume that banks are competitive in order to attract depositors.

Optimal interbank lending Let us now assume that banks make advance provisions for such interbank lending by lending to each other from the in the initial time period and not only once the withdrawals are realised. Assuming each bank lends an amount of $M = \frac{2}{N}(p_H - p) D$ to each of the other $N - 1$ banks, the resulting allocation is efficient as will be shown. A bank facing withdrawal rate p_H, would have to repay deposits $p_H D$ and the interbank loans the other banks facing p_H would recall to meet these withdrawals. There are a total of $\frac{N}{2}$ banks with a high withdrawal rate as we assumed high and low withdrawal rates are equally likely, thus $\frac{N}{2} - 1$ such loans are recalled as the bank considered is one of the banks with such a high withdrawal rate. To meet these withdrawals, banks use their cash reserves C and call in all $N - 1$ outstanding interbank loans from other banks. Hence we require that

$$p_H D + \left(\frac{N}{2} - 1\right) \frac{2}{N} (p_H - p) D = C + (N - 1) \frac{2}{N} (p_H - p) D, \tag{32.57}$$

which simplifies to equation (32.55). Similarly, at maturity of the loan, the remaining deposits are to be repaid with interest, $(1 - p_H)(1 + r_D) D$. The bank also needs to repay the interbank loans given by the $\frac{N}{2}$ banks facing the low withdrawal rate p_L, who will not have found it necessary to withdraw their interbank loans. This is financed by the proceeds of the loans they have given. Note that all interbank loans the bank has given have already been recalled in the previous time period as depositors withdrew their funds. Thus we have

$$(1 - p_H)(1 + r_D) D + \frac{N}{2}\frac{2}{N} (p_H - p)(1 + r)D) D = \pi (1 + r_L) L, \tag{32.58}$$

which simplifies to equation (32.56). Note that the interest on deposits and interbank loans are identical as they are perfect substitutes for banks. We thus have established that for banks facing high withdrawal rates the provision of loans is efficient.

If the bank has low withdrawal rate p_L, then it will face the withdrawal of deposits as well as of the interbank loans of all banks facing p_H. They would not call in interbank loans themselves, though, and make these payments from their existing cash reserves C. In this case we have

$$p_L D + \frac{N}{2}\frac{2}{N}\left(p_H - p\right) D = C, \tag{32.59}$$

which again simplifies to equation (32.55). At the maturity of the loans the bank will repay the remaining depositors, with interest, the interbank loans to other banks facing p_L that have not been withdrawn and have funds from the loan and the interbank loans it has not called in. Hence we obtain

$$\begin{aligned}(1 - p_L)(1 + r_D) D + \left(\frac{N}{2} - 1\right)\frac{2}{N}(p_H - p)(1 + r_D) D \\ = \pi (1 + r_L) L + (N - 1)\frac{2}{N}(p_H - p)(1 + r_D) D,\end{aligned} \tag{32.60}$$

again simplifying to equation (32.56). Hence this use of interbank loans yields the efficient outcome for banks facing a low withdrawal rate. Hence it is optimal for banks to provide interbank loans of size M, as defined above, to all banks.

Losses in the banking system If due to the deposit withdrawals a bank were to sell all its assets, it would obtain its cash reserves, the value of the sold loans, $\lambda\pi (1 + r_L) L$, and the interbank loans it has provided and which we assume are repaid at a fraction γ_i, hence $\frac{2}{N}(p_H - p) D \sum_{j=1, j\neq i}^{N} \gamma_i$. As all banks are identical, the γ_i are all identical, $\gamma_i = \gamma$ and hence the amount of interbank loans repaid becomes $2\frac{N-1}{N}\gamma (p_H - p) D$. The total liabilities consist of the deposits D and the interbank loans obtained, $(N - 1)\frac{2}{N}(p_H - p) D$. Hence the fraction of interbank loans that can be repaid for the bank, such that assets and liabilities are equal, would be

$$\gamma = \frac{C + \lambda\pi (1 + r_L) L + 2\frac{N-1}{N}(p_H - p)\gamma D}{D + 2\frac{N-1}{N}(p_H - p) D}. \tag{32.61}$$

This expression solves for

$$\gamma = \frac{C + \lambda\pi (1 + r_L) L}{D}. \tag{32.62}$$

If $\gamma < 1$ the liquidity shortage of the bank would impose losses on other banks as interbank loans are not fully repaid; hence the bank facing a liquidity shock would spread losses across the banking system.

Banks can sell their loans to meet their obligations if facing losses; they require for those depositors that retain their funds an amount of $(1 - p_i) D$. From the loan they obtain a repayment of $\pi(1+r)L)\, L$ and therefore need to retain loans to the amount of $\frac{1-p_i}{\pi(1+r_L)}D$ to cover the repayment of their non-withdrawn deposits and can thus sell loans to the amount of $L - \frac{1-p_i}{\pi(1+r_L)}D$ to yield a cash reserves of $\lambda\left(L - \frac{1-p_i}{\pi(1+r_L)}D\right)$. If $p_i = p_L$ a smaller amount of cash is raised from selling loans as fewer deposits are withdrawn and hence less loans can be sold. The losses from interbank loans not being repaid fully are given by $(1-\gamma)\frac{2}{N}(p_H - p)\, D$, reflecting the fraction $1-\gamma$ of the totalk interbank loans M that is not repaid. These losses can be covered without causing another bank to fail if

$$(1-\gamma)\frac{2}{N}(p_H - p)\, D \leq \lambda\left(L - \frac{1-p_L}{\pi(1+r_L)}D\right). \tag{32.63}$$

The losses the bank is exposed to from interbank loans not being repaid has to be covered by the remaining loans being liquidated

Using equation (32.61) with $\gamma = 1$, thus assuming that all interbank loans are repaid, and noting that $L = (1-p)\, D$ representing the loans that can be given after deposits have been reduced by early withdrawals, we can rewrite equation (32.63) as

$$\begin{aligned}\Delta p&\left(\frac{1-\lambda\pi(1+r_L)}{N+(N-1)\Delta p}\left(1 - p_H + \frac{1}{2}\Delta p\right) - \frac{1}{2}\lambda\frac{\pi(1+r_L)-2}{\pi(1+r_L)}\right) \\ &< \lambda(1-p_H)\left(1 - \frac{1}{\pi(1+r_L)}\right).\end{aligned} \tag{32.64}$$

We see that if $\Delta p = 0$ and hence both withdrawal rates are identical, equation (32.64) is always fulfilled given that the right-hand side is positive for the reasonable assumption that providing loans is socially desirable, $\pi(1+r_L) > 1$, and the left-hand side is zero. If, on the other hand, $\Delta p = 1$ and the differences between withdrawal rates is maximal, implying that $p_H = 1$, $p_L = 0$ and hence $p = \frac{1}{2}$, equation (32.64) becomes

$$\frac{1-\lambda\pi(1+r_L)}{2N-1} - \lambda\frac{\pi(1+r_L)-2}{\pi(1+r_L)} < 0. \tag{32.65}$$

With the reasonable assumption that loan rates are not excessive and $\pi(1+r_L) < 2$, the second term is negative. Thus if $\lambda\pi(1+r_L) < 1$, as is reasonable to ensure loan liquidations are causing losses, then equation (32.65) can never be fulfilled. consequently, there will exist a Δp^* defined such that equation (32.64) holds with equality. If $\Delta p < \Delta p^*$, then no other bank will suffer losses and hence no other bank will fail. When deriving equation (32.56) we had assumed that banks are competitive and make no profits and we have not introduced equity; thus any loss would induce a bank to fail such that the absence of banks making an actual loss will ensure that all banks survive. In this case this is because with sufficiently small differences between possible withdrawal rates, a small Δp, the amount of interbank lending is sufficiently small that they can be covered by the profits from ordinary loans that banks have provided.

Similarly, small interbank loans emerge if N becomes sufficiently large. In this case, $N \to +\infty$, the first term on the left-hand side of equation (32.64) vanishes and the condition becomes

$$\Delta p < -\frac{(1-p_H)\left(\pi\left(1+r_L\right)-1\right)}{2-\pi\left(1+r_L\right)}, \tag{32.66}$$

noting our assumption that $\pi\left(1+r_L\right) < 2$. As long as the differences between withdrawals rates is sufficiently small, the interbank lending to each bank will be suffic9iently small such that any losses can be covered by losses from the ordinary loans.

We have thus seen that as long the interbank loans are sufficiently small, losses will not become so large that they cause a bank to fail. We will now explore the possibility of a single bank receiving a liquidity shock and how this affects potential contagion to other banks.

Liquidity shock to a single bank Assume now that the withdrawal rate of one bank only increases to $\hat{p}_H > p_H$ unexpectedly. With the high withdrawal rate of this bank now being $\hat{p}_H$, from equation (32.56) we see that losses can become $\lambda\left(L-\frac{1-\hat{p}_H}{\pi(1+r_L)}D\right)$ before they impose losses on other banks. Banks expect a withdrawal rate of p_H, hence the additional funds required by this bank are $\left(\hat{p}_H-p_H\right)D$; this amount needs to be covered from selling additional loans. Hence if

$$\left(\hat{p}_H-p_H\right)D > \lambda\left(L-\frac{1-\hat{p}_H}{\pi\left(1+r_L\right)}D\right), \tag{32.67}$$

losses would not be covered by the profits from ordinary loans and would therefore spread and cause other bank to fail. This condition can be solved for

$$\hat{p}_H > \frac{\pi\left(1+r_L\right)\left(\lambda\left(1-p\right)+p_H\right)-\lambda}{\pi\left(1+r_L\right)-\lambda}, \tag{32.68}$$

using that $L=(1-p)\,D$. Thus if the withdrawal rate of a bank becomes sufficiently high, contagion will occur in the banking system. If ordinary loans are fully illiquid, $\lambda = 0$, we immediately see that the condition simplifies to $\hat{p}_H > p_H$ and any increase in withdrawal rates will lead to losses spreading as no funds can be raised from selling loans. In this case we see that condition (32.64) is also violated, causing bank failures to spread in the banking system.

Thus, if withdrawal rates at banks increase too much, they are not able to repay their interbank loans and impose losses on other banks, which causes them to fail and thus cause contagion.

Summary Hence we see that liquidity shortages from higher withdrawal rates of deposits will lead to losses being imposed on banks through the partial repayment of interbank loans; this is due to ordinary loans having to be sold by such banks to raise additional funds to return the withdrawn deposits. The losses such sales accrue or larger than the profits from the retained loans and hence interbank loans cannot be repaid in full. To instigate this contagion, no initial losses to a bank are needed, only a liquidity shortage.

Provided differences between withdrawals rates are not too large, losses that accrue from banks facing high withdrawal rates of deposits can be covered from the profits on ordinary loans as long as these higher withdrawal rates are expected by banks. It is the unexpected, and hence unaccounted for increase the withdrawal rate at a bank that will cause contagion and cause banks to fail as a result.

Reading Allen & Gale (2000)

32.1.4 Counterparty risk

Inter bank loans are risky in that they may not be repaid if the borrowing bank fails, due to high default rates on their loans or a bank run; this is commonly referred to as counterparty risk. We will investigate here how such counterparty risk can be spread within the banking system, causing a bank to fail, even if its counterparty has not failed.

We consider two banks who lend to each other in order to hedge against the risk of depositors withdrawing early; banks face either a high withdrawal rates p_H or a low withdrawal rate $p_L < p_H$. Banks do now know the withdrawal rate they are facing. To cover the expected withdrawals, banks hold cash reserves C and bank i lends bank j the amount M_{ij}. As banks are assumed to be equal, we have $M_{ij} = M_{ji}$. The deposits D are invested into non-liquid loans $L_i = D - C + M_{ji} - M_{ij}$, that are repaid with interest r_L with probability π. The risk of loans, π, is not known to banks, but they are aware its distribution $F(\cdot)$. Loans and interbank loans can be sold to outside parties at a discount λ to raise cash reserves.

Let us first focus on the bank facing the high withdrawal rate, p_H. If, part from those depositors that withdraw early, depositors retain their monies with this bank, the amount of monies the depositors wishing to withdraw late, $(1 - p_H) D$, are given by

$$\begin{aligned}(1 - p_H) D &= C + (1 + r_M) M_{ij} + (1 + r_L) L - (1 + r_M) M_{ji} - p_H D \\ &\equiv C^H_{i2G} (1 - p_H)\end{aligned} \tag{32.69}$$

if the loan is repaid. The funds available to repay depositors are from the cash reserves held, the interbank loan repaid from the other bank and the repaid loan. From this the interbank loan obtained from the other bank has to be paid and the repayments of those withdrawing early. We define this expression by C^H_{i2G}.

If the loan is not repaid, then we have equivalently

$$(1 - p_H) C^H_{i2B} = (1 - p_H) D = C + (1 + r_M) M_{ij} - (1 + r_M) M_{ji} - p_H D, \tag{32.70}$$

with the only difference to equation (32.69) being that the loan has not been repaid.

If late depositors were to withdraw early and hence cause a bank run, the total funds available are

$$C_{i1}^{H} = C + \lambda\overline{\pi}\left(1 + r_L\right)L + M_{ij}, \tag{32.71}$$

where banks use their cash reserves, loans are sold off to generate cash reserves and the interbank loan granted to the other bank is recalled without interest; the interbank loan obtained is not repaid as it was not called in by that bank at the time. By $\overline{\pi}$ we denote the expected repayment rate of the loans, $\overline{\pi} = E\left[\pi\right]$.

Depositors would retain their monies if the expected repayment of doing so exceeds that of withdrawing early. Hence combining the two cases with and without the ordinary loan being repaid, we require that

$$\pi C_{i2G}^{H} + \left(1 - \pi\right) C_{i2B}^{H} \geq C_{i1}^{H}, \tag{32.72}$$

which solves for

$$\begin{aligned} \pi \geq \pi^* &= \frac{C_{i1}^{H} = C_{i2B}^{H}}{C_{i2G}^{H} - C_{i2B}^{H}} \\ &= \frac{\left(1 - p_H\right)\lambda\overline{\pi}\left(1 + r_L\right) + \left(1 - p_H\right)M_{ij} + p_H\left(D - C\right)}{\left(1 + r_L\right)L}. \end{aligned} \tag{32.73}$$

Thus, if the repayment rate of ordinary loans is sufficiently high, no bank run occurs and the bank survives. If the repayment rate is too low, the bank will fail and given the distribution of repayment rates, this happens with probability $\gamma_H = F\left(\pi^*\right)$. Thus depending on the realisation of the repayment rate, a bank run will occur in the bank facing a high withdrawal rate.

We now consider the second bank in our banking system and assume that this bank faces the low withdrawal rate. Let us first assume that the other bank does not face a bank run, we then get similar to equations (32.69), (32.70), and (32.71) that

$$\begin{aligned} C_{j2GS}^{L} &= \frac{C + \left(1 + r_M\right)M_{ji} + \left(1 + r_L\right)L - \left(1 + r_M\right)M_{ij} - p_L D}{1 - p_L}, \\ C_{j2BS}^{L} &= \frac{C + \left(1 + r_M\right)M_{ji} - \left(1 + r_M\right)M_{ij} - p_L D}{1 - p_L}, \\ C_{j1S}^{L} &= C + \lambda\overline{\pi}\left(1 + r_L\right) + \lambda\left(1 + r_M\right)M_{ji} - M_{ij}, \end{aligned} \tag{32.74}$$

where in the case of early withdrawals the interbank loan is sold early, too, and the interbank loan obtained is recalled early.

Similarly, for the case that the other bank faces a bank run, we have

$$\begin{aligned} C_{j2GF}^{L} &= \frac{C - \left(1 + r_M\right)M_{ij} + \left(1 + r_L\right)L - p_L D}{1 - p_L}, \\ C_{j2BF}^{L} &= \frac{C - \left(1 + r_M\right)M_{ij} - p_L D}{1 - p_L}, \\ C_{j1F}^{L} &= C - M_{ij} + \lambda\hat{\pi}\left(1 + r_L\right)L, \end{aligned} \tag{32.75}$$

noting that as the interbank loan the bank has obtained is not repaid as the other bank will not be able to repay their interbank loan due to their bank run.

We now assume that it is known to depositors of bank j whether the loan in bank i is repaid and whether the bank faces a bank run. If the other bank survives, we require

$$\pi C_{j2GS}^{L} + (1-\pi) C_{j2BS}^{L} \geq C_{j1F}^{L} \tag{32.76}$$

to avoid a bank run on this bank. This condition gives us

$$\begin{aligned} \pi \geq \pi_S^{**} &= \frac{C_{jiF}^{F} - C_{j2BS}^{L}}{C_{j2GS}^{L} - C_{j2BS}^{L}} \\ &= \frac{\left\{ \begin{array}{c} p_L (D-C) + (1-p_L) \lambda \overline{\pi} (1+r_L) L \\ + (1-p_L) (\lambda (1+r_M) - 1) M \end{array} \right\}}{(1+r_L) L}. \end{aligned} \tag{32.77}$$

In the case that the other bank fails, following the same steps this becomes

$$\begin{aligned} \pi \geq \pi_F^{**} &= \frac{C_{jiF}^{L} - C_{j2BF}^{L}}{C_{j2GF}^{L} - C_{j2BF}^{L}} \\ &= \frac{\left\{ \begin{array}{c} p_L (D-C) + (r_M + p) M_{ij} \\ + (1-p_L) \lambda \overline{\pi} (1+r_L) L \end{array} \right\}}{(1+r_L) L} > \pi_S^{**}. \end{aligned} \tag{32.78}$$

We make the approximation that

$$F\left(\pi_i^{**}\right) = F\left(\pi^{**}\right) + f\left(\pi^{**}\right)\left(\pi_i^{**} - \pi^{**}\right), \tag{32.79}$$

where $f\left(\pi^{**}\right) = \frac{F(\pi^{**})}{\partial \pi_i^{**}}$ denotes the density function and π^{**} is defined such that $\pi_F^{**} > \pi^{**} > \pi_S^{**}$. The expected probability of the probability of failure of this bank is then with the approximation in equation (32.79) given by

$$\begin{aligned} \gamma_L &= \gamma_H F\left(\pi_F^{**}\right) + (1-\gamma_H) F\left(\pi_S^{**}\right) \\ &= F\left(\pi^{**}\right) + f\left(\pi^{**}\right)\left(p_H \pi_F^{**} + (1-p_H) \pi_S^{**} - \pi^{**}\right), \end{aligned} \tag{32.80}$$

where $\gamma_h = F\left(\pi^*\right)$. Let us compare this result to that if depositors have no information on the failure of the other bank. In this case, the expected payoffs are used and thus

$$\begin{aligned} &\pi\left(\gamma_H C_{j2GF}^{L} + (1-\gamma_H) C_{j2GS}^{L}\right) \\ &\quad + (1-\pi)\left(\gamma_H C_{j2BF}^{L} + (1-\gamma_H) C_{j2BS}^{L}\right) \\ &\quad \geq \gamma_H C_{jiF}^{L} + (1-\gamma_H) C_{j1S}^{L}, \end{aligned} \tag{32.81}$$

which solves for

$$\pi \geq \pi^{**} = \frac{\left\{ \begin{array}{c} \gamma_H \left(C^L_{j1F} - C^L_{j1S} - C^L_{j2BF} + C^L_{j2BS} \right) \\ +C^L_{j1S} - C^L_{j1BS} \end{array} \right\}}{\left\{ \begin{array}{c} \gamma_H \left(C^L_{j2GF} - C^L_{j2GS} - C^L_{j2BF} + C^L_{j2BS} \right) \\ +C^L_{j2GS} - C^L_{j2BS} \end{array} \right\}} = \gamma_H \frac{(1 - \lambda (1 - p_L)) (1 + r_M) M}{(1 + r_L) L} + \pi^{**}_S > \pi^{**}_S, \tag{32.82}$$

when inserting from equations (32.74), (32.75), and (32.77). As

$$\pi^{**}_F - \pi^{**}_S = \frac{(1 - \lambda (1 - p_L)) (1 + r_M) M}{(1 + r_L) L} \tag{32.83}$$

we have that $\pi^{**} = \gamma_H \left(\pi^{**}_F - \pi^{**}_S \right) + \pi^{**}_S = \gamma_H \pi^{**}_F + (1 - \gamma_H) \pi^{**}_S$ and hence $\pi^{**}_F \geq \pi^{**} \geq \pi^{**}_S$. Inserting this relationship into equation (32.80), we get that $\gamma_L = F(\pi^{**})$ and hence information on the other bank does not effect the average number of failures of the second bank. As the loans are already withdrawn by the first bank facing the high withdrawal rate, the failure of the second bank does not affect the first bank at all.

in the absence of any interbank loans, $M_{ij} = 0$, we get from equations (32.73) and (32.82) that $\pi^{**} < \pi^{*}$ if $\lambda \overline{\pi} (1 + r_L) < 1$. Thus if the liquidation of loans can only happen at a loss, the bank with the lower withdrawal rate of deposits will be less risky than the bank with the higher withdrawal rate. This is because the threshold on the repayment rate of the ordinary loan to ensure no bank run occurs is lower, and hence the probability of such a bank run is reduced.

We get that the bank with the lower withdrawal rate is more likely to be subjected to a bank run than the bank with a high withdrawal rate of deposits, $\pi^{*} < \pi^{**}$, if the interbank loans are sufficiently high. The condition solves for

$$M_{ij} > \frac{(p_H - p_L) \lambda \overline{\pi} (1 + r_L) - 1}{\left\{ \begin{array}{c} (1 - p_H) - (1 - p_L) (\lambda (1 + r_M) - 1) \\ +\gamma_H (1 + r_M) (\lambda (1 - p_L) - 1) \end{array} \right\}}, \tag{32.84}$$

where we assume that p_H is sufficiently large for the denominator to be negative. Here the large exposure of the bank facing the low withdrawal rate of deposits, makes them more vulnerable to losses spreading from the other bank facing a higher withdrawal rate. It is this exposure to interbank loans that requires a high repayment rate from ordinary loans to cover any potential losses to avoid a bank run. Such a high repayment rate is less likely to be realised and hence bank runs are more likely than in the other bank facing a low withdrawal of deposits. They are less vulnerable as a spill over from the other bank, having a low withdrawal rate, is less likely. We have to note that it is not necessary for banks to actually make a loss or face a bank run before the losses can spread in the banking system; the mere possibility of such losses are sufficient to cause a bank run at the bank less affected by the risk of deposit withdrawals.

The higher likelihood of bank runs for banks with low deposit withdrawals arises from the a high exposure to interbank loans, which make the imposition of losses from a bank run due to sufficiently high defaults at the other bank more significant and can lead to losses that induce a bank run in this bank, even though it is unaffected by high defaults on their ordinary loans. Thus this exposure to counterparty risk induces is more likely to induce a bank run in this bank than the other bank who is more likely to experience a bank run due to high default rates. If on the other hand, the interbank loans are small, the bank facing higher deposit withdrawals will more likely be facing a bank run as lower default rates are required to induce losses on depositors as their withdrawal is higher and hence more loans need to be sold at a loss.

Reading Ahnert & Georg (2018)

Résumé

Interbank loans can be seen similar to deposit and as much as we can observe a bank run arising from depositors withdrawing funds, so could we observe a creditor run with interbank loans being withdrawn. The motivation and incentives for such creditor runs would be identical to that of a bank run as discussed in chapter 15. However, interbank lending has the potential to spread the failure of one bank to other banks. The origin of the initial bank run can be a fundamental reason in that banks face losses from their ordinary loans and these losses will then not allow banks to repay interbank loans in full, imposing losses on interbank lenders. Such losses can then lead to the failure of the bank if no other sources of profits can cover such losses, or they might lead toa bank run at that bank due to the uncertainty whether deposits can be repaid in full anymore. But losses can not only be transmitted in response to a bank facing higher defaults on their ordinary loans, a liquidity shortage in one bank, for example due to depositors withdrawing more funds than anticipated for exogenous reasons, can then lead to losses as banks seek to raise cash reserves, which will then be transmitted to the lending bank, who may fail due to these losses or the resulting bank run.

We have seen in various settings that a sufficiently large exposure to interbank loans makes the banking system vulnerable to contagion in that the failure of one bank, whether due to high loan defaults, a bank run, or only a higher than expected liquidity need, can spread to other banks who are not affected by these events. Interbank loans can therefore be cause of systemic risk, making banking systems more vulnerable the more widespread interbank lending is, notwithstanding the benefits interbank lending provides as we have discussed in chapter 16 at length.

32.2 Common exposures

Banks can reduce their risks by diversifying their loan portfolios, that is giving smaller loans to many different borrowers. If the demand by borrowers is for larger loans than a single bank can provide if it seeks to diversify, then the borrower will seek loans from multiple banks. The consequence if that banks will have exposure to the same company and hence the default of such a company can impose losses on multiple banks; these losses might then lead to the failure of several banks, while if the loan was given by a single undiversified bank, only this single bank would fail. Banks giving loans to the same companies is known as banks having a common exposure to such borrowers. If we interpret the failure of multiple banks as systemic risk, then commone exposure can increase systemic risk.

In chapter 32.2.1 we will explicitly look at the effect diversification has on systemic risk and what the optimal level of diversification for banks is. In contrast to this, chapter 32.2.2 will focus on the effect information has on the preferences of depositors for the diversification strategy of banks. We will explore how banks having access to more information about the risks they are facing when diversifying their loan portfolio and adjusting deposit rates accordingly will affect depositors and subsequently the choice by banks.

32.2.1 Diversification

Banks can affect the level of risk they are taking not only through the risk of each loan they are providing, but also the diversification of their loan portfolio. Thus rather than giving a small number of large loans, they could provide a large number of small loans; this diversified loan portfolio will generally have a lower risk. However, if banks provide smaller loans to more companies, companies will require loans from multiple banks such that they can obtain loans to requisite size. The consequence is that multiple banks will provide loans to the same companies, giving them a similar risk profile and which makes them vulnerable to the same adverse events and they are more likely to fail together, rather than only a single bank failing from their unique risk exposure. We will here investigate how banks will choose the optimal level of diversification and thereby determine the systemic risk of the banking system, which we consider here as the probability of multiple banks failing together.

Let us consider a situation in which two banks provide loans to two companies. Each bank can provide a loan of size L to both companies combined and each company requires a loan of size L. Companies repay loans with probability π_i, including interest r_L; the defaults of companies are independent of each other. This probability is random with an identical and independent distribution of $F(\cdot)$ and expected value π.

Banks can decide how to split the provision of loans among the two companies. Bank 1 would give a fraction γ of their loans to company 1 and consequently a fraction $1-\gamma$ to company 2. Complementary, Bank 2 provides a fraction $1-\gamma$ of their loans to company 1 and a fraction γ to company 2; we have thus that both companies receive a loan from banks of size L and both banks provide a loan of size L. With a deposit rate r_D and deposits D fully financing loans, we have the profits of the two banks given by

$$\Pi_B^1 = \gamma\pi_1 (1 + r_L) L + (1 - \gamma) \pi_2 (1 + r_L) L - (1 + r_D) L, \qquad (32.85)$$
$$\Pi_B^2 = (1 - \gamma) \pi_1 (1 + r_L) L + \gamma\pi_2 (1 + r_L) L - (1 + r_D) L,$$

where we use the realization of the default rate π_i here. In order for a bank to not fail, we need it to be profitable, thus $\Pi_B^i \geq 0$. From equation (32.85) we obtain that this requires companies to have repayment rates that fulfill the requirement that

$$\pi_1 \geq \pi^* = \frac{1 + r_D}{1 + r_L}\frac{1}{\gamma} - \frac{1 - \gamma}{\gamma}\pi_2, \qquad (32.86)$$
$$\pi_2 \geq \pi^{**} = \frac{1 + r_D}{1 + r_L}\frac{1}{1 - \gamma} - \frac{\gamma}{1 - \gamma}\pi_2.$$

Figure 32.3 shows the areas in which one or both banks fail, depending on which combination of repayment rates the two companies have. We can easily see from equation (32.86) that as the fraction of the loan to company 1 by bank 1, γ, increases from 0 to $\frac{1}{2}$, the slopes slowly converge and the failure of a single bank becomes less and less likely, while the area in which both banks fail, increases.

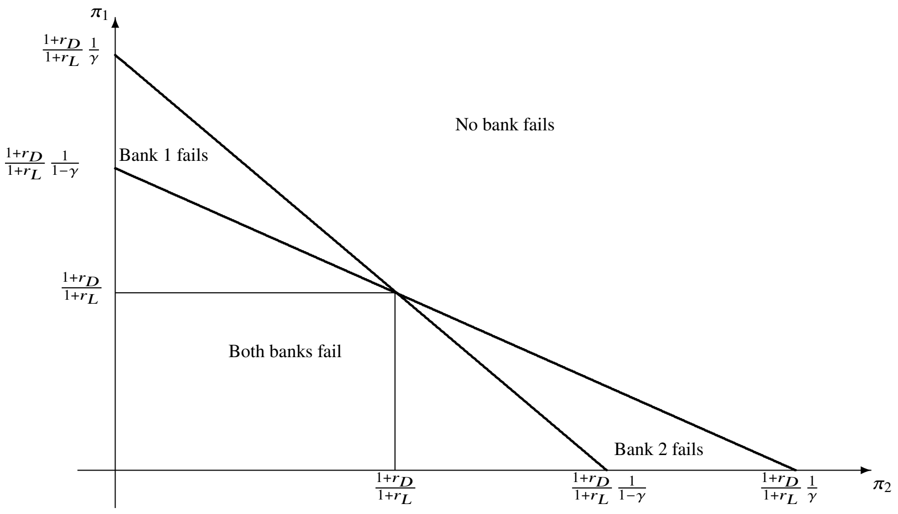


Fig. 32.3: Bank failures with diversified loans

If a bank fails, we assume that the other bank can take on the loan at a discount of $\lambda < 1$ of its expected value $\pi (1 + r_L) L$, where we assume that the other bank doe snot know the repayment rate of this loan. If both banks fail, outside investors can buy the loans at a larger discount $\hat{\lambda} < \lambda$, again not being aware of the risks of the loans they are purchasing.

The probability of bank 2 failing is the area under π^{**} in figure 1, which is given by

$$p_2 = \int_0^{\frac{1+r_D}{1+r_L}\frac{1}{\gamma}} \int_0^{\pi^{**}} dF(\pi_1)\, dF(\pi_2) \tag{32.87}$$
$$= \frac{1}{\left(1-\underline{\pi}\right)^2} \int_0^{\frac{1+r_D}{1+r_L}\frac{1}{\gamma}} \pi^{**} d\pi_2$$
$$= \frac{1}{2\left(1-\underline{\pi}\right)^2} \left(\frac{1+r_D}{1+r_L}\right)^2 \frac{1}{\gamma(1-\gamma)},$$

where for the second equality we assume that π_i is uniformly distributed over $\left[\underline{\pi};1\right]$. Similarly the probability of both banks failing is given as

$$p_{12} = \int_0^{\frac{1+r_D}{1+r_L}} \int_0^{\pi^{**}} dF(\pi_1)\, dF(\pi_2) \tag{32.88}$$
$$+ \int_{\frac{1+r_D}{1+r_L}}^{\frac{1+r_D}{1+r_L}\frac{1}{1-\gamma}} \int_0^{\pi^{*}} dF(\pi_1)\, dF(\pi_2)$$
$$= \frac{1}{2\left(1-\underline{\pi}\right)^2} \left(\frac{1+r_D}{1+r_L}\right)^2 \frac{3-\gamma}{1-\gamma}.$$

Assume the bank seeks to minimize the losses from failure as it has to sell off the loans at a discount of λ and $\hat{\lambda}_1$, for it alone failing or both banks failing, respectively. We know that the probability of only bank 2 failing is $p_2 - p_{12}$, given p_2 includes the case of both banks failing. Thus the total costs of failing to bank 2 are given by

$$C = (p_2 - p_{12})(1-\lambda)\pi(1+r_L)L \tag{32.89}$$
$$+ p_{12}\left(1-\hat{\lambda}\right)\pi(1+r_L)L$$
$$= \frac{1}{2\left(1-\underline{\pi}\right)^2} \left(\frac{1+r_D}{1+r_L}\right)^2 \pi(1+r_L)$$
$$\times \left(\frac{1-\lambda}{\gamma(1-\gamma)} + \frac{\left(\lambda-\hat{\lambda}\right)(3-\gamma)}{1-\gamma}\right).$$

Minimizing these costs over the best allocation of loans across the two companies, we obtain the first order condition as

$$\frac{\partial C}{\partial \gamma} = \frac{1}{2\left(1-\underline{\pi}\right)^2} \left(\frac{1+r_D}{1+r_L}\right)^2 \frac{\pi(1+r_L)L}{(1-\gamma)^2} \tag{32.90}$$
$$\times \left(2\left(\lambda-\hat{\lambda}\right) - \frac{(1-\lambda)(1-2\gamma)}{\gamma^2}\right)$$
$$= 0.$$

Solving equation (32.90) we derive the optimal share of the loan bank i gives company i as

$$\gamma^* = \frac{1}{1 + \sqrt{\frac{(1-\hat{\lambda})+(\lambda-\hat{\lambda})}{1-\lambda}}}. \tag{32.91}$$

We see that $0 < \gamma^* < \frac{1}{2}$ and hence the maximal level of diversification, $\gamma = \frac{1}{2}$ is not achieved; given that both companies have the same distribution of repayment rates and defaults are independent, an equal investment into both loans would reduce the overall risk most. This deviation from a full diversification arises because the benefits of reducing the failure of a single bank implies an increase in the failure of both banks as they become more alike. As this has higher costs through the higher discount if both banks fail $\hat{\lambda}$, these two effects have to be balanced. From equation (32.90) we see that if $\lambda = \hat{\lambda}$, we get $\gamma^* = \frac{1}{2}$, as the cost differential vanishes and hence companeis would diversify fully. If $\lambda = 1$ and thus a single bank failing would not make a loss when selling their loans, then $\gamma^* = 0$ and no diversification occurs, the joint failure of both banks is minimized as these are the only costs to the bank.

If we interpret the failure of both banks as an instant of systemic risk due to their chosen level of common exposure to companies, diversification increases this risk. Comparing the minimized systemic risk at $\gamma = 0$ with the systemic risk at the optimal level of diversification, $\gamma = \gamma^*$, then the ratio of these systemic risks is given by

$$\frac{p_{12}(\gamma = \gamma^*)}{p_{12}(\gamma = 0)} = 1 + \frac{2}{3}\sqrt{\frac{1-\lambda}{(1-\hat{\lambda}) + (\lambda - \hat{\lambda})}}. \tag{32.92}$$

We see that as the costs of a single bank failing reduce, λ increases, this ratio becomes approaches 1 as $\lambda = 1$ and the systemic risk is minimized; this is because it is optimal to not diversify the loan provision as this minimizes the costs from systemic risk. On the other hand if the costs of both banks failing does not increase beyond the costs of a single bank failing, $\hat{\lambda} = \lambda$, and thus systemic risk does not impose any additional costs, then this ratio becomes $\frac{5}{3}$ and the systemic risk increases by $\frac{2}{3}$ due to the diversification of banks.

We thus see that while there is no causal relationship between the two banks failing simultaneously, they both experience a default on loans they have obtained, their choice of loans will affect the likelihood of both banks failing together. It is the choice of banks to diversify their loan provision, and thus diversify their loan portfolio, that causes systemic risk to increase. Banks are doing this to reduce their risks by having a smaller exposure to the risk from each lender, making it less likely they will fail. However, such a wider spread of loans is only achievable if companies have loans from several banks, thus banks will have common exposures to companies, making the banks more alike. Thus, if one bank fails, it is quite likely that the other banks will also fail as they have an exposure to similar risks. Thus, the risk of banks failing might be reduced due to diversification, but if banks fail, there might be more banks failing; this can be interpreted as a higher systemic risk.

Reading Wagner (2010)

32.2.2 Information spillover

Banks gather information about the economy not only from observing macroeconomic data but also the ability of companies they provide with loans to repay these. Having a more information in the first instance should improve the decision-making of banks but also of depositors. However, we will assess here whether the additional information that depositors can generate from observing the defaults on the loans their bank has provided benefits them and how this affects the systemic risk that banks create by following the preferences of their depositors.

Let us assume that banks can give loans to two companies. The ability of these companies to repay their loan depends on the state of the economy, which is unobservable to banks. It either is in a good state G with probability p or a bad state B with probability $1 - p$. In the good state, such a well-performing economy, the companies can repay their loans more easily and we assume that this happens with probability $\pi > \frac{1}{2}$. In the bad state, for example during recessions, the companies repay their loans only with probability $1 - \pi < \frac{1}{2}$. Depositors can only observe whether a loan has been repaid in the past, the outcome H, or wether the company has defaulted, outcome L, but not the state of the economy itself.

As deposits ore only repaid if the company repays the bank, depositors will use past repayments as an indicator for the probability of this event S, given the history of observations Ω. Thus the deposit rate would be such that

$$Prob\left(S|\Omega\right)\left(1 + r_D^\Omega\right) D = D, \tag{32.93}$$

ensuring depositors break even, given they are promised a deposit rate r_D^Ω. From this condition we easily obtain the deposit rate to be

$$1 + r_D^\Omega = \frac{1}{Prob\left(S|\Omega\right)}. \tag{32.94}$$

If the bank invests only into one of the two companies, its depositors will only observe a signal of this one company such that $\Omega \in \{H; L\}$. Using Bayes' Theorem we easily get the probability of the state of the economy, given the observed outcome from past loans as

$$\begin{aligned}
Prob\left(G|H\right) &= \pi_S\left(G|H\right) = \frac{p\pi}{p\pi + (1-p)(1-\pi)}, \\
Prob\left(B|H\right) &= \pi_S\left(B|H\right) = \frac{(1-p)(1-\pi)}{p\pi + (1-p)(1-\pi)}, \\
Prob\left(G|L\right) &= \pi_S\left(G|L\right) = \frac{p(1-\pi)}{p(1-\pi) + (1-p)\pi}, \\
Prob\left(B|L\right) &= \pi_S\left(B|L\right) = \frac{(1-p)\pi}{p(1-\pi) + (1-p)\pi},
\end{aligned} \tag{32.95}$$

where the numerator determines the probability of the stated outcome (H or L), given the state (G or H) and the denominator represents the stated outcome and alternative outcome. We can now easily get

$$\begin{aligned} Prob\,(S|H) &= \pi_S\,(G|H)\,\pi + \pi_S\,(B|H)\,(1-\pi) \\ &= \frac{p\pi^2 + (1-p)\,(1-\pi)^2}{p\pi\,(1-p)\,(1-\pi)}, \\ Prob\,(S|L) &= \pi_S\,(G|L)\,\pi + \pi_S\,(B|L)\,(1-\pi) \\ &= \frac{\pi\,(1-\pi)}{p\,(1-\pi) + (1-p)\,\pi}. \end{aligned} \tag{32.96}$$

Similarly, if the bank provides loans to both firms, the depositors have two signals, one from each company, such that $\Omega = \{HH; HL; LH; LL\}$. Hence

$$\begin{aligned} \pi_D\,(G|HH) &= \frac{p\pi^2}{p\pi^2\,(1-p)\,(1-\pi)^2}, \\ \pi_D\,(B|HH) &= \frac{(1-p)\,(1-\pi)^2}{p\pi^2 + (1-p)\,(1-\pi)^2}, \\ \pi_D\,(G|LL) &= \frac{p\,(1-\pi)^2}{p\,(1-\pi)^2 + (1-p)\,p^2}, \\ \pi_D\,(B|LL) &= \frac{(1-p)\,\pi^2}{p\,(1-\pi)^2 + (1-p)\,p^2}, \\ \pi_D\,(G|HL) &= \pi_D\,(G|LH) = p, \\ \pi_D\,(B|HL) &= \pi_D\,(B|LH) = 1-p. \end{aligned} \tag{32.97}$$

Furthermore, as above we get

$$\begin{aligned} Prob\,(S|HH) &= \pi_D\,(G|HH)\,\pi + \pi_D\,(B|HH)\,(1-\pi)\,, \\ Prob\,(S|LL) &= \pi_D\,(G|LL)\,\pi + \pi_D\,(B|LL)\,(1-\pi)\,, \\ Prob\,(S|LH) &= Prob\,(S|HL) \\ &= \pi_D\,(G|LH)\,\pi + \pi_D\,(B|LH)\,(1-\pi)\,. \end{aligned} \tag{32.98}$$

We know further that

$$\begin{aligned} Prob\,(HH) &= p\pi^2 + (1-p)\,(1-\pi)^2\,, \\ Prob\,(HL) &= p\pi\,(1-\pi) + (1-p)\,(1-\pi)\,\pi = \pi\,(1-\pi)\,, \\ Prob\,(H) &= p\pi + (1-p)\,(1-\pi)\,, \end{aligned} \tag{32.99}$$

and can easily show that

$$\begin{aligned} Prob\,(HH) + Prob\,(HL) &= Prob\,(H)\,, \\ Prob\,(S \wedge HH) + Prob\,(S \wedge HL) &= Prob\,(S \wedge H)\,, \end{aligned} \tag{32.100}$$

with $\wedge$ denoting that both events are occurring. With the definition of conditional probabilities

$$Prob\,(S|\Omega) = \frac{Prob\,(S \wedge \Omega)}{Prob\,(\Omega)} \tag{32.101}$$

and equation (32.94), we finally get

$$1 = r_D^{\Omega} = \frac{Prob\,(\Omega)}{Prob\,(S \wedge \Omega)}. \tag{32.102}$$

Let us now consider the deposit rate that would be required if depositors observe the repayments of a single company only; we obtain

$$\begin{aligned} 1 + r_D^H &= \frac{Prob\,(H)}{Prob\,(S \wedge H)} \\ &= \frac{Prob\,(HH)}{Prob\,(S \wedge H)} + \frac{Prob\,(HL)}{Prob\,(S \wedge H)} \\ &= \frac{Prob\,(S \wedge HH)}{Prob\,(S \wedge HH) + Prob\,(S \wedge HL)}\left(1 + r_D^{HH}\right) \\ &\quad + \frac{Prob\,(S \wedge HL)}{Prob\,(S \wedge HH) + Prob\,(S \wedge HL)}\left(1 + r_D^{HL}\right) \end{aligned} \tag{32.103}$$

using equations (32.100) and (32.102).

If depositors are risk averse, they would prefer receiving deposit rate r_D^H over receiving either r_D^{HH} or r_D^{HL}, even though the expected return is the same. This is due to the risk in receiving an uncertain return, which will depend on the combinations of signals the depositor will observe. In the same way we can show that depositors prefer to receive r_D^L over a combination of either r_D^{LL} or r_D^{LH}. If depositors prefer to receive r_D^L and r_D^H over the alternatives when observing two signals, then banks can charge lower deposit rates when investing into a single company than diversifying into two companies. Hence, if the return on loans are identical, the bank would prefer giving loans only to a single company and not diversify their loan portfolio.

The reason for this result lies in the information spillover. Receiving more signals from the success or failure of loans to different companies leads to more uncertainty on the deposit rate without affecting the average deposit rate that depositors receive, and hence depositors prefer banks that invest into one company only. With banks investing into a single company only, each bank will have different borrowers and hence the failure of one borrower is most likely to affect only this one bank. Whereas if banks were to invest into two companies, they would share the loan and thus both be exposed to the same companies, making it more likely that both banks will fail together if their loans default. If we define systemic risk as the likelihood of both banks failing simultaneously, the systemic risk is reduced due to the information spillover.

We have thus seen that it will be optimal for banks to not diversify their loan portfolio. The additional information that is available about the state of the economy only increases uncertainty to depositors about the return they will obtain and thus they would prefer banks to reduce the information available. Less information about the state of the economy will reduce systemic risk.

Reading Acharya & Yorulmazer (2008)

Résumé

We have seen that banks diversify they loan portfolio suboptimally and deliberately expose themselves to systemic risk. In order to minimise systemic risk, banks should not diversify their loan portfolio; this way banks are only exposed to the risks of those loans they have provided and other banks are exposed to different risks. Multiple banks failing would only occur if losses accumulate in several loans simultaneously. However, for banks it is optimal to have some common exposure as this minimises their costs from failing; this will increase systemic risk as banks now will have a common exposure and are subjected to potential losses from multiple sources, which may cause many of these banks to fail from just a single risk being realised.

However, the exposure to multiple risks also induces uncertainty into the returns that banks can generate and thus the probability of their own default. Many factors need to be considered and information on these factors need to be evaluated, leading to constant revision of assessments and hence of the probability of default of the bank. This will adversely affect risk-averse depositors who would prefer a stable deposit rate that is less commonly changed. It is for this reason that depositors would prefer banks to not diversify their loan portfolio.

As so often, we have contradictory results about the effect of common exposure on systemic risk. On the one hand banks prefer to diversify their loan portfolio to some extent in order to reduce risks, but depositors would prefer them to not do so to ensure a more stable deposit rate. The overall result will in most cases be a compromise between these two contrasting result with the demands from depositors resulting in less diversification of the loan portfolio than the bank would optimally choose.

32.3 Fire sales

Banks facing a liquidity shortage will have to raise additional funds, either through borrowing from central banks, the interbank market, or the sale of their assets, most notably their loan book. The sale of assets in such a situation is referred to as a fire sale as the sale has to be completed fast and will commonly involve a substantial discount on the actual value of the asset. We will look here how such a fire sale can affect the systemic of the banking system.

If assets are sold by banks, they will have to be bought; given the nature of the assets, loans, the most likely purchasers will be other banks. Having other banks purchase these assets, will reduce their liquidity and could affect their ability to repay depositors if they demand so. Chapter 32.3.1 will explore such a scenario and see under which conditions the liquidity shortage of one bank can spread to other banks and the banking system will face multiple bank runs. Even if banks could sell assets to meet their immediate liquidity needs, this sale might be at a loss and the liquidity shortage may turn into a loss from the sale of assets, causing the bank to fail for this reason as chapter 32.3.2 will show.

32.3.1 Liquidity shortages from late loan repayments

Banks may face a liquidity shortage as cash inflows are lower than expected, which then can cause a bank run if depositors loose confidence in the ability of the bank to repay their funds. Such a bank run will require banks to sell some of their loans, which will require other banks to purchase them, reducing their liquidity and the liquidity shortage spreads and more banks fail in bank runs. We will investigate here under which conditions such bank runs can spread and how this relates to the overall liquidity position in the banking system.

Assume that borrowers will want to repay their loans either early with probability p_i, or at maturity with probability $1 - p_i$. This probability p_i is initially unknown to banks, but become known at the time the early repayments are made. All market participants know the distribution of p_i, which is denoted by $F(\cdot)$.

Loans that are repaid at maturity can be sold at a fraction λ of its final value, which is $\pi (1 + r_L) L$, where π denotes the probability the loan is repaid at all and r_L the loan rate. If loans are sold, the bank receives a fraction γ of the loan proceeds $\lambda\pi (1 + r_L) L$ at the same time as the early repayments and a fraction $1-\gamma$ at maturity of the loan. Banks will decide to sell a fraction μ_i of the loans that are not repaid early.

The proceeds the bank obtains at the early repayment date are reinvested into the banking system as deposits in other banks, earning interest r_D. As we assume no cash or equity in our model, loans are fully financed by deposits and the bank profits are given by

$$\begin{aligned}\Pi_B^i = {} & p_i\pi (1 + r_L)(1 + r_D) L \\ & +\mu_i (1 - p_i)(\gamma\lambda\pi (1 + r_L)(1 + r_D) L \\ & + (1 - \gamma)\lambda\pi (1 + r_L) L) \\ & + (1 - \mu_i)(1 - p_i)\pi (1 + r_L) L - (1 + r_D) D,\end{aligned} \tag{32.104}$$

where the first term represents the return on early repayments that have been reinvested into deposits at other banks, the second term shows the returns on the loans that are only going to be repaid at maturity but that have been sold and proceeds obtained early and late, respectively. The third term denotes the repayments on unsold loans repaid at maturity, and the final term the payments to depositors.

Banks decide how much of the loans repaid at maturity to sell. They will obtain a value of $\pi (1 + r_L) L$ if they do not sell their loans and $\gamma\lambda (1 + r_D)\pi (1 + r_L) L + (1 - \gamma)\lambda\pi (1 + r_L) L$ if they sell these loans. Given this incentive structure banks will either sell none of these loans, $\mu_i = 0$, or sell all such loans, $\mu_i = 1$, depending which expression is larger. Thus they will retain their loans if

$$1 + r_D < 1 + r_D^* = \frac{1 - \lambda (1 - \gamma)}{\lambda \gamma} \tag{32.105}$$

and sell these loans otherwise. Higher deposit rates make it more attractive for banks to sell their loans as they can earn interest on the proceeds they receive early. This is to be balanced against the loses from only receiving a fraction λ of the value with a fraction γ being paid early.

If depositors cannot be repaid fully, they would instigate a bank run to withdraw their funds just before the early repayments are made. To accommodate these deposit withdrawals all loans would be sold in case of a bank run. To avoid a bank run, the bank needs to be able to repay all depositors at the maturity of the loans, which will happen if they are making a profits, hence we need $\Pi_B^i \geq 0$. In the case of no loans being sold, $\mu_i = 0$, this requires

$$p_i \pi (1 + r_L) (1 + r_D) L + (1 - p_i) \pi (1 + r)L) - (1 + r_D) \geq 0$$

or

$$p_i \geq p^* = \frac{(1 + r_D) - \pi (1 + r_L)}{r_D \pi (1 + r_L)}. \tag{32.106}$$

If selling all loans being repaid at maturity, $\mu_i = 1$, we obtain the condition that

$$\begin{aligned} 0 \leq\ & (1 + r_D) p_i \pi (1 + r_L) L \\ & + (1 - p_i) (\gamma \lambda (1 + r_D) \pi (1 + r_L) + (1 - \gamma) \lambda \pi (1 + r_L) L) \\ & (1 + r_D) L, \end{aligned}$$

from which we can derive that this requires

$$p_i \geq p^{**} = \frac{(1 + r_D) - \lambda (\gamma (1 + r_D) + (1 - \gamma)) \pi (1 + r_L)}{((1 + r_D) (1 - \lambda \gamma) + \lambda (1 - \gamma)) \pi (1 + r_L)}. \tag{32.107}$$

If the bank does not face a bank run, depositors would obtain $(1 + r_D) L$ at the maturity of the loans, when deposits are also due to be repaid. As the interest accrues between the early repayment date and the maturity of loans and depositors can withdraw their deposits and use another bank, this needs to be discounted by $1 + r_D$ to give its value at the early repayment date. This gives us a demand of L for liquidity in banks not facing a bank run.

For banks facing a bank run, the demand is given by the total proceeds the bank can receive from selling all loans, thus $\frac{\lambda \pi (1+r_L) L (\gamma (1+r_D) + (1-\gamma))}{1+r_D} = \lambda \pi (1 + r_L) L \left(\gamma + \frac{1-\gamma}{1+r_D}\right)$.

The available funds for banks not facing a bank run at the early repayment date consists of the loans repaid early and the early funds obtained from the sold loans, hence $p_i \pi (1 + R) L + (1 - p_i) \mu_i \lambda \gamma \pi (1 + r_L) L$, where the first term also includes the borrower's profits reinvested at a return R. those banks that face a bank run sell all loans and obtain a fraction γ at the early repayment date, such that they receive $\lambda \gamma \pi (1 + r_L) L$.

Thus the net demand for funds in the market is in the case that no loans are sold, $\mu_i = 0$, and hence $r_D < r_D^*$, given by

$$\Delta = \int_0^{p^*} \lambda\pi (1 + r_L) L \left(\gamma + \frac{1-\gamma}{1+r_D}\right) dF(p_i) + \int_{p^*}^1 L dF(p_i) - \int_0^{p^*} \lambda\gamma\pi (1 + r_L) L dF(p_i) - \int_{p^*}^1 p_i\pi (1 + R) L dF(p_i) \tag{32.108}$$

with the first two terms denoting the demand by depositors in banks facing banks runs and not facing bank runs, respectively, and the final two terms the available funds in each case. Note that for $p_i < p^*$ the bank faces a bank run and for larger values no bank run occurs. We can transform equation (32.108) into

$$\Delta = \lambda\pi (1 + r_L) L \frac{1-\gamma}{1+r_D} \int_0^{p^*} dF(p_i) + L \int_{p^*}^1 (1 - p_i\pi (1 + r_L)) dF(p_i), \tag{32.109}$$

from which we get that

$$\frac{\partial\Delta}{\partial(1+r_D)} = -\frac{\lambda(1-\gamma)\pi(1+r_L) LF(p^*)}{(1+r_D)^2} + f(p^*) \frac{\partial p^*}{\partial(1+r_D)} L \left(\frac{\lambda(1-\gamma)\pi(1+r_L)}{1+r_D} - 1 + \pi(1+R) p^*\right), \tag{32.110}$$

where

$$\frac{\partial p^*}{\partial(1+r_D)} = \frac{\pi(1+r_L) - 1}{r_D^2 \pi(1+r_L)} > 0. \tag{32.111}$$

In a competitive equilibrium, we would expected the deposit rate r_D to adjust such that no excess fund exist $\Delta = 0$. If $r_D = 0$, then from equation (32.106) we obtain that $p^* = 1$ and hence $\Delta = \lambda\pi (1 + r_L) L (1 - \gamma) > 0$, requiring an increase in the deposit rate to obtain $\Delta = 0$.

As the deposit rate increases, the value of payments to failing banks reduces as the discount rate of payments at maturity increases; this reduces the excess demand as shown in the first term of equation (32.110). The second term looks at the effect of bank failures. As the discount rate increases, more banks experience bank runs as the threshold p^* increases, which is shown in equation (32.111). Failing banks sell all loans and only obtain a fraction γ at the early date, hence require funds as the remaining payment of a fraction $1-\gamma$ of the proceeds are only paid later, as expressed by the first expression of the final term of equation (32.110), this is balanced by the early repayment of the non-failing banks.

Banks with a low fraction of early repayments, p_i, do not produce much liquidity at this stage, but absorb them by holding on to loans and not selling them as long as they do not face a bank run; such banks failing and being forced to sell their loans might actually increase the funds available and reduce excess demand. Which effect dominates will depend on the exact parameter constellation; if the second term in equation (32.110) is sufficiently positive, the excess demand for liquidity might well increase in the deposit rate r_D and hence an equilibrium becomes unattainable. A similar result is obtained if $r_D > r_D^*$ and $\mu_i = 1$. Here all loans are already sold, thus failures of banks due to a bank run do not create additional liquidity from the sale of late loans.

Thus we have established the possibility of a viscous circle in which excess demand for liquidity due to a bank run increases the deposit rate, which then reduces the value of assets of the bank due to higher discount rates, which in turn cause more bank runs as . As these banks absorb funds to repay depositors, but do not generate sufficient funds from the sale of loans, excess demand will increase further, causing even more increases in the deposit rate as banks seek liquidity.

From equation (32.107) we see that for ever increasing deposit rates we obtain the threshold for the minimum early repayment of loans becomes

$$\lim_{r_D \to +\infty} p^{**} = \frac{1 - \gamma\lambda\pi (1 + r_L)}{(1 - \gamma\lambda) \pi (1 + r_L)} \tag{32.112}$$

and if $\gamma\lambda$ is sufficiently small such that $\gamma\lambda\pi (1 + r_L) < 1$ then $0 < p^{**} < 1$ as negative values are impossible and we assume that lending is desirable as $\pi (1 + r_L) > 1$. Hence banks with a too low early repayment rate of loans would fail and we could have a failure of the entire banking system if all repayment rates are sufficiently low.

Even though individual banks are failing, the banking system overall can well be solvent. Define $\overline{p} = \int_0^1 p_i dF (p_i)$ as the average early repayment rate, then the aggregate bank profits, using equation (32.104) are given by

$$\begin{aligned} \int_0^1 \Pi_B^i dF (p_i) = {} & (1 + r_D) ((\overline{p} + \mu (1 - \overline{p}) \gamma\lambda) \pi (1 + r_L) - 1) L \\ & + (\mu (1 - \gamma) (1 - \overline{p}) \lambda \\ & + (1 - \mu) (1 - \overline{p})) \pi (1 + r_L) L, \end{aligned} \tag{32.113}$$

where $\mu_i = \mu$ for all banks as their decisions whether to sell late loans is identical for all banks due to all of them having the same threshold on the deposit rate in equation (32.105).

In order for $\int_0^1 \Pi_B^i dF (p_i) \geq 0$ to make the banking system as a whole solvent, we need from equation (32.113) that

$$1 + r_D < (1 - \overline{p}) \pi (1 + r_L) \frac{\mu\lambda (1 - \gamma) + (1 - \mu)}{1 - \pi (1 + r_L) (\overline{p} + \mu\lambda\gamma (1 - \overline{p}))}. \tag{32.114}$$

We thus need to ensure that deposit rates that are sufficiently low. A viable solution to attract any deposits requires $1 + r_D > 1$, such that we need

$$\overline{p} > \frac{1 - \pi (1 + r_L)(1 - \mu (1 - \lambda))}{\pi (1 + r_L) \mu (1 - \lambda)} \tag{32.115}$$

If early repayments are sufficiently common, then we can find a deposit rate such that the banking system is solvent as a whole is solvent and liquidity could be redistributed in the banking system to avoid bank runs.

The failure of banks is the result of banks with low early repayments facing liquidity shortages due bank runs, which they can only accommodate through the sale of loans, which however increases excess demand for funds, pushing up deposit rates, that cause the vicious cycle described above and leading to the failure of banks even though sufficient funds may be available in the banking system as a whole and could be distributed to avoid bank runs.

Some banks may face a shortage of liquidity due to insufficient loans being repaid early; this will reduce the returns that banks generate as they cannot re-invest sufficient repayments. This may cause a bank run in this bank, which it can only address by selling loans at a loss. This sale, however, requires liquidity from other banks to purchase these loans and to do so, they will increase the deposit rate in order to attract more funds to their bank. But this increased deposit rate will increase the liability to depositors, requiring even higher early repayments by borrowers at other banks to avoid the bank run spreading. It is thus the fire sale of assets by a bank facing a bank run that absorbs the liquidity of other banks, who then may be subjected to a bank run themselves due to them not holding sufficient liquidity. We have therefore that the bank run in one bank necessitating the sale of loans will reduce the liquidity in other banks and the bank run can spread throughout the banking system causing a systemic banking crisis.

Reading Diamond & Rajan (2005)

32.3.2 Losses from adverse selection

Banks facing a liquidity shortage may be required to sell assets to cover such shortfalls; the most common asset banks have are loans. The risks of such loans might be well known to the bank itself, but any potential buyer of these loans will be much less well informed, giving rise to adverse selection. We will investigate how such adverse selection imposes losses on banks, that might cause them to fail due to insolvency rather than illiquidity.

We have banks that for exogenous reasons face a withdrawal of a fraction λ of their deposits D. These banks have provided loans that can be sold into a market consisting of informed investors, uninformed investors and so-called noise traders. The loans the bank gives out are repaid with probability π_i, including interest r_L. With probability p the probability of repayment is π_H and with probability $1 - p$ it is $\pi_L < \pi_H$, such that we have $\pi_H (1 + r_L) \geq \pi_L (1 + r_L) \geq 1$. Hence loans do, on average, not cause any losses to banks, regardless of the repayment probability. Any funds obtained from the sale of loans that are not used to repay depositors can be invested into a risk-free asset which yields a return of r; at the time of selling the loans this return is uncertain with only its distribution $F(\cdot)$, where $r \geq 0$, known.

Banks know the realization of the repayment rate π_i and sell the loan at a price of P. With the amount being sold denoted by $\hat{L}$, we get the profits of banks as

$$\begin{aligned}\Pi_B = \pi_i (1 + r_L) \left(L - \hat{L}\right) + (1 + r) \left(E [P|\pi_i] \hat{L} - \lambda D\right) \\ - (1 + r_D) (D - \lambda D) .\end{aligned} \tag{32.116}$$

Where the first term denotes the return on the retained loans, the second term the return on the cash reserves raised that are not used to repay depositors and the final term denotes the repayment of the retained deposits, including interest r_D. The bank will form expectations about the price it will receive, given its knowledge of the repayment rate, π_i.

The optimal amount of loans to sell is then leading to the first order condition

$$\frac{\partial \Pi_B}{\partial \hat{L}} = -\pi_i (1 + r_L) + (1 + r) E [p|\pi] . \tag{32.117}$$

This gives rise to corner solutions depending on the sign of this expression:

$$\hat{L} = \begin{cases} L \text{ if } E [P|\pi_i] \geq \frac{\pi_i(1+r_L)}{1+r} \\ 0 \text{ if } E [P|\pi_i] < \frac{\pi_i(1+r_L)}{1+r} \end{cases} . \tag{32.118}$$

The informed trader also knows the realization of π_i and will buy or short-sell the amount $\hat{L}_I$. Informed traders have wealth W and we assume that the most they can invest is this amount and that any short position need to be fully covered by this wealth. Hence with uninvested monies yielding no return, we get their profits as

$$\Pi_I = \pi_i (1 + r_L) \hat{L}_I + \left(W - E [P|\pi_i] \hat{L}_I\right) , \tag{32.119}$$

where the first term denotes the expected return from the loans they obtained and the second term the wealth not invested into purchasing loans. Informed investors maximize their profits, which yields the following first order condition for the optimal investment amount:

$$\frac{\partial \Pi_I}{\partial \hat{L}_I} = \pi_i (1 + r_L) - E [P|\pi_i] . \tag{32.120}$$

This again gives us a corner solution, where we define $\hat{W} = \frac{W}{E[P|\pi_i]}$, hence

$$\hat{L}_I = \begin{cases} \hat{W} \text{ if } E [P|pi_i] \leq \pi_i (1 + r_L) \\ -\hat{W} \text{ if } E [P|pi_i] > \pi_i (1 + r_L) \end{cases} . \tag{32.121}$$

We thus see that if the loan is undervalued, the informed trader will take a long position and if it is overvalued, a short position. As π_H and π_L are known to be the only possibilities, it is obvious that the loan with the high repayment rate, π_H, cannot be overvalued and the loan with low repayment rate, π_L, cannot be undervalued. Hence we have as the investment amount by investors

$$\hat{L}_I = \begin{cases} \hat{W} & \text{if } \pi_i = \pi_H \\ -\hat{W} & \text{if } \pi_i = \pi_L \end{cases}. \tag{32.122}$$

From equation (32.118) we know that $\hat{L} = L$ if $1 + r \geq \frac{\pi_i(1+r)}{E[P|\pi_i]}$, thus using that the risk-free rate r has a distribution $F(\cdot)$, we get the probability that the bank sells loans with a high repayment rate as

$$\mu_H = Prob\left(\hat{L} > 0|\pi_H\right) = 1 - F\left(\frac{\pi_H(1+r_L)}{E[P|\pi_H]} - 1\right). \tag{32.123}$$

As the loan with the low repayment rate π_L cannot be undervalued, we have $E[P|\pi_L] = \pi_L(1+r_L)$, implying that this is fulfilled for any $r > 0$, hence

$$\mu_L = Prob\left(\hat{L} > 0|\pi_L\right) = 1, \tag{32.124}$$

implying that loans with low repayment rates are always sold by banks.

Using Bayes' Theorem, we can get

$$\begin{aligned} \hat{p} &= Prob\left(\pi_H|\hat{L} > 0\right) \qquad (32.125) \\ &= \frac{Prob(\pi_H)\, Prob\left(\hat{L} > 0|\pi_H\right)}{Prob(\pi_H)\, Prob\left(\hat{L} > 0|\pi_H\right) + Prob(\pi_L)\, Prob\left(\hat{L} > 0|\pi_L\right)} \\ &= \frac{p\mu_H}{p\mu_H + (1-p)\mu_L} \\ &= \frac{p\mu_H}{p\mu_H + (1-p)}. \end{aligned}$$

In addition to informed traders we also have noise traders in the market; noise traders demand loans for exogenous reasons unrelated to information or utility maximisation, such as a need to conduct investments. Suppose noise traders demand loans of the size L^* and such a demand occurs with probability η. This probability η is not known but a random variable with density $g(\cdot)$.

Other market participants are investors who are uninformed about the quality of the sold loans, thus they do not know whether the loans offered have a high or low probability of being repaid; their only source of information is the aggregate demand of the informed and noise traders, defined by

$$\xi = \hat{L}_I + \eta M. \tag{32.126}$$

With $\eta = \frac{\xi - \hat{L}_I}{M}$, we can use Bayes' Theorem to derive the probability of a loan having a high repayment rate, given the observation of the uninformed investor, as

$$\begin{aligned}\hat{\hat{p}} &= Prob\left(\pi_H|\xi\right) \\ &= \frac{\hat{p}g\left(\frac{\xi-\hat{W}}{M}\right)}{\hat{p}g\left(\frac{\xi-\hat{W}}{M}\right)+\left(1-\hat{p}\right)g\left(\frac{\xi+\hat{W}}{M}\right)},\end{aligned} \tag{32.127}$$

where we used that $\hat{L}_I = \hat{W}$ for π_H and $\hat{L}_I = -\hat{W}$ for π_L. Making the same substitution into η of equation (32.126), we can rewrite this probability as

$$\hat{\hat{p}} = \begin{cases} \frac{\hat{p}g(\eta)}{\hat{p}g(\eta)+(1-\hat{p})g\left(\eta+\frac{2\hat{W}}{M}\right)} & \text{if } \pi_i = \pi_H \\ \frac{\hat{p}g\left(\eta-\frac{2\hat{W}}{M}\right)}{\hat{p}g\left(\eta-\frac{2\hat{W}}{M}\right)+(1-\hat{p})g(\eta)} & \text{if } \pi_i = \pi_L \end{cases}. \tag{32.128}$$

Let us consider only loans with high repayment rates, $\pi_i = \pi_H$, and rather than assuming η to be given, we consider it as a random variable such that with $\hat{\hat{p}}_H$ denoting the case of $\pi_i = \pi_H$, we have

$$E\left[\hat{\hat{p}}_H\right] = \int_0^1 \hat{\hat{p}}_H g\left(\eta\right) d\eta. \tag{32.129}$$

As $g(\cdot)$ is the density of a probability, it is only defined for arguments below 1, higher values of its argument having a density of zero. Thus we need $\eta + \frac{2\hat{W}}{M} \leq 1$, or

$$\eta \leq \eta^* = 1 - \frac{2\hat{W}}{M}. \tag{32.130}$$

If $\eta > \eta^*$, then $g\left(\eta + \frac{2\hat{W}}{M}\right) = 0$ and hence $\hat{\hat{p}} = 1$. We can rewrite equation (32.129) as

$$\begin{aligned}E\left[\hat{\hat{p}}_H\right] &= \int_0^{\eta^*} \hat{\hat{p}}_H g\left(\eta\right) d\eta + \int_{\eta^*}^1 g\left(\eta\right) d\eta \\ &= 1 - F\left(\eta^*\right) + \int_0^{\eta^*} \hat{\hat{p}}_H g\left(\eta\right) d\eta.\end{aligned} \tag{32.131}$$

If we set $\eta^* = 0$, then $E\left[\hat{\hat{p}}_H\right] = 1$ as $F\left(0\right) = 0$ and the integral in the above expression vanishes. This case also implies that $\hat{W} = \frac{M}{2}$, and for $\hat{W} \geq \frac{M}{2}$ it is $E\left[\hat{\hat{p}}_H\right] = 1$ and for $\hat{W} < W^*$, thus $\eta^* > 0$, we have $E\left[\hat{\hat{p}}_H\right] < 1$.

For a fair price we need that

$$\begin{aligned}E\left[P\right] &= E\left[\hat{\hat{p}}_H\right]\pi_H\left(1+r_L\right) \\ &\quad + \left(1 - E\left[\hat{\hat{p}}_H\right]\right)\pi_L\left(1+r_L\right)\end{aligned} \tag{32.132}$$

The first term considers the loan to have a high repayment rate, which is inferred to be the case with probability $E\left[\hat{\hat{p}}_H\right]$ and the second term the alternative that the loan has a low repayment rate. Thus the price would reflect fairly the information of the uninformed investors.

If $E\left[\hat{\hat{p}}_H\right] = 1$, then $E[P] = \pi_H(1 + r_L)$ and the bank obtains the full value for their loans. If $E\left[\hat{\hat{p}}_H\right] < 1$, then $E[P] < \pi_H(1 + r_L)$ and the bank would sell loans at a loss as long as $E[P] > \frac{\pi_H(1+r_L)}{1+r}$ as implied by equation (32.118), despite the bank making a loss from this sale.

In the case that $E[P] = \pi_H(1 + r_L)$, we have from equation (32.123) that $\mu_H = 1 - F(0) = 1$ and hence from equation (32.125) that $\hat{p} = p$. Therefore only loans with high-repayment rates are offered for sale; if $E[P] < \pi_H(1 + r_L)$, then we find $\hat{p} < p$ and loans with high repayments are sold, p, alongside loans with low repayment rates, $1 - p$.

If $\hat{W} < \frac{M}{2}$, there are not sufficient informed investors in the market to allow banks to sell their loans with a high repayment rate at their full value; instead they have to offer a discount on the value of these loans to entice uninformed investors. These investors will not be willing to pay the full value of the loan due to adverse selection, given that banks are also selling a fraction $1 - p$ of loans with the low repayment rate. If they were to purchase such loans they would make a loss, which is compensated for by the profits from purchasing a loan with a high repayment rate at a discount. Having to sell loans at a discount is akin to a fire sale of the assets as bank have to sell these in order to cover their liquidity shortfall. The less the demand by informed traders, the larger this discount becomes; the lower price becomes the less willing banks will become to sell loans with high repayment rates, making them less common in the market compared to loans with low repayment rates, reducing the price the bank achieves even further. Thus, in order to cover the withdrawal of deposits, banks are forced to sell assets at a loss; the initial liquidity shortfall has been turned into a loss for the bank. The origin of this loss is the adverse selection as uninformed investors do not know the repayment rates of a loan and banks have an incentive to sell loans with low repayment rates.

We can now see how this fire sale of assets can initiate a systemic banking crisis. The initial liquidity shortage has induced losses on the bank; these losses can now spread through interbank loans, as discussed in chapter 32.1. We therefore have established another way a liquidity crisis in a single bank can cause to impose losses on other banks who do not experience a liquidity shortage: the fire sale of assets by the bank at a loss due to potential purchasers being less well informed about the qualities of the assets than the bank selling it.

Reading Dow & Han (2018)

Résumé

We have seen that a liquidity shortage can lead to losses by banks. Such losses may accrue through competition of these sold assets with deposits for funds, causing deposit rates to rise, worsening the liquidity shortage, and imposing a loss on banks. It can also be that the sale of assets is not possible at full value due to the uncertainty by purchasers about the quality of these assets, causing banks to sell these at a loss to cover an liquidity shortage in a forced sale. It is thus a that a liquidity shortage in a bank can lead to it making a loss arising from the forced sale of assets in a so-called fire sale. These losses can then spread through interbank loans to other banks as discussed in chapter 32.1 and thereby cause a systemic banking crisis.

However, the sale of assets imposing a loss on those banks selling the assets and subsequent exposure through interbank market is not the only way banks are affected. Often banks are required to value their assets using market prices as a benchmark; if due to a bank selling their assets below its value, this will be recorded as a market price and other banks holding the same or similar assets will have to write down their assets, causing them a loss. This will include the assets remaining with the bank selling the assets, increasing their losses by having to write down assets they retain. These banks' losses are incurred not because they sold assets themselves or are exposed to the loss-making bank through interbank loans, but arise from the revaluation of the assets of a bank. Such losses may then cause banks to fail and instigate a more widespread systemic crisis, even without any links between banks.

Conclusions

The failure of one bank can spread to other banks. Of course it is possible that banks runs can spread if the observance of a bank run in one bank reduces either the confidence falls in the stability of other banks or the confidence that other depositors will retain their deposits as a result of their observation. However, this is not the only way that other banks can be affected. We have seen that banks can transmit their failure through the interbank market. Banks facing a loss might not be able to repay the loans that other banks have given them, imposing losses on these banks, who of course in turn could fail for this reason; the result is that the failure of one bank has spread to other banks, even though they faced no comparable losses, their only losses arise from the interbank loans. It is not essential for a systemic banking crisis that losses from loans are actually incurring, the mere inference that such losses might be realised can be sufficient to withdraw interbank loans, similar to deposits being withdrawn and causing the failure of banks as they do not have sufficient liquidity reserves to repay the recalled interbank loans or cannot raise sufficient funds from selling assets. This would then impose a loss on the bank providing the interbank loan, who will have failed, even without the bank whose stability had been questioned was sound, bar the unwarranted withdrawal of interbank loans.

However, it is not necessarily the losses of a bank in the first instance causing such a systemic banking crisis. A bank may face a liquidity shortage and in order to increase their cash reserve sell assets in a fire sale; the result will be that banks do not realise the full value of their assets, which leads to them making a loss. This resultant loss may not enable them to repay interbank loans, imposing losses on other banks and their initial liquidity shortage has resulted in them incurring a losses throughout the banking system due to them selling their assets below value. We thus see that liquidity shortages can result in actual losses, either directly through their inability to repay loans or the losses they make when selling assets in a fire sale; these losses can then spread through interbank markets. Banks will also be affected if the low realised asset values are taken as the basis for the valuation of the bank's assets. As these asset values are below the true value of the assets, banks will make a loss, which in turn can cause a bank to fail, even if they have no direct or indirect exposure to the bank initially failing through interbank loans.

With limited liability banks can also have an incentive to obtain exposure to the same risks as other banks rather than seeking out different risks. Hence, all banks will be susceptible to the same risks and if losses due to such risks are realised, all banks will be affected, which can easily result in the failure of many banks. Had they sought exposures to different risks, such a widespread failure would be less likely. While we do not observe contagion in that the failure of one bank affects the failure of other banks, the bank's deliberate choice to select risks that are similar across banks, increases the risks of multiple banks failing, and thus increases the risks to the banking system. The connection of banks through interbank loans, the losses due to fire sales of banks with a liquidity deficit, and the exposure to common loan risks can all make a banking system vulnerable to widespread failure by banks, either as the result of one bank facing losses or a liquidity shortage, or that a bank is perceived by other banks or depositors to be in that situation.

Chapter 33

Bailout decisions

A bailout happens if a bank fails and the government, a central bank, or regulator provide funds to the bank in order that they can repay the deposits, and other liabilities such as interbank loans, they have held. While bailouts can happen in any scenario where a bank fails, it is of particular importance in the context of a systemic banking crisis. Bailouts might be used in order to prevent a wider spread of the failure of a bank, for example through bank runs by depositors or other banks through the interbank market, or by passing on losses through the interbank market directly to other banks.

In this chapter we will explore the consequences of such bailouts for systemic risk itself. Looking at the costs and benefits of bailouts from the government's perspective as well as the possibility of a bank being taken over by another bank, chapter 33.1 will investigate under which conditions a bailout is optimal and under which conditions a bank should be liquidated. Chapter 33.2 will then look at how bailouts affect the incentives to provide interbank loans and how the provision of ordinary loans is affected if banks can be bailed out. A bailout is not the only option for governments, they could also force depositors to incur losses if a bank fails, while not liquidating the bank; such a bail-in is considered in chapter 33.3. Finally, chapter 33.4 will investigate how much resources a government should be providing for bailouts; hence we do not only consider that all banks affected will be able to obtain a bailout, but that the resources of governments will be limited and not all banks can be bailed out, even if the government wanted to do this.

33.1 Optimal bailout decisions

Government need to fund bailouts and should only do so if this is a better option than to let the banks fail. We will here consider the conditions under which such a bailout is optimal. We consider that banks may fail for exogenous reasons and as a result of this failure may be liquidated with their assets, the loans, sold at a discount

λ. Alternatively they might be taken over by another bank, who will purchase these loans and the deposits, or the bank could be bailed out by the government.

Let us assume that depositors make no profits and that the deposit rate r_D is determined such that when the bank succeeds with probability π as its loan L is repaid with interest r_L, it fulfills $\pi(1+r_L)D = D$. We thus obtain the deposit rate as

$$1 + r_D = \frac{1}{\pi}. \tag{33.1}$$

If a bank fails and is liquidated, we assume that the government has to cover the repayment of the deposits from its deposit insurance. When liquidating a bank, the assets are sold at a discount and the government obtains a fraction λ of the value of the loans. From these proceeds they have to repay the deposits, $(1+r_D)D = \frac{D}{\pi}$, and we assume that there are financing costs c on the net funds the government spends, which is the deposits less the price obtained from selling the assets of the bank, which is given by $\lambda\pi(1+r_L)D$. Thus the outcome for the government is given by

$$\begin{aligned}\Pi_G^F &= (\lambda\pi(1+r_L) - (1+r_D))D - c((1+r_D)D - \lambda\pi(1+r_L)D) \\ &= (1+c)\left(\lambda\pi(1+r_L) - \frac{1}{\pi}\right)D,\end{aligned} \tag{33.2}$$

where we inserted from equation (33.1) for the deposit rate. For each of the two banks the government obtains cash from the sale of the assets and uses these to repay depositors and faces the costs of financing this repayment of deposits.

If a bank is bailed out, then no loans are sold and they remain at full value when taken over by the government. The government now has to finance the deposit repayments without obtaining any revenue from the sale of the loans, hence they obtain

$$\begin{aligned}\hat{\Pi}_G^F &= \pi((1+r_L) - (1+r_D))D - c(1+r_D)D \\ &= \left(\pi(1+r_L) - \frac{c+\pi}{\pi}\right)D.\end{aligned} \tag{33.3}$$

The case where one of the banks is bailed out and the other liquidated is giving government pay outs gives us

$$\hat{\hat{\Pi}}_G^F = \frac{1}{2}\Pi_G^{FF} + \frac{1}{2}\hat{\Pi}_G^{FF}. \tag{33.4}$$

hence it will always be dominated by the larger of both banks being liquidated or bailed out. The government liquidates the banks if it is more beneficial to do so, $\Pi_G^F > \hat{\Pi}_G^F$, which inserting from equations (33.2) and (33.3) yields

$$\lambda > \lambda^* = \frac{\pi^2(1+r_L) + (1-\pi)}{(1+c)\pi^2(1+r_L)}. \tag{33.5}$$

We thus see that a failing bank is liquidated if the discount on selling the loans is sufficiently low.

If only one bank survives and the other bank fails, the other bank can buy the bank at a price $P = \lambda\pi\,(1 + r_L)\,D$, accounting for the sale of the loan at a discount; but the buying bank will get the full value of the loans, thus their profits are

$$\begin{aligned}\Pi_B^S &= 2\pi\,((1 + r_L) - (1 + r_D))\,D - P \\ &= ((2 - \lambda)\,\pi\,(1 + r_L) - 2)\,D.\end{aligned} \tag{33.6}$$

If not purchasing the failing bank, the profits are given by

$$\hat{\Pi}_B^S = \pi\,((1 + r_L) - (1 + r_D))\,D \tag{33.7}$$

$$= (\pi\,(1 + r_L) - 1)\,D. \tag{33.8}$$

A surviving bank will purchase a failing bank if it is more profitable to do so, hence $\Pi_B^S \geq \hat{\Pi}_B^S$, which easily becomes

$$\lambda < \lambda^{**} = \frac{\pi\,(1 + r_L) - 1}{\pi\,(1 + r_L)}. \tag{33.9}$$

The sale of a bank only happens if the discount for purchasing the loans is sufficiently high. We assume here that the costs of a bailout and liquidation for the government are higher than the benefits, thus $\hat{\Pi}_G^F < 0$ and $\Pi_G^F < 0$. However, the social costs of a bank failing uncontrolled without depositors being repaid would be higher though, such that the government will consider either a controlled liquidation or a bailout.

We can now summarize the optimal decision to resolve a failing bank as in figure 33.1. We see that banks are bailed out by the government if the discount on selling loans are sufficiently small and the repayment rates of banks low, thus risks are high. In this case selling the loans is not attractive to the government and the high risks make the loans not profitable enough for other banks to purchase these loans. Once the risks of the failing bank is reduced, the bank will be purchased by other banks. Once the discount when selling the loans becomes small, liquidating these loans becomes more attractive to the government, while it becomes less attractive for other banks to purchase the failing bank as the proce they have to pay is higher.

We thus see that banks are saved from failure if the discount on selling the loans of this bank are sufficiently high such that either the government would make too much of a loss when doing so or the price is sufficiently attractive for another bank to step in. In cases where loans can be sold at a high price, and the failing bank is not too risky, the bank will be liquidated with depositors being covered. In this case, their failure may spread to other banks through interbank loans, for example, as only depositor's funds are guaranteed by the government.

Reading Acharya & Yorulmazer (2007)

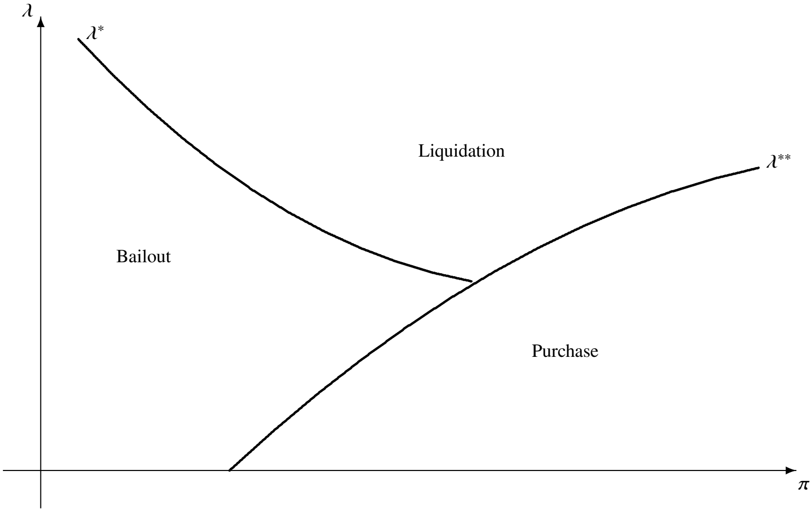


Fig. 33.1: Lending characteristics and bail-out decisions

33.2 Indirect bailouts

We assume that the government may bail out failing banks with probability p; they will only do so for some banks, but not other banks. The criteria which bank may be bailed out could be dependent on their size or their importance in the economy, for example a bank specialising in providing loans to sectors of economy that are seen as essential for future economic growth might be included in a bail out, while another bank who is one of many loan providers in other sectors of the economy will not receive such a bailout. Governments may also take into account the composition of depositors, providing a bailout to companies whose depositors are individuals, while banks with deposits mostly originating from companies and wealthy individuals might not receive a bailout. We will focus our analysis on the behaviour of banks with respect to interbank loans and how the presence of a bailout for some banks will affect the provision of such loans.

Let us initially consider a bank who does not provide nor receive any interbank loans. If we assume that loans L are repaid with probability π, including interest r_L, and such loans are fully financed by deposits on which interest r_D^i is payable, we get the profits of a bank as

$$\Pi_B^i = \pi \left((1 + r_L) - \left(1 + r_D^i\right)\right) D. \tag{33.10}$$

We here assumed that banks have limited liability and will not have to cover their losses if the loans are not repaid; thus, depositors might not be fully repaid in this instance. If the bank receives a bailout, we assume that bank owners do not benefit, they lose their entire bank holdings, but the bailout is directed at depositors who will be repaid in full, despite the bank failing. Depositors will fund the bank if

$$\pi\left(1+r_D^i\right)D+(1-\pi)\,p\left(1+r_D^i\right)D\geq D, \tag{33.11}$$

where the first term denotes the case where the bank is solvent as the loan is repaid and hence the bank repays depositors, and the second term shows the bailout of the insolvent bank, where depositors are repaid only if a bailout occurs; this repayment to depositors has to exceed the initial investment made. Giving depositors the minimal possible interest, this condition will be fulfilled with equality such that the deposit rate is given by

$$1+r_D^i=\frac{1}{\pi+(1-\pi)\,p}. \tag{33.12}$$

The higher the probability of a bailout, p, the lower the deposit rate will be as the risk of not having their deposit returns becomes smaller.

We will now introduce interbank loans and analyse whether these will be accepted and given.

Accepting interbank loans Let us consider the case where we have only two banks. Suppose banks 2 obtains an interbank loan M from bank 1 at interest rate $\hat{r}_M$. It then can provide loans to the amount of $L+M$, where L denotes the amount of lending possible without interbank loans. The profits of the bank accepting the interbank loan are given by

$$\hat{\Pi}_B^2=\pi\left((1+r_L)\,(L+M)-\left(1+\hat{r}_D^2\right)D-\left(1+\hat{r}_M^1\right)M\right). \tag{33.13}$$

Provided the loan is repaid, π, the bank obtains these repayments and uses them to repay their depositors and the interbank loan.

The incentives for depositors are unchanged, whether the bank obtains an interbank loan or not, hence the deposit rate will be identical to the above case without interbank loans, $1+\hat{r}_D^2=1+r_D^2$. We therefore from inserting equation (33.12) into equation (33.13) that the profits of the bank are given by

$$\begin{aligned}\hat{\Pi}_B^2&=\pi\left((1+r_L)-\left(1+r_D^2\right)\right)D+\pi\left((1+r_L)-\left(1+\hat{r}_M^1\right)\right)M \qquad (33.14)\\ &=\Pi_B^2+\pi\left((1+r_L)-\left(1+\hat{r}_M^1\right)\right)M.\end{aligned}$$

To obtain the second line we have inserted the deposit rate from equation (33.12) into equation (33.10), which represents the profits of the bank without interbank loans. We now easily see that as long as the loan rate exceeds the interbank loan rate, $1+r_L>1+\hat{r}_M^1$, the profits when accepting interbank loans are larger than when not accepting interbank loans, $\hat{\Pi}_B^2>\Pi_B^2$ and bank 2.

While we have shown that interbank loans are accepted by the bank, they also need to be given by the other bank. We will investigate the conditions for this to occur next.

Giving interbank loans Bank 1 would give the interbank loan M if it is more profitable for them than to provide an ordinary loan. If it is more profitable to do so, then the bank would use their entire deposits to provide interbank loans; as the profits of the bank obtaining these loans are increasing in the size of these loans as we can see from equation (33.14), bank 2 will accept such a large interbank loan and we have $M = L = D$. The profits of this bank are then given by

$$\begin{aligned}\hat{\Pi}_B^1 = {} & \pi\left(\left(1+\hat{r}_M^1\right)M - \left(1+\hat{r}_D^1\right)D\right) + (1-\pi)\,p\left(\left(1+\hat{r}_M^1\right)M\right. \\ & \left. - \left(1+\hat{r}_D^1\right)D\right),\end{aligned} \tag{33.15}$$

where the first term denotes the profits if the other bank, bank 2, does not fail, π, and repays the interbank loan. The second term denotes the repayment of the interbank loan through the bail-out if bank 2 fails, $1-\pi$, but is bailed out, p. We assume here that interbank loans are treated the same way as deposits and are included in the bailout.

For depositors we require that

$$\pi\left(1+\hat{r}_D^1\right)D + (1-\pi)\left(p\left(1+\hat{r}_D^1\right)D + (1-p)\,p\left(1+\hat{r}_D^1\right)D\right) \geq D. \tag{33.16}$$

The first term denotes the repayment of deposits if bank 2 repays the interbank loan and the final term if it fails and cannot repay the interbank loan itself and hence bank 1 cannot repay depositors due to the loss from not receiving payment on their interbank loans; the first expression covers the case if the other bank is bailed out the interbank loan is then repaid, allowing bank 1 to repay depositors. The second expression represents the case of the bank 2 not being bailed out, but bank 1 is bailed out. Note that if bank 2 is bailed out, bank 1 cannot fail and hence we do not need to consider the possible bailout of bank 1. the minimal deposit rate fulfilling this requirement is then given by

$$1+\hat{r}_D^1 = \frac{1}{\pi + (1-\pi)(2-p)\,p} \leq 1 + r_D^1. \tag{33.17}$$

The benefits of the bank providing interbank loans instead of ordinary loans become apparent here. The potential bailout of the interbank loan through the other bank, reduces the risk of depositors as now they do not only have to consider the possibility of the bank they hold their deposits at being bailed out, but in addition the other bank being bailed out would also ensure they are repaid their deposits.

Inserting equation (33.17) into equation (33.15) we get the profits of the bank providing the interbank loan as

$$\hat{\Pi}_B^1 = (\pi + (1-\pi)\,p)\left(\left(1+\hat{r}_M^1\right) - \frac{1}{\pi + (1-\pi)(2-p)\,p}\right)D. \tag{33.18}$$

An interbank loan will only be granted if it is more profitable than providing ordinary loans, thus we require $\hat{\Pi}_B^1 \leq \Pi_B^1$, which solves for

$$1+\hat{r}_M^1 \leq 1+r_M^* = \frac{\pi}{\pi+(1-\pi)\,p}\left((1+r_L) - \frac{1}{\pi+(1-\pi)\,p}\right) - \frac{1}{\pi+(1-\pi)\,(2-p)\,p} \tag{33.19}$$

Thus if $1+r_L \geq 1+\hat{r}_M^1 \geq 1+r_M^*$, an interbank loan can be generated that increases profits for both the bank providing this interbank loan and the bank receiving it; we can show that for $p>0$ the inequalities are strict.

The reason for banks preferring to provide and accept interbank loans is that the lower deposit rate that emerges for bank 1 as a result of depositors at this bank having the protection of the bailout of its own bank and the bailout of the other bank, allows banks to accept returns on their interbank loans that are below that of ordinary loans, where they would not be afforded such a double protection, but depositors would only have the protection from the bailout of their own bank and thus require a higher deposit rate.

Having thus established that banks would be happy to provide and receive interbank loans from each other, we will explore if the receiving bank will actually provide an ordinary loan or would want to provide another interbank loan instead.

Bilateral interbank loans Bank 2 has accepted the interbank loan provided by bank 1 if $1+r_L \geq 1+\hat{r}_M^1 \geq 1+r_M^*$. Rather than investing the full proceeds into the ordinary loan, it could now seek to return the interbank loan to bank 1.

Let us denote by ρ the probability that neither bank fails. The probability of the other bank only failing, is then given by $\pi-\rho$, the probability of this bank not failing, less the probability neither fails, giving us the probability of only the other bank failing. Consequently, $(1-\pi)-(\pi-\rho)=1-2\pi+\rho$ will denote the probability of both banks failing; it is this bank failing, $1-\pi$, less exactly the other bank only failing, $\pi-\rho$, gives that both banks fail.

We then have the profits of the bank given by

$$\begin{aligned}\hat{\hat{\Pi}}_B^2 = {} & \rho\left((1+r_L)\,L + \left(1+\hat{\hat{r}}_M^1\right)M - \left(1+\hat{\hat{r}}_D^2\right)D \right.\\ & \left. - \left(1+\hat{\hat{r}}_M^2\right)M\right)\\ & + (\pi-\rho)\left((1+r_L)\,L + p\left(1+\hat{\hat{r}}_M^1\right)M\right.\\ & + (1-p)\,\alpha_M\left(1+\hat{\hat{r}}_M^1\right)M - \left(1+\hat{\hat{r}}_D^2\right)D\\ & \left. - \left(1+\hat{\hat{r}}_M^2\right)M\right),\end{aligned} \tag{33.20}$$

where $\alpha_M = \frac{\left(1+\hat{\hat{r}}_M^1\right)M}{\left(1+\hat{\hat{r}}D^1\right)D+\left(1+\hat{\hat{r}}_M^1\right)M}$ denotes the fraction of assets that are allocated to interbank loans if a bank fails and is not bailed out. The first term denotes the profits if both banks are not failing; the bank will obtain its loan and the interbank loan given to the other bank, from which it will pay its own interbank loan and depositors. The second term represents the profits if the other bank is failing. Here the bank collects its own loan, receives the interbank loan if the other bank is bailed out, its fraction of the assets, which will be its own interbank loan only as the loan has not been repaid, if it is not bailed out; in addition the interbank loan and deposits are repaid.

For depositors we have with $\alpha_D = 1 - \alpha_M$ denoting the fraction of assets that are allocated to depositors if a banks fails and is not bailed out, that

$$\begin{aligned}D \leq \rho \left(1 + \hat{\hat{r}}_D^2\right) D + (\pi - \rho) \left(1 + \hat{\hat{r}}_D^2\right) D \\ + (\pi - \rho) \left(p \left(1 + \hat{\hat{r}}_D^2\right) D + (1 - p)\, \alpha_D \left(1 + \hat{\hat{r}}_M^1\right) M\right) \\ + \frac{1}{2} (1 - 2\pi + \rho) \left(p \left(1 + \hat{\hat{r}}_D^2\right) D + (1 - p)\, p\alpha_D \left(1 + \hat{\hat{r}}_M\right) M\right. \\ + p^2 \left(1 + \hat{\hat{r}}_D^2\right) D + (1 - p)\, p \left(1 + \hat{\hat{r}}_D^2\right) D \\ \left. + p\, (1 - p)\, \alpha_D \left(1 + \hat{\hat{r}}_M^1\right) M\right)\end{aligned} \tag{33.21}$$

The first term gives the return if both banks do not fail and the second term if the other bank only fails. The third term denotes this bank failing, being bailed out or not being bailed out and sharing the proceeds. The final term denotes the case that both banks fail. Here the order of bailout becomes relevant and the first two terms cover the case of this bank being bailed out first, while the final 3 expressions are if this bank is second in the bailout. Then if both banks are bailed out, the deposits are returned in full, the same is the case if this bank only is bailed out. If only the other bank is bailed out, the depositors share the proceeds. This condition now solves for

$$\begin{aligned}1 + \hat{\hat{r}}_D^2 &= \left(1 + r_D^2\right) \\ &\quad - (1 - p)\, \frac{(\pi - \rho) + (1 - 2\pi + \rho)\, p}{\pi + (1 - \pi)\, p} \alpha_D \left(1 + \hat{\hat{r}}_M^1\right) \\ &< 1 + r_D^2.\end{aligned} \tag{33.22}$$

In this case we again seeing the benefits of additional bailouts in the other bank as this reduces deposit rates.

We can now derive that

$$\frac{\partial \alpha_D}{\partial \rho} = \frac{\partial \left(1 + \hat{\hat{r}}_D^2\right)}{\partial \rho} \frac{1 + \hat{\hat{r}}_M^2}{\left(\left(1 + \hat{\hat{r}}_D^2\right) + \left(1 + \hat{\hat{r}}_M^2\right)\right)^2} \tag{33.23}$$

and

$$\begin{aligned}\frac{\partial\left(1+\hat{\hat{r}}_D^2\right)}{\partial\rho} &= \frac{(1-p)^2}{\pi+(1-\pi)\,p}\alpha_D\left(1+\hat{\hat{r}}_M^1\right) \\ &\quad -(1-p)\,\frac{(\pi-\rho)+(1-2\pi+\rho)\,p}{\pi+(1-\pi)\,p}\frac{\partial\alpha_D}{\partial\rho}\left(1+\hat{\hat{r}}_M^1\right),\end{aligned} \tag{33.24}$$

which using equation (33.23) solves for

$$\begin{aligned}\frac{\partial\left(1+\hat{\hat{r}}_D^2\right)}{\partial\rho} &= \frac{(1-p)^2}{\pi+(1-\pi)\,p}\alpha_D\left(1+\hat{\hat{r}}_D^1\right) \\ &\quad \times\left(1+(1-p)\,\frac{(\pi-\rho)+(1-2\pi+\rho)\,p}{\pi+(1-\pi)\,p}\right. \\ &\quad \left.\frac{\left(1+\hat{\hat{r}}_M^1\right)^2}{\left(\left(1+\hat{\hat{r}}_D^1\right)+\left(1+\hat{\hat{r}}_M^1\right)\right)^2}\right)^{-1} > 0\end{aligned} \tag{33.25}$$

Using this result we can easily show that the profits of the bank are increasing in the probability of both banks failing,

$$\frac{\hat{\hat{\Pi}}_B^2}{\partial\rho} > 0. \tag{33.26}$$

This implies that bank would like to choose the probability of them not failing together, ρ, to be as large as possible, thus $\rho=\pi$, which is the probability of a bank not failing. In this case the banks invest into identical loans as $\pi-\rho=0$ and hence the chance of only one of the banks failing is zero. Therefore it is optimal for banks to not diversify their loan provision across banks, but as they both fail or are do not fail, they maximize the benefits of the bailout.

In this case we have we have the profits of the bank given from equation (33.20) as

$$\hat{\hat{\Pi}}_B^2 = \pi\left((1+r_L)-\left(1+\hat{\hat{r}}_D^2\right)\right) > \Pi_B^2, \tag{33.27}$$

where the inequality emerges due to $1+\hat{\hat{r}}_D^2 < 1+r_D^2$ and comparison with equation (33.10). As both banks are identical and give interbank loans to each other, they will charge the same interbank rates and these two terms cancel each other out. Furthermore if $1+\hat{\hat{r}}_M^1 = 1+r_L$, we easily can verify from equations (33.10) and (33.14) that $\hat{\Pi}_B^2 = \Pi_B^2$ and $\hat{\hat{\Pi}}_B^2 > \hat{\Pi}_B^2$.

A decrease in the interbank loan rate, $1+\hat{\hat{r}}_M^1$, will increase deposit rates, $1+\hat{\hat{r}}_D^2$, as we can from equation(33.22) and the fact that

$$\frac{\partial\alpha_D\left(1+\hat{\hat{r}}_M^1\right)}{\partial\left(1+\hat{\hat{r}}\right)} = \alpha_D^2 > 0 \tag{33.28}$$

As the deposit rate $1+\hat{\hat{r}}_D^2$ increases, the bank profits in equation (33.27) will decrease as we can easily observe. Thus there will be some interbank loan rate $1+r_M^{**}$ such that for all $1+\hat{\hat{r}}_M^1 > 1+r_M^{**}$ we have $\hat{\hat{\Pi}}_B^2 > \hat{\Pi}_B^2$. Hence, we will find that in this case bank 2 wants to return an interbank loan to bank 1 as this allows the bank to benefit from the bailout of the other bank, such that both banks are partially covered by the bailout of the other bank. As banks can set their interbank rates accordingly, they can ensure the conditions are fulfilled, namely the interbank loan rate much fulfill the condition that

$$1 + r_L > 1 + r_M > \max\left\{1 + r_M^*; 1 + r_M^{**}\right\}. \tag{33.29}$$

Of course, banks 1 needs to accept this reciprocated interbank loan from bak 2. Using the relationships between the profits of the banks in the various situations, namely

$$\begin{aligned} &\hat{\hat{\Pi}}_B^2 > \hat{\Pi}_B^2 > \Pi_B^2 \\ &\hat{\Pi}_B^1 > \Pi_B^1, \end{aligned} \tag{33.30}$$

and the symmetry of banks, we can see from figure 33.2 that both banks providing interbank loans is the equilibrium of this strategic game between banks. The entries represent the profits of each bank for all the possible combinations of providing and accepting the reciprocated interbank loan or not having such a loan. The arrow indicate which profits are higher and using iterated strict dominance we easily see that the only equilibrium is for the reciprocated interbank loan to be provided and accepted.

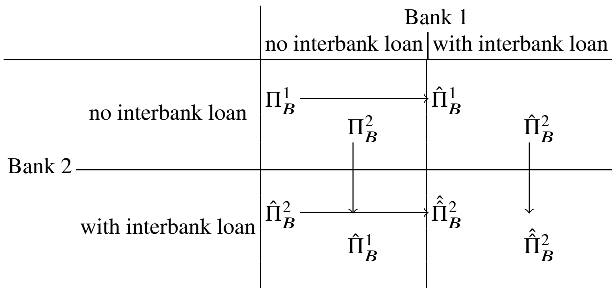


Fig. 33.2: Strategic game of interbank loan provision

Summary We have established that a possible bailout for a failing bank provides incentives for banks to increase the coverage of this protection through reciprocative interbank loans. Banks will rather provide interbank loans than ordinary loans; the protection against default on this loan through the bailout of the other bank, which is not given for ordinary loans, will reduce the risk of failure for their bank, and hence reduce the deposit rate they have to pay due to the lower risk of depositors making a loss from the failure of the bank. While in the model all funds were invested into interbank loans, in reality there needs to be investment into ordinary loans or no external revenue is generated to obtain profits and no banks in the traditional sense would exist. Banks seek protection from failure not only through a bailout of their own bank in case they are failing, but they reduce the risk of failure in the first place by relying on the other bank, which they have granted interbank loans to, being bailed out. This way, banks have two chances of benefitting from a bailout, their own and that of the other bank.

In addition, the banks will seek to maximize the correlation of their investments. That is as to maximize the value of the bailout they can obtain, as in this case both banks will have a probability p of obtaining a bailout. If both banks fail, either will receive a bailout with probability $1-(1-p)^2$, which can easily be shown to be larger than p, the probability of receiving a bailout if only one bank fails. The consequence is that compared to banks providing diversified loans such that the probability of both banks failing is $(1-\pi)^2$ would be lower than if they granted the same loans with the probability of both banks failing given by $\rho = \pi$, which is larger if $\pi > \frac{3-\sqrt{5}}{2} \approx 0.38$. We can thus conclude that the possibility of a bailout increases the likelihood of a systemic banking crisis where all banks fail due to investing into the same loan portfolio.

Reading Eisert & Eufinger (2019)

33.3 Bail-in

A bank that is not able to repay all its depositors would fail as it has to sell its assets, the loans it has provided, at a loss in a fire sale to liquidate the bank. These losses would reduce the amount it is able to pay its depositors. Of course, to prevent losses to depositors banks could be bailed out by the government, but an alternative is to impose some losses on depositors as well in order to limit the losses from such asset sales. If depositors are required to accept a loss to ensure the bank is not liquidated, this is referred to as a bail-in. We will here discuss under which condition a bail-in is preferable to a bailout.

We consider an economy in which a bank provides loans over two time periods, but if these loans are nor repaid at the end of the first time period, the bank can be bailed out by the government, be forced to bail in depositors, or be liquidated. Banks are subject to moral hazard in that the repayment rate of their loans is given as π_L, but when putting in additional effort at cost cL this can increase to $\pi_H > \pi_L$. The decision to exert effort to monitor loans is made in each time period separately.

If in the second time period the bank is to monitor the loans, we need for a bank that has previously been successful, and thus not failed, that the profits when exerting effort and obtaining the higher repayment rate exceed the profits if they do not do so and obtain a low repayment rate. Hence we require that

$$\begin{aligned} &\pi_H\left(\left(1+r_L\right)L-\left(1+r_D^2\right)D\right)-cL \\ &\geq \pi_L\left(\left(1+r_L\right)L-\left(1+r_D^2\right)D\right) \end{aligned} \tag{33.31}$$

where r_L denotes the loan rate and r_D^2 the deposit rate in the second time period, L the loans and D deposits. We assume banks hold no cash or equity and hence loans are fully financed by deposits such that $D = L$. We can solve the condition in equation (33.31) for

$$1+r_D^2 \leq \left(1+r_L\right)-\frac{c}{\Delta\pi}, \tag{33.32}$$

with $\Delta\pi = \pi_H - \pi_L$ representing the difference in the repayment rates of loans with and without monitoring. A sufficiently low deposit rate ensures that the profits with monitoring exceed those without. As we assume that after the second time period there is no bailout or bail-in, depositors will provide deposits as long as the repayments they receive exceed their initial investment, $\pi_H\left(1+r_D^2\right)D \leq D$. This becomes

$$1+r_D^2=\frac{1}{\pi_H}, \tag{33.33}$$

if we assume depositors are offered the minimum interest rate that will entice them to provide deposits.

If the bank has failed after the first period and been bailed out or depositors being bailed in, we assume that the government only allows the bank owners to retain a fraction α of the bank, the remainder is taken by the government who provides the bailout or the depositors who are bailed in. In this case, in order to induce banks to monitor their loans, we require that

$$\begin{aligned} &\alpha\pi_H\left(\left(1+r_L\right)L-\left(1+\hat{r}_D^2\right)D\right)-cL \\ &\leq \alpha\pi_L\left(\left(1+r_L\right)L-\left(1+\hat{r}_D^2\right)D\right), \end{aligned} \tag{33.34}$$

which easily solves for

$$1+\hat{r}_D^2 \leq \left(1+r_D\right)-\frac{c}{\alpha\Delta\pi} < 1+r_D^2. \tag{33.35}$$

This deposit rate is lower than the deposit rate if no bailout or bail-in had occurred as the existing bank owners have to fully cover the costs of monitoring, rather than being shared out among all owners equally; implicitly we assume here that the government or depositors who now own a part of the bank do not cover the losses, this would be the case for owner-managers of the bank, for example. As the incentives for depositors to retain their deposits are unchanged, we can insert for $1 + r_D^2$ from equation (33.33) and solve this constraint for

$$\alpha \geq \alpha^* = \frac{\pi_H}{\pi_H\left(1+r_L\right)-1}\frac{c}{\Delta\pi}. \tag{33.36}$$

Hence if a bank is bailed out or depositors bailed in, it must be provided with a minimum share α^* of the bank after the rescue of the bank to ensure that banks monitor loans and depositors retain funds. If we assume here that $\pi_L (1 + r_L) < 1 < \pi_H (1 + r_L)$ such that monitoring is essential to ensure depositors are fully repaid after the second time period.

Bail-in decision Let us now consider the decision to bail-in depositors after the bank has failed in the first time period. In this case depositors convert their deposits in to equity of the bank, and the government might add funds of T to support this bail-in, also in return for a share of the bank's equity. The bank owners retain a fraction α and the remaining fraction $1 - \alpha$ is shared between depositors obtaining $(1 - \alpha) \beta$ and the government obtaining $(1 - \alpha)(1 - \beta)$; the government funds are to be financed with interest r. The total welfare of bailing in the bank is then given by

$$\begin{aligned} \Pi_W &= \pi_H \left((1 + r_L) L - \left(1 + r_D^2\right) D \right) - rT \\ &= (\pi_H (1 + r_L) - 1) D - rT, \end{aligned} \tag{33.37}$$

where the second equality is obtained using that loans are fully funded by deposits and we insert for the deposit rate from equation (33.33). We see that the total welfare is maximized if government funds are minimal, $T = 0$, and as the government does not make any contribution, it implies $\alpha = 1$, such that the depositors will hold a share of $1 - \beta$ of the bank. With the liquidation of the bank yielding no payment, we see that a bail-in is preferred to liquidation as long as $\pi_H (1 + r_L) > 1$, which we had assumed to be the case.

Knowing that a failing bank will be bailed in, we can now turn to the decision by depositors in the first time period. In order for banks to monitor loans, we need that

$$\begin{aligned} &\pi_H \left((1 + r_L) L - \left(1 + r_D^1\right) D + \pi_H \left((1 + r_L) L - \left(1 + r_D^2\right) D \right) \right) \\ &\quad + \beta (1 - \pi_H) \pi_H \left((1 + r_L) L - \left(1 + r_D^2\right) D \right) - 2cL \\ &\geq \pi_L \left((1 + r_L) - \left(1 + r_D^1\right) D + \pi_H \left((1 + r_L) L - \left(1 - r_D^2\right) D \right) \right) \\ = &\quad + \beta (1 - \pi_L) \pi_H \left((1 + r_L) - \left(1 + r_D^2\right) D \right) - cL, \end{aligned} \tag{33.38}$$

where the first term denotes the case the first time period is successful and the bank continues to operate in the second time period. The second term denotes the case of failure in the first time period, where the bank is bailed in and bank owners retain a fraction β of bank's second period profits. Note that in this case the first period will not provide profits as the bank failed and the bank monitors loans in the second time period due to the previous analysis. .

We can now solve equation (33.38) and insert for the deposit rate from equation (33.33) and obtain the condition for banks to monitor in the first time period as

$$1 + r_D^1 \leq 1 + \hat{r}_D^1 = (1 + r_L) + (1 - \beta)\, \pi_H \left((1 + r_L) - \frac{1}{\pi_H} \right) - \frac{c}{\Delta\pi}. \tag{33.39}$$

Depositors are repaid with probability π_H and in the case of failure obtain a fraction $1 - \beta$ of the future profits of the bank. Hence for deposits to be provided in the first time period, we require that

$$\pi_H \left(1 + r_D^1\right) D + (1 - \pi_H)(1 - \beta) \left(\pi_H (1 + r_L) - \frac{1}{\pi_H} \right) D \geq D, \tag{33.40}$$

solving for

$$1 + r_D^1 = \frac{1}{\pi_H} - (1 - \pi_H)(1 - \alpha) \left((1 + r_L) - \frac{1}{\pi_H} \right) \tag{33.41}$$

if we again assume that banks pay the lowest possible deposit rate to ensure deposits are made. Inserting this expression into equation (33.39), we get

$$\beta \leq \beta^{**} = \frac{2\left((1 + r_L) - \frac{1}{\pi_H} \right) - \frac{c}{\Delta\pi}}{(1 + r_L) - \frac{1}{\pi_H}}. \tag{33.42}$$

In order to obtain a viable solution we need that the fraction the current bank owners retain, α and β, respectively. The fraction retained must exceed α^* defined in equation (33.36) such that banks monitor loans and obtain deposits in the second time period, and it much be below β^{**} defined in equation (33.42) such that a bail in is feasible while obtaining deposits in the first time period. Hence a viable solution can only be found if $\alpha^* \leq\leq \beta^{**}$, which ensures that banks monitor loans in both time periods and a bail-in can be conducted. Inserting from equations (33.36) and (33.42) we obtain this condition to require that

$$c \leq c^* = \left((1 + r_L) - \frac{1}{\pi_H} \right) \Delta\pi \tag{33.43}$$

and we see that for sufficiently low monitoring costs, bail-ins are feasible.

Having established the condition under which a bail-in is feasible, we will now consider when a bailout can be conducted.

Bailout decision In a full bail-out, the government transfers $T = \left(1 + r_D^1\right) D$ to the bank to repay its depositors and obtains a fraction $1 - \alpha$ of the future bank profits. As depositors see deposits as risk-free in this case, we easily obtain that

$$1 + r_D^1 = 1. \tag{33.44}$$

The incentives for banks to monitor their loans in the first time period are thus given by

$$\pi_H \left((1+r_L) L - \left(1+r_D^1\right) D + \pi_H \left((1+r_L) L - \left(1+r_D^2\right) D \right) \right) \tag{33.45}$$
$$+ (1-\pi_H) \pi_H \alpha \left((1+r_L) L - \left(1+r_D^2\right) D \right) - 2cL$$
$$\geq \pi_L \left((1+r_L) L - \left(1+r_D^1\right) D + \pi_H \left((1+r_L) L - \left(1+r_D^2\right) D \right) \right)$$
$$+ (1-\pi_L) \pi_H \alpha \left((1+r_L) L - \left(1+r_D^2\right) D \right) - cL.$$

Using that $1 + r_D^1 = 1$ and $1 + r_D^2 = \frac{1}{\pi_H}$, this requirement can be solved for

$$\alpha \leq \alpha^{**} - \frac{(1+\pi_H)(1+r_L) - 2 - \frac{c}{\Delta\pi}}{\pi_H (1+r_L) - 1}. \tag{33.46}$$

Similar to the case of a bail-in, we require $\alpha^* \leq \alpha^{**}$ to ensure that monitoring can occur in both time periods. This implies that we have to ensure that

$$c \leq \hat{c}^{**} = \frac{(1+\pi_H)(1+r_L) - 2}{2\pi_H} \Delta\pi. \tag{33.47}$$

We can easily see that $c^* < \hat{c}^{**}$ when comparing the conditions in equations (33.43) and (33.47), we see that bailouts are possible under less restrictive conditions than bail-ins. As bail-outs are costly to governments, they would prefer bail-ins and due to the associated financing costs rT, total welfare would be improved, too, as we have seen from equation (33.37).

An alternative to a bail-in or a bailout would be to liquidate the bank, which we will consider next.

Liquidation If a bank is liquidated, then the incentives to monitor in the first time period are such that

$$\pi_H \left((1+r_L) L - \left(1+r_D^1\right) D + \pi_H \left((1+r_L) L - \left(1+r_D^2\right) D \right) \right) - 2cL \tag{33.48}$$
$$\geq \pi_L \left((1+r_L) L - \left(1+r_D^1\right) D + \pi_H \left((1+r_L) L - (1+r_D) D \right) \right) - cL,$$

which when using that $1 + r_D^2 = \frac{1}{\pi_H}$, becomes

$$1 + r_D^1 \leq (1+\pi_H)(1+r_L) - 1 - \frac{c}{\Delta\pi}. \tag{33.49}$$

The incentives for depositors are in this case that they require that their expected payments if the bank does not fail exceeds their initial investment, $\pi_H \left(1 + r_D^1\right) D \geq D$. Thus, when again assuming that banks only pay the minimum deposit rate required, we obtain the deposit rate as

$$1 + r_D^1 = \frac{1}{\pi_H}. \tag{33.50}$$

Inserting this expression into equation (33.49), we obtain that

$$c \leq c^{**} = (1 + \pi_H) \left((1 + r_L) - \frac{1}{\pi_H} \right) \Delta\pi \tag{33.51}$$

If monitoring costs are sufficiently low, liquidation is feasible while deposits are made with the bank. We find that $c^* < c^{**}$ and, if $\pi_H > \frac{1}{2}$, then $\hat{c}^{**} < c^{**}$ as we can see from comparing equations (33.43), (33.47), and (33.51) and liquidations are more easily feasible than bail-ins or bailouts. Rather than focussing on bailouts and bail-ins exclusively, we might combine these two possibilities.

The optimal policy mix In the general case, the government pays the bank the amount T and obtains a fraction $(1 - \alpha)(1 - \beta)$ of the profits of the bank, while depositors get a fraction $(1 - \alpha)\beta$ of these profits, while bank owners obtain a fraction α. In contrast to a pure bailout, depositors are only repaid a fraction of the future profits rather than receiving their full deposits returned as we had assume din the case of a full bailout. The bailout consisted of repaying depositors, while here the government injects an amount T into the bank.

Depositors are having their deposits returned if the bank does not fail and otherwise obtains their share of the profits of the bank, enhanced by the government payment T. Hence in order to provide deposits we require that

$$\pi_H \left(1 + r_D^1\right) D + (1 - \pi_H) \left(T + (1 - \alpha)\beta \left((1 + r_L) L - \left(1 + r_D^2\right) D\right)\right) \geq D, \tag{33.52}$$

hence using that $1 + r_D^2 = \frac{1}{\pi_H}$ we obtain

$$1 + r_D^1 = \frac{1}{\pi_H} - \frac{1 - \pi_H}{\pi_H} T - (1 - \pi_H)(1 - \alpha)\beta \left((1 + r_L) - \frac{1}{\pi_H}\right). \tag{33.53}$$

Solving equation (33.45) for $1 + r_D^1$, we get the requirement for banks monitoring in time period 1 as

$$1 + r_D^1 \leq (1 + r_L) + (1 - \alpha)\pi_H \left((1 + r_L) - \frac{1}{\pi_H}\right) - \frac{}{\Delta\pi}. \tag{33.54}$$

Let us choose a policy that maximizes welfare, which from equation (33.37) is given as $\Pi_W = (\pi_H (1 + r_L) - 1) D - rT$. We see that T reduces welfare and should thus be minimized. We further see from equation (33.53) that β reduces the deposit rate of $1 + r_D^1$, and hence to make equation (33.54) the least restrictive, we would choose $\beta = 1$ and depositors obtain the highest possible compensation. Furthermore, as α increases equation (33.53) and decreases equation (33.54), we choose $\alpha = \alpha^*$ to be least restrictive. Then, inserting equation (33.53) into equation (33.54), we obtain that at equality of this constraint we have

$$T^* = \frac{2\pi_H}{(1 - \pi_H)\Delta\pi}(c - c^*) = \frac{2\pi_H}{(1 - \pi_H)\Delta\pi}\left(c - (1 + r_L) + \frac{1}{\pi_H}\right), \tag{33.55}$$

using $c^* = \left((1 + r_L) - \frac{1}{\pi_H}\right) \Delta\pi$ from equation (33.43)

This policy mix is feasible as long as $\Pi_W \geq 0$, or

$$T^* \leq \frac{\pi_H (1 + r_L) - 1}{r} D. \tag{33.56}$$

Inserting from equation (33.55) for the optimal government payment, we can determine a π_H^*, such that for $\pi_H \leq \pi_H^*$ we have $\Pi_W \geq 0$.

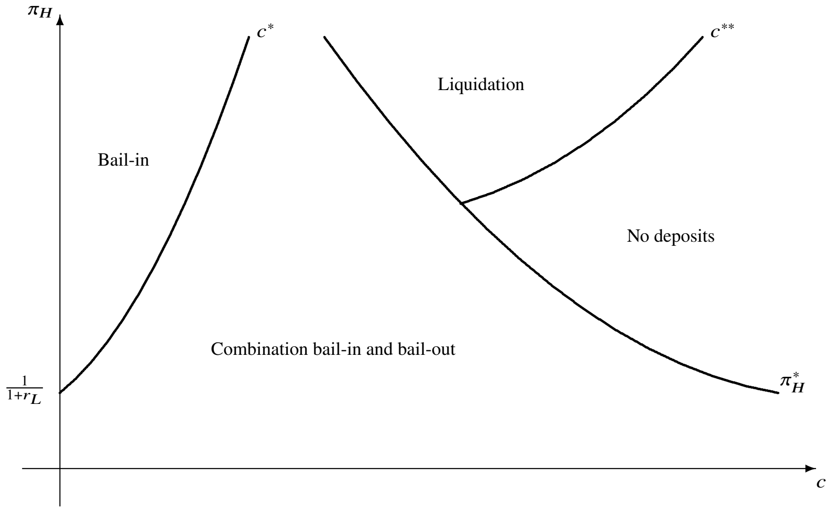


Fig. 33.3: Optimal bail-in decisions

Summary Figure 33.3 illustrates the optimal decisions about bailouts and bail-ins. We see that for small monitoring costs it is generally optimal to bail in depositors, while for high success rates and high monitoring costs, banks that fail are liquidated, and if success rates are sufficiently low, no deposits are made in the first instance. In the intermediate region a mix of bailouts and bail-ins are optimal.

Bail-ins are more difficult to achieve as they increase the costs to banks and hence lower their profits as deposit rates are higher due to the potential loss of deposits. In contrast to that, a bailout by the government injects additional resources T into the banking system that makes such a combination of bail-in and bailout more easily feasible. If neither bailouts nor bail-ins are feasible, then the bank would be liquidated. In this case, too risky banks would not receive any deposits as depositors would never be able to break even while the bank is still profitable. It is thus that as long as banks have sufficient incentives to monitor loans and consequently reduce the risk of failure, requiring low monitoring costs, most bank failures can be avoided as a bail-in or a combination of a bail-in with a bailout in the form of government funds is feasible.

Reading Pandolfi (2021)

33.4 Optimal bailout resources

It is often assumed that bailouts will happen if it is beneficial for the government or social welfare. However, governments will have limited resources to to finance such a bailout and hence if banks fail, they might not be able to bail out all banks. Banks knowing that a bailout might be offered, will take this into account when deciding on the risks they are taking. If bailouts are very likely, then banks might increase their risks and the resources required for governments will be high, making is difficult to bail out all failing banks; on the other hand if bailouts are unlikely banks might be pursuing less risky strategies and thus fewer will fail, making a bailout more likely as resources are used less. We will therefore investigate the optimal amount of resources government should be providing for bailouts.

Let us assume that an infinite number of atomistic banks in perfect competition provide loans that are repaid with probability π, including interest r_L; these repayments are independent of each other across banks. Banks are fully financed by deposits, which are paid interest r_D and hence with limited liability, their expected returns from lending the amount of L are $\pi (r_L - r_D) L$. The risk banks take, measured by the success rate π, can be affected by effort and in order to obtain a success rate π, the bank faces costs of $\frac{1}{2} c\pi^2 L$, that are always borne by the bank. If the bank fails, which happens with probability $1 - \pi$, they may be bailed out with probability p, in which case the income from the loans, $(1 + r_L) L$ is restored to the bank at no costs to them. Hence bank profits are given by

$$\Pi_B = \pi (r_L - r_D) L + (1 - \pi) p (r_L - r_D) L - \frac{1}{2} c\pi^2 L. \tag{33.57}$$

We now propose that the government bails out banks only if the fraction of banks failing, $1-\pi$, exceeds a threshold minimum $\underline{\lambda}$; such a threshold may be justified in that a small number of banks failing will not impose large costs on the economy and the likelihood of any of their losses spreading to other banks is low. Once this threshold is surpassed, the threat of more banks failing and causing a systemic banking crisis increases and thus failing banks are bailed out. However, there are limited resources available and at most a fraction of $\overline{\lambda}$ can be bailed out fully. If even more banks are failing, not all will be bailed out and only a fraction $\frac{\overline{\lambda}}{1-\pi}$ can be rescued. It is therefore that we interpret $\overline{\lambda}$ as the resources available for a bailout as this determines the fraction of banks that can be rescued. Hence we have the probability of such a bailout given as

$$p(\pi) = \begin{cases} 0 & \text{if } \quad 1 - \pi < \underline{\lambda} \\ 1 & \text{if } \underline{\lambda} \leq 1 - \pi \leq \overline{\lambda} \\ \frac{\overline{\lambda}}{1-\pi} & \text{if } \quad 1 - \pi > \overline{\lambda} \end{cases}, \tag{33.58}$$

where we assume that all failing banks are treated alike.

The government faces total costs of $C_F L$ if banks fail and are not bailed out, consisting of covering the costs to depositors and wider macroeconomic costs. Bailing out a bank will cost $C_B L$, being the costs of funding the loans $(1 + r_L) L$ and the associated financing costs. We assume that $C_F > C_B > 1$. Then, if banks are not bailed out, the costs to government are $(1 - \pi) C_F L$ and if they are fully bailed out, they are $(1 - \pi) C_B L$. If not all banks can be bailed out, but only a fraction $\overline{\lambda}$ of all banks, we have costs of $\overline{\lambda} C_B L + \left(1 - \pi - \overline{\lambda}\right) C_F L$, that is the costs of bailing out a fraction $\overline{\lambda}$ of banks and letting the remaining banks fail. Thus the payments by the government are given by

$$\Pi_G = \begin{cases} -(1-\pi) C_F L & \text{if } \quad 1-\pi < \underline{\lambda} \\ -(1-\pi) C_B L & \text{if } \underline{\lambda} \leq 1-\pi \leq \overline{\lambda} \\ -\overline{\lambda} C_B L - \left(1-\pi-\overline{\lambda}\right) C_F L & \text{if } \quad 1-\pi > \overline{\lambda} \end{cases}. \tag{33.59}$$

Let us now consider the optimal risk-taking by banks, π, as well as the optimal amount of resources the government puts in, $\overline{\lambda}$. As banks are atomistic, the decision of one bank, does not affect the fraction of banks failing overall. Thus the likelihood of a bailout will not be affected by the decision of a single bank, but merely by the aggregate decision of the banks overall. Therefore we can take $p(\pi)$ as given for a bank optimising their risk-taking. Then the first order condition for banks maximizing their profits by choosing the optimal risk level is given as

$$\frac{\partial \Pi_B}{\partial \pi} = ((r_L - r_D) - c\pi - p(\pi)(r_L - r_D)) L = 0, \tag{33.60}$$

which we can re-write as

$$p(\pi) = 1 - \eta\pi, \tag{33.61}$$

with $\eta = \frac{c}{r_L - r_D}$. This can be interpreted as the ratio of the costs increasing the success rate (c) and the benefits from doing so, $r_L - r_D$. As all banks are alike, in equilibrium we require that the equilibrium repayment rate of loans, π^* is identical for all banks. Using equation (33.61) to insert into equation (33.57), we get that in equilibrium the bank profits are given by

$$\Pi_B^* = \pi^* (r_L - r_D) L - \frac{1}{2} c\pi^{*2} L + (1 - \pi^*)(1 - \eta\pi^*)(r_L - r_D) L. \tag{33.62}$$

This gives us $\frac{\partial \Pi_B^*}{\partial \pi^*} = -c(1 - \pi^*) L < 0$ and hence if there are multiple equilibria, that is multiple values if π fulfilling the first order condition in equation (33.60), banks would prefer the solution with the lower success rate.

Looking at possible solutions to the first order condition (33.61), can consult figure 33.4 which shows the right-hand side of equation (33.61) as the straight descending line indicated by η_i and its left-hand side as the increasing line p_i as defined in equation (33.58). We see immediately that any equilibrium must be such that $\pi < 1 - \overline{\lambda}$. The potential second equilibrium at $\pi > \underline{\lambda}$ would be inferior as equilibria with lower success rates are preferred as shown above. Hence the equilibrium is given by $\frac{\overline{\lambda}}{1-\pi} = 1 - \eta\pi$, which solves for

$$\pi^* = \frac{1+\eta}{2\eta} - \sqrt{\frac{(1+\eta)^2}{4\eta^2} - \frac{1-\overline{\lambda}}{\eta}}. \tag{33.63}$$

Given that in this case with $\pi < 1 - \overline{\lambda}$ we have $\Pi_G = -\overline{\lambda} C_B L - \left(1 - \pi - \overline{\lambda}\right) C_F L$, we find that $\underline{\lambda}$ is irrelevant in equilibrium and can take any value $\underline{\lambda} \in \left[0; \overline{\lambda}^*\right]$. The optimal $\overline{\lambda}^*$ we obtain from maximizing Π_G, which gives us as the first order condition that maximises the benefits of the government over the optimal resources $\overline{\lambda}$ as

$$\frac{\partial \Pi_G}{\partial \overline{\lambda}} = C_B L - C_F L - C_F \frac{\partial \pi^*}{\partial \overline{\lambda}} L = 0. \tag{33.64}$$

From equation (33.63) we easily get that $\frac{\partial \pi^*}{\partial \overline{\lambda}} = -\frac{1}{2\eta\sqrt{\frac{(1+\eta)^2}{4\eta^2} - \frac{1-\overline{\lambda}}{\eta}}}$. Defining $\hat{C} = \frac{C_F - C_B}{C_F}$ as the relative costs to the government of banks failing and being bailed out, we then can rewrite the first order condition in equation (33.64) as $\frac{\partial \pi^*}{\partial \overline{\lambda}} = -\hat{C}$, which after inserting solves for the optimal bailout resources to be

$$\overline{\lambda}^* = 1 - \frac{(1+\eta)^2 \hat{C}^2 - 1}{4\eta\hat{C}^2}. \tag{33.65}$$

As we need to ensure that $0 \leq \overline{\lambda}^* \leq 1$, we require that $\hat{C}^2 (1-\eta)^2 \leq 1 \leq \hat{C}^2 (1+\eta)^2$. If these conditions are violated, the optimal solutions are $\overline{\lambda}^* = 1$ or $\overline{\lambda}^* = 0$, corresponding to the case of a guaranteed full bailout and no possible bailout, respectively. For simplicity we assume that this condition is fulfilled.

Using from the first order condition that $\sqrt{\frac{(1+\eta)^2}{4\eta^2} - \frac{1-\lambda}{\eta}} = \frac{1}{2\eta\hat{C}}$, we can rewrite the optimal risk banks take from equation (33.63) as

$$\pi^* = \frac{(1+\eta)\hat{C} - 1}{2\eta\hat{C}}. \tag{33.66}$$

Using equation (33.61) we then easily get that the fraction of banks bailed out in equilibrium will be

$$p^* = \frac{\hat{C}(1-\eta) + 1}{2\hat{C}}. \tag{33.67}$$

In order to analyse the equilibrium outcomes of the risk-taking by banks and their likelihood of being bailed out, we can now derive the following partial derivatives of our equilibrium repayment rates and the probability of a bank being bailed out. We find that

$$
\begin{aligned}
\frac{\partial \overline{\lambda}^*}{\partial \hat{C}} &= -\frac{1}{2\eta \hat{C}^3} < 0, \\
\frac{\partial \overline{\lambda}^*}{\partial \eta} &= \frac{1}{4\eta^2}\left(1 - \frac{1}{\hat{C}^2} - \eta^2\right) < 0, \\
\frac{\partial \pi^*_{\overline{\lambda}}}{\partial \hat{C}} &= \frac{1}{2\eta \hat{C}^2} > 0, \\
\frac{\partial \pi^*_{\overline{\lambda}}}{\partial \eta} &= \frac{1}{2\eta^2}\left(\frac{1}{\hat{C}} - 1\right) > 0, \\
\frac{\partial p^*_{\overline{\lambda}}}{\partial \hat{C}} &= -\eta \frac{\partial \pi^*}{\partial \hat{C}} < 0, \\
\frac{\partial p^*_{\overline{\lambda}}}{\partial \eta} &= -\pi^* - \eta \frac{\partial \pi^*}{\partial \eta} < 0,
\end{aligned}
\tag{33.68}
$$

where for the second and fourth results we note that $\hat{C} < 1$ and the final two results can be derived from directly from equation (33.61).

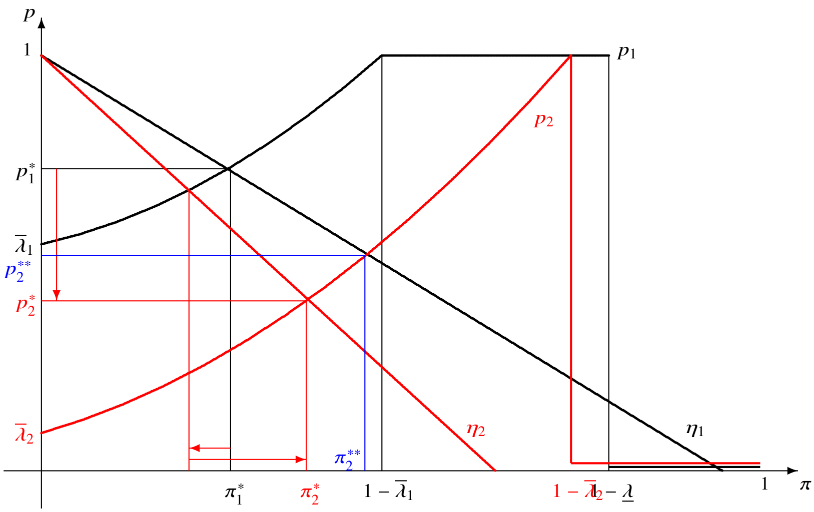


Fig. 33.4: Comparative statics of the equilibrium success rates

If, for the government, letting banks fail becomes relatively more expensive, $\hat{C}$ increases, the government will reduce the resources available for bailouts, $\overline{\lambda}^*$. This seems counterintuitive at first as reducing the resources available for bailouts on its own would increase the number of banks failing as less can be bailed out, and thus costs to the government would increase. There is, however, a secondary effect in that the reduced resources for bailouts also affect the banks' risk-taking. In the light of increased costs of not bailing out failed banks, governments could employ two strategies to reduce these costs. Firstly, they could increase the resources available for bailouts, but this would increase their overall costs. The second strategy would be the reduce the number of failing banks, which would reduce their costs. The only mechanism to pursue the second strategy for the government is to reduce the resources available for bailouts. In this case, as from above we have $\frac{\partial \pi^*_{\overline{\lambda}}}{\partial \overline{\lambda}} < 0$, the risks banks take reduce and therefore the costs of not bailing out banks will reduce, this is the resource effect. Of course, this reduction in costs due to lower risk-taking by banks has to be balanced against the reduced resources available to bail out banks and thus the government having to let banks fail, which is more costly, the cost effect. This second effect of increasing costs is lower than the resource effect of lowering government costs. In figure 33.4 we indicate this overall effect where we assume that overall the resources available reduce to $\overline{\lambda}_2$.

If we increase the costs to banks of reducing risks, c, or reduce the benefits of such reduced risk, $r_L - r_D$, both increasing η, we again observe two effects. Firstly, taking risks becomes more attractive to banks, increasing the risk they are taking and the likelihood of them failing, the cost effect,. In figure 33.4 this is indicated by increasing the slope of the curve indicated η_2, which would lead to an equilibrium with a lower success rate, thus higher risks. However, the government will also react and reduce the resources available for bailouts, $\overline{\lambda}$, as this is the only way the government can induce less risky behaviour and thus reduce the overall costs to the government. This resource effect reduces the risk-taking by banks. The resource effect overall dominates the cost effect. Figure 33.4 illustrates this result where the bank costs increase to η_2, causing the government to reduce bailout resources to $\overline{\lambda}_2$, which reduces the risks taken overall.

We have seen that banks react to the likelihood of a bailout by increasing their risks if bailouts become more likely, increasing costs to the government from failing banks. Balancing the costs of bailing out banks and the social costs of banks failing and a systemic banking crisis, the government would provide resources for bailouts that minimises these costs. These government incentives interact with the incentives for banks to determine their optimal level of risk. If taking risks becomes more attractive to banks, this can actually reduce the overall risk bank are taking as the government will react to this change in the incentives to banks by reducing the resources for a bailout; this will then reduce the incentives of banks to increase risks and could even reduce the overall risks banks take due to reduced bailouts.

Reading Krause (2025a)

Conclusions

We have seen that bailouts of failing banks can be effective in reducing the impact of bank failures on depositors and spreading any losses to other banks, thus preventing a systemic banking crisis. Bailouts have been shown to be of particular value where the liquidation of banks would yield low revenue relative to the value bank can generate when continuing to operate. However, bailouts are not necessarily the most efficient way of preventing bank failures, in many cases it would be preferable to include depositors into the restructuring of banks, a so-called bail-in. The anticipation of a bailout will also affect the incentives of banks to increase the risks they are taking, a moral hazard, as depositors will not require compensation for the risks the bank is taking, given they will be repaid due to the bailout. Governments should take these incentives into account and limit the resources for bailouts such that banks cannot be sure to obtain a bailout and thus parts of the costs arising from taking higher risks will be internalised through higher deposit costs.

Bailouts are often assumed to be limited to depositors, but a main risk of a bank failure arises from the spread of any losses to other banks and thereby causing a systemic banking crisis. Thus a bailout might not only encompass depositors but also the interbank loans provided by other banks. In this case, there are incentives for banks to provide such interbank loans over the provision of ordinary loans as interbank loans are then protected from default by the bailout of the other bank, reducing the risk of this bank failing, but also reducing the costs of deposits due to the lower risks to depositors; this will increase profits of banks.

We thus see that bailouts can be an effective, although costly, tool to reduce systemic risk. It has however, undesirable effects on the risk-taking of banks and for this reason bailouts should be limited. Bailouts may also distort incentives as it becomes more profitable to provide funds to banks that are covered by bailouts, such as interbank loans, rather than providing ordinary loans. This will increase the balance sheet of banks through high interbank lending, but reduce the amount of funds available for investment and thus can affect economic growth.

Review

Systemic risk is concerned with the spread of bank failures to other banks. Such failures can be propagated mainly through direct exposures of banks to each other; this might take the form of interbank loans, but also through exposures in derivatives markets, the so-called counterparty risk. Such contagion does not necessarily have to be triggered by a bank making losses from higher than expected loan defaults or other losses, but can have its origin in a lack of liquidity due to high deposit withdrawals. Such deposit withdrawals do not need to have their causes in the bank being not sound, but could be due to the desire by depositors to purchase consumption goods. In order to raise sufficient cash reserves such that these deposit withdrawals can be met, banks may have to sell assets at a loss, which then can cause the bank to transmit these losses to other banks. Also, facing the withdrawals of deposits and the subsequent liquidity shortage, may cause interbank loans to be withdrawn, exacerbating the liquidity shortage and the potential losses from selling assets.

While the transmission of bank failures is through the exposure of banks to each other, the number of banks failing can be increased by banks making the decision to provide loans to companies with similar characteristics and therefore, if one bank fails, it is likely that other banks will also face significant losses, making the transmission of losses that lead to bank failures across banks, all weakened from losses, more likely. While it is optimal for banks to behave like this, it increases the vulnerability of the banking system to a systemic crisis. The transmission of losses through inte5rbank loans and other exposures between banks can be very complex. With banks providing interbank loans to many banks and obtaining interbank loans from other banks, a complex network such interbank exposures emerge, where the failure of one or a small number of banks can cause a cascade of losses that spread through this network, causing many banks to fail, even if the initial losses are only small.

While bailouts can be an effective tool to prevent a systemic banking crisis by either preventing the failure of a single bank or intervening if a critical number of banks has failed. However, such interventions are not only costly, but their anticipation causes a moral hazard in that banks will increase the risk they are taking, given that they do not have to bear the full costs of their potential failure; banks may also divert loans towards interbank loans if they are included into the bailout scheme as such loans are less risky than ordinary loans. It is therefore desirable to limit the level of bailouts to strike a balance between the costs of letting banks fail and the costs of bailing out banks that take on additional risks.

Overall, we have established that systemic risk can originate from banks incurring losses directly or from their attempt to raise additional cash reserves due a liquidity shortage. These losses are then transmitted to other banks through interbank loans. If banks provide loans to similar borrowers, losses will occur in many banks simultaneously, making it more likely that any transmitted losses will cause other banks to fail. Using bailouts to prevent the spread of such failures need to be implemented carefully to ensure they not incentivise banks to take higher risks.

Part VII
Banking regulation

Banking is probably the most heavily regulated industry. In order to operate a banking business a licence to do so is commonly required and the granting of such licences is subject to strict conditions in terms of minimum equity requirements, conditions that directors of the bank have to meet strict requirements in term of experience and qualifications, as well as corporate governance, including risk management. Such strict requirements on establishing and operating are rarely imposed on other industries, apart from insurance companies and some protected professions such as lawyers, accountants, and doctors. The justification for such strict requirements for operating,a nd further regulatory constraints we introduce below is that banks play a central role in the economy and any bank failing imposes significant social costs through its own failure but also the risk of failures to spread to other banks, systemic risk.

However, banks are not only subject to strict licensing requirements, the way they conduct their business is also heavily regulated. Unlike non-financial companies, banks are limited on how much loans they can provide and how much other risky assets they can hold, thus the extent of their business, as the amount of loans cannot exceed a certain multiple of the equity they are holding. This is known as the capital requirements for banks. How these capital requirements are calculated can be complex, but in general they are determined by taking into account the risks of the loans the bank has provided and the risks of any other assets. The aim of capital regulation is to reduce the risk of banks failing due to losses on their loans and other assets, but imposing such restrictions will impact the incentives of banks when providing loans and chapter 34 will discuss the impact of minimum capital requirements on bank behaviour.

In addition to capital requirements, banks are also subject to minimum liquidity requirements, typically based on the deposits a bank has obtained. Such minimum liquidity requirements have the goal to ensure banks are able to repay any deposits that are withdrawn as well as provide resilience to interbank loans not being extended. In practice these liquidity requirements are seen as less stringent than capital requirements, but we will see in chapter 35 that they can nevertheless impact lending decisions by banks.

A concern for any regulator is that despite the stringent regulation, banks may fail. While in practice this will commonly result in the bank being taken over by another bank to avoid an outright bank failure, decisions need to be made on when and how to resort to measures that lead to the liquidation or sale of a bank. Chapter 36 will therefore discuss the way such decisions should be taken.

Banking regulations are enacted by each country and might therefore differ; with banks operating across borders, they might exploit the different regulatory requirements set in each country. The consequences of exploiting regulatory differences is discussed in chapter 37, along side the possibility of common regulation across countries.

Chapter 34

Capital regulation

Banks are subject to minimum capital requirements that often take the form of banks having to hold equity that exceeds a certain fraction of the risk-weighted assets, principally the loans banks have provided. With loans mainly financed by deposits, this indirectly limits the amount of deposits banks take and hence the leverage of banks. It is thus often more convenient to analyse minimum capital requirements as the maximum leverage a bank can have. This chapter explore the implications of such capital requirements.

In chapter 34.1 we will firstly look at the way binding capital requirements affect the incentives of banks to take on risks and investigate if higher capital requirements reduces the risk banks are taking before than in chapter 34.2 we will specifically look at the regulation of credit risk, given the importance it has in banking. Capital requirements are not the only regulation banks are impacted by and for this reason chapter 34.3 we will look at some other regulatory measures and how they affect bank behaviour in combination with capital requirements. Chapter 34.4 will then investigate the relationship between capital requirements and the other regulatory actions such as the granting of a bank licence and a regulatory regime using risk-based capital requirements. If, despite the regulatory constraints, a bank fails, it may be bailed out by a regulator, central bank, or government, and chapter 34.5 will look at the consequences of such bailouts being considered. Using debt that is written down if the bank is not able to repay their deposit9ors and this debt is often referred to as a bail-in and chapter 34.6 will consider the optimality of such bail-in bonds. In many cases, regulations and the provision of loans are directly or indirectly influenced by politicians and we will explore the relationship between banks and politicians in chapter 34.7.

34.1 Incentives to increase risks

We will look specifically at the incentives banks have to reduce or increase risks as capital requirements are tightened. We will determine in chapter 34.1.1 how the use of risk weights can distort the provision of loans towards more risky loans, while chapter 34.1.2 will focus on the costs of losses to depositors if a bank fails. While most models assume that banks, regulators, and depositors are aware of the risks a bank takes, chapter 34.1.3 will investigate the implications of regulators and depositors not being fully informed about these risks.

34.1.1 Distorted loan allocation

The provision of loans can be interpreted as a portfolio decision, where each loan L_i comprises one element of the resulting loan portfolio. The risk-free asset for banks would be cash reserves C on which they earn the risk-free interest rate r. The characteristics of the loans, their default rates, as well as the default correlations will determine the risks the bank is taking. Thus by providing loans with different characteristics, banks can reduce or increase the risks they are taking. We will use such a framework to determine the reaction of banks to a change in capital requirements.

We will at first determine the bank's decision on the provision of loans if it does not face any capital requirements and use this as a basis for the further analysis.

No capital requirements Let us assume that each of the N loans L_i are repaid with probability π_i and the bank sets the loan rate at r_L^i. Any funds not invested into loans is held as cash reserves. The bank's profits are then given by

$$\Pi_B = \sum_{i=1}^{N} \pi_i \left(1 + r_L^i\right) L_i + (1 + r) C - (1 + r_D) D, \tag{34.1}$$

where $\sum_{i=1}^{N} L_i + C = D + E$, represents the equality of assets and liabilities with D denoting the deposits on which interest r_D is payable, and E the bank's equity. Defining μ_i as the excess returns of the loans, $\mu_i = \pi_i \left(1 + r_L^i\right) - (1 + r)$, we can rewrite equation (34.1) as

$$\Pi_B = (1 + r) E + \sum_{i=1}^{N} \mu_i L_i. \tag{34.2}$$

In doing so we have assumed that the deposit rate equals the risk-free rate; this can be justified if we assume that depositors are covered by deposit insurance that is provided for free by the government. This assumption allows us to focus on the risks banks pose without the added complication that deposit rates depend on this risk.

Using matrix notation with μ denoting the vector of expected excess returns, $\mathbf{L}$ the vector of loan amounts and $\mathbf{\Sigma}$ the covariance matrix of these excess returns, we can write the expected profits of banks and their variance as

$$\begin{aligned} E\left[\Pi_B\right] &= (1 + r) E + \mu^T \mathbf{L}, \\ Var\left[\Pi_B\right] &= \mathbf{L}^T \mathbf{\Sigma} \mathbf{L}. \end{aligned} \tag{34.3}$$

The expected utility of the bank is then given by $E\left[U\left(\Pi_B\right)\right] = U\left(E\left[\Pi_B\right] - \frac{1}{2}zVar\left[\Pi_B\right]\right)$, with z representing the absolute risk aversion of the bank. The first order condition for the optimal loan allocation is then given by

$$\frac{\partial E\left[U\left(\Pi_B\right)\right]}{\partial \mathbf{L}} = \mu - z\mathbf{\Sigma}\mathbf{L} = \mathbf{0}, \tag{34.4}$$

or

$$\mathbf{L}^* = \frac{1}{z}\mathbf{\Sigma}^{-1}\mu. \tag{34.5}$$

Inserting the optimal allocation of loans into equation (34.3) we get the variance of bank profits as $Var\left[\Pi_B\right] = \frac{1}{z^2}\mu^T$
$mathbfSigma^{-1}\mu$ and the expected profits as

$$\begin{aligned} E\left[\Pi_B\right] &= (1+r)\,E + \frac{1}{z}\mu^T\mathbf{\Sigma}^{-1}\mu \\ &= (1+r)\,E + Var\left[\Pi_B\right]. \end{aligned} \tag{34.6}$$

We will now use this allocation of loans without any restrictions as a benchmark as we introduce capital requirements on banks.

Capital requirements with generic risk weights Capital requirements are introduced here by restricting the leverage of the bank, thus taking the amount of equity as given and limiting the amount of loans a bank can provide. We introduce a vector of weights, α, representing the risk weights for each loan such that $\alpha^T\mathbf{L}$ will give the risk-weighted loans the bank has provided. This risk-weighted amount of loans then cannot exceed κE, where κ denotes the leverage of the bank. This $\frac{1}{\kappa}\alpha^T\mathbf{L}$ would represent the equity the bank would need to hold. If we assume that these capital requirements are binding at the optimal unconstrained loan portfolio, $\mathbf{L}^*$, the bank will choose a loan allocation that meets this requirement exactly; hence we have

$$\kappa E = \alpha^T L. \tag{34.7}$$

With capital requirements being a constraint, the objective function of the bank becomes $\mathcal{L} = E\left[U\left(\Pi_B\right)\right] + \xi\left(\kappa E - \alpha^T\mathbf{L}\right)$, where ξ represents the Lagrange multiplier. The first order condition for a bank seeking the optimal loan allocation $\mathbf{L}$ then becomes

$$\frac{\partial \mathcal{L}}{\partial \mathbf{L}} = \mu - z\mathbf{\Sigma}\mathbf{L} - \xi\alpha = \mathbf{0}. \tag{34.8}$$

If this constraint is binding, we have a positive Lagrange multiplier and hence $\xi > 0$ such that the optimal loan allocation is given by

$$\mathbf{L} = \frac{1}{z}\mathbf{\Sigma}^{-1}\left(\mu - \xi\alpha\right). \tag{34.9}$$

If $\alpha_i \geq 0$ and thus loans obtain a positive risk weight, the allocation of loans is comparable to that of the unconstrained allocation in equation (34.5), but with a smaller expected return $\hat{\mu} = \mu - \xi\alpha$ which reduces the slope of the capital market line as shown in figure 34.1, commencing from the point where the constraint becomes binding.

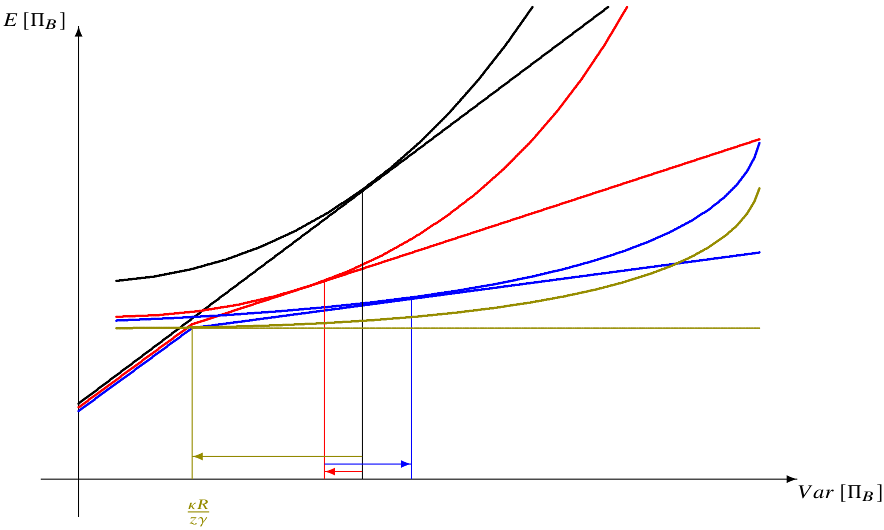


Fig. 34.1: Risk shifting with capital requirements

Using equations (34.7) and (34.9) in equation (34.3), we get the expected profits and their variance in this constrained loan allocation as

$$\begin{aligned} Var\left[\hat{\Pi}_B\right] &= \frac{1}{z^2}\left(\mu^T\mathbf{\Sigma}^{-1}\mu - \xi\mu^T\mathbf{\Sigma}^{-1}\alpha\right) - \frac{\xi}{2}\kappa E \\ &= Var\left[\Pi_B\right] - \frac{\xi}{z^2}\left(\mu^T\mathbf{\Sigma}^{-1}\alpha + z\kappa E\right), \\ E\left[\hat{\Pi}_B\right] &= (1 + r + \xi\kappa)\, E + zVar\left[\hat{\Pi}_B\right]. \end{aligned} \tag{34.10}$$

Thus compared to the unconstrained optimum from equation (34.6), the capital market line starts higher, but its slope will be lower if $\mu^T\mathbf{\Sigma} +^{-1} \alpha + z\kappa E > 0$. Comparing equations (34.6) and (34.10), we see that the constraint becomes binding if $\mu^T\mathbf{\Sigma}^{-1}\alpha > 0$ as then the constrained capital market line is below the unconstrained. Figure 34.1 shows that, depending on the preferences of banks, the resulting optimal portfolio may have a higher variance than the unrestricted portfolio of loans. If we introduce capital requirements, the capital market line becomes flatter from the point where it become binding, and the higher these capital requirements are, the lower this slope becomes. We clearly see that for lower capital requirements, the risk banks take are reduced; however, if we increase the capital requirements further, the risks banks take will increase again and may even exceed the risks the bank was exposed to without any capital requirements. The reason for this result is that banks would allocate larger loan amounts to those that have higher excess return, relative to the capital allocation, such as a higher ratio of $\frac{\mu_i}{\alpha_i}$. As these loans are generally more risky, the total risk increases.

The cause of the bank risk increasing is that the risk weights do not accurately reflect the risks of loans, in addition to the preferences of banks in term of the risk-return relationship enabling that the risks banks are taking in equilibrium will increase. We will therefore consider now the appropriate risk weights that ensure banks reduce the risks they are taking, regardless of the preferences of banks.

Using appropriate risk weights While with generic weights α there is the possibility that constraints on capital can actually increase the total risk taken by the bank, we can develop risk weights that prevent such a shift into loans with a high ratio of excess returns and risk weights, $\frac{\mu_i}{\alpha_i}$. Let us set $\alpha_i = \gamma\mu_i$, such that the constraint is proportional to the excess returns. If the Capital Asset Pricing Model holds and $\mu = \beta\mu_M$, with $\beta = \frac{\mathbf{\Sigma L}}{\mathbf{L}^T\mathbf{\Sigma L}}$ and μ_M denoting the market excess return, we have $\alpha = \gamma\beta\mu_M$ and capital requirements are proportional to the systematic risk of the loan.

Setting $\alpha = \gamma\mu$, we get from equation (34.9) that $\mathbf{L} = \frac{1-\xi\gamma}{z}\mathbf{\Sigma}^{-1}\mu$ and hence from equations (34.7) and (34.3) that

$$\kappa E = \alpha^T\mathbf{L} = \gamma\mu^T\mathbf{L} = \gamma\left(E\left[\hat{\Pi}_B\right] - (1+r)\,E\right), \tag{34.11}$$

such that the expected profits of the bank is given by

$$E\left[\hat{\hat{\Pi}}_B\right] = \left(1 + r + \frac{\kappa}{\gamma}\right)E, \tag{34.12}$$

which is independent of the risk and hence the capital market line will be horizontal. As is apparent from figure 34.1, in this case the risk will always be reduced as the optimal solution will be a corner solution where the capital requirements become binding. When comparing equation (34.12) with equation (34.6), we see that the variance of profits is given by

$$Var\left[\hat{\hat{\Pi}}_B\right] = \frac{\kappa E}{z\gamma}. \tag{34.13}$$

Thus only if capital constraints are implemented proportionally to excess returns can we guarantee that banks will reduce risks; in all other cases, including a leverage ratio when using $\alpha = \mathbf{L}$, we see a reduction in the amount of lending, but this effect is at least partially offset by a shift towards assets with higher excess returns and thus often higher risks as these excess returns are often the result of banks taking higher risks. This can lead to the total risks bank take increasing as capital requirements become more strict. If the risk weight assigned to loans is proportional to their excess returns, the overall risk will always reduce as less loans are given, without distorting the allocation towards more risky loans.

Summary We have thus seen that imposing capital requirements on banks can increase the risks banks are taking. While capital requirements may reduce the amount of loans that banks can provide, they may provide more risky loans; the reduced risk from providing less loans is outweighed by the loans that are provided being more risky. This may be the case if the risk weights that are assigned to loans, and on which the calculation of capital requirements are based, do not align with the actual risks of the loans. In addition, the preferences of banks regarding the risk-return relationship have to be such that an equilibrium with higher risks is preferred. It is only in the case where the risk weights assigned to loans reflect the excess returns banks generate from these loans that the kapital requirements can ensure that risks are reduced.

Reading Rochet (1992)

34.1.2 The costs of deposit insurance

A concern when regulating banks can be the costs to a regulator or the government of the bank failing, most notably the costs of providing deposit insurance, whether implicit deposit insurance or a premium funded explicit scheme. We will now investigate how capital requirements can affect these costs.

Assume that a bank can provide loans with a high repayment rate, π_H^j or a low repayment rate, π_L^j. The repayment rates depend on the state of the economy, which can be either good, and hence give a repayment rate of π_i^G, or bad, with repayment rate $\pi_i^B < \pi_i^G$. With a loan rate r_L and deposit rate r_D, we assume that we have

$$\pi_L^B (1 + r_L) < \pi_H^B (1 + r_L) < 1 + r_D < \pi_H^G (1 + r_L) < \pi_L^G (1 + r_L) . \qquad (34.14)$$

Firstly we observe that if the economy is in a good state, G, the repayments from loans always allow the bank to cover their deposits and hence the bank can never fail and the deposit insurance will not be required to make any payments. Furthermore, we see that the bank providing a loan with a lower repayment rate, π_L^j will be more risky as it obtains a very low return in the bad state of the economy, but a very high return in the good state of the economy. As we are only concerned about the possibility of the bank failing and the deposit insurance having to make payments, we will ignore the good state of the economy; it follows from the above assumption that $\pi_L^B < \pi_H^B$ and hence we can identify these two types of loans as having low and high repayment rates, respectively.

Suppose the bank invests a fraction γ of its funds into the high-risk loans, those loans with a low repayment rate; the funds available for banks to lend out, L, will consist of deposits D and equity E, such that $L = D + E$. The costs of a bank failing consists of the amount of deposits that have not been repaid and have thus to be repaid by the deposit insurance or the government. Thus we have in the bad state of the economy that concerns us these costs given by

$$\Pi_G = (1 + r_D) D - \left(\gamma \pi_L^B + (1 - \gamma) \pi_H^B\right) (1 + r_L) (D + E) . \qquad (34.15)$$

From the total payment to depositors the government can deduct the repayments received from the loans the bank has given, which will be a combination of high-risk and low-risk loans.

Constant equity If we hold the amount of equity the bank holds constant and increase the leverage through increasing deposits, we see unsurprisingly that higher deposits increase the costs of a bank failing; we easily obtain that

$$\frac{\partial \Pi_G}{\partial D} = (1 + r_D) - \left(\gamma \pi_L^B + (1 - \gamma)\, \pi_H^L\right) (1 + r_L) > 0, \tag{34.16}$$

where the sign of this derivative follows from our assumptions in equation (34.14). Similarly, an increase in the risks taken by increasing the fraction of the high-risk loans, γ, increases the costs. With equation (34.14) we obtain not surprisingly that the costs to government are increasing as

$$\frac{\partial \Pi_G}{\partial \gamma} = \left(\pi_H^B - \pi_L^B\right) (1 + r_L)\,(D + E) > 0. \tag{34.17}$$

Suppose we do impose a constraint on the risks banks can take by limiting their leverage, thus the amount of deposits they can obtain for a given amount of equity, they may react by increasing the risks they are taking, a higher γ. Thus the lower leverage, lower deposits D, reduce the risks banks are taking, but this may well be at least partially offset by banks providing more high-risk loans to retain their profitability. The effect on government costs are given by

$$\frac{\partial^2 \Pi_G}{\partial D \partial \gamma} = \pi_H^B - \pi_L^B > 0. \tag{34.18}$$

The marginal costs of reducing deposits, and hence leverage, are becoming smaller the more restrictive capital requirements become, thus the lower the amount of deposits banks obtain; the incentives for banks to increase their risks by providing high-risk loans most likely increases the more restrictive capital requirements are, which will in turn increase the cost of the bank failing due to less resources being available to repay depositors. Let us nevertheless assume that the net effect remains such that risks are reduced when the capital requirements are increased. In this case the positive effects of imposing leverage restrictions are reduced and the government would need to impose ever tight capital requirements to reduce their costs.

The government cannot find the optimal leverage, thus the optimal amount of deposits, that minimises their costs; the smaller the amount of deposits are, the lower government costs will become and the government would want banks to obtain no deposits.

Increasing equity Rather than assuming that banks hold their equity constant and thus any tighter restrictions on leverage will result in reduced deposits, banks can, at least in the long run, increase their equity to maintain their capital requirements without reducing deposits. Total assets are given by the loans provided, $L = D + E$, and capital requirements set as $L = D + E = (1 + \kappa) E$ if we define leverage as κ through $D = \kappa E$. Reducing the leverage of banks, will reduce the costs to governments as we can easily see from inserting these expressions into equation (34.15) that

$$\frac{\partial \Pi_G}{\partial \kappa} = \left((1 + r_D) - \left(\gamma \pi_L^B + (1 - \gamma) \pi_H^B\right) (1 + r_L)\right) E > 0, \tag{34.19}$$

with the final inequality arising from the assumption in equation (34.14). Here we replace deposits by equity rather than reduce deposits to meet the capital requirements. Similarly, the effect of increasing equity would yield

$$\frac{\partial \Pi_G}{\partial E} = (1 + r_D) \kappa - \left(\gamma \pi_L^B + (1 - \gamma) \pi_H^B\right) (1 + r_L). \tag{34.20}$$

The costs to government can be minimized if this equation is set equal to zero as the first order condition for minimizing costs, giving us an optimal leverage of

$$\kappa^* = \frac{\left(\gamma \pi_L^B + (1 - \gamma) \pi_H^B\right) (1 + r_L)}{(1 + r_D) - \left(\gamma \pi_L^B + (1 - \gamma) \pi_H^B\right) (1 + r_L)}. \tag{34.21}$$

We thus obtain an optimal leverage that minimizes the costs to the government of providing deposit insurance. We can easily see that $\frac{\partial \kappa^*}{\partial \gamma} < 0$ and hence banks choosing more risky loans will face tighter capital requirements.

Using that $D = \kappa E$ and inserting the optimal leverage, κ^* into the profits of the bank in equation (34.15), we obtain $\Pi_G = 0$ and hence the costs are minimized such that the deposit insurance does not have to make any net payments. We also see that increasing risks by banks in reaction to stricter capital requirements will not affect the costs to the government; this will be taken into consideration by making capital requirements for banks more strict, thus only a reduced leverage, κ^*, will be allowed.

Summary We thus see that if banks react to the imposition of stricter capital requirements by reducing the amount of loans they provide, and hence the amount of deposits they obtain, while holding their amount of equity constant, the costs of deposit insurance may increase as banks increase the risks of the loans they provide. It is only if banks react to stricter capital requirements by increasing the amount of equity and replace deposits, that introducing stricter capital requirements will reduce the costs to government while maintaining a positive leverage. In this case, the leverage restriction will take into account an incentives of the bank to increase risks and the costs to the government of providing deposit insurance are unaffected. Hence, the effectiveness of capital requirements in limiting the costs of deposit insurance are dependent on the reaction of banks to stricter requirements.

Reading Furlong & Keeley (1989)

34.1.3 Incomplete information on risk

It is common to assume that regulators and depositors know the risks that banks are taking, but this will often not be the case as the disclosure requirements for risks are limited and hence regulators might be subject to adverse selection by banks being better informed about the risks they are taking than the regulator imposing capital requirements based on this risk. We will therefore here consider the impact uncertainty about the risks banks take has on the regulation of capital requirements.

As a benchmark, we will initially consider the case where risks banks take are known by all parties.

Known risks Let us begin by assuming that the risks banks take are known. Loans L_i are repaid with probability π_i, including interest r_L; these are financed by deposits D_i, yielding interest r_D, and equity E_i. In addition, banks face costs C_i of granting loans and attracting deposits to finance these. These costs are deducted from the amount that can be lent out as we assume that they were incurred before the loans were granted and thus reduce the resources available for lending, thus $L_i = D_i + E_i - C_i$.

In the social optimum we would seek to maximize the proceeds from granting these loans, less the costs of deposits to finance the loans, hence

$$\Pi_S^i = \pi_i \left(1 + r_L\right) L_i - \left(1 + r_D^i\right) D_i, \tag{34.22}$$

such that the optimal amount of deposits are given by the first order condition

$$\frac{\partial \Pi_S}{\partial D_i} = \pi_i \left(1 + r_L\right) \left(1 - \frac{\partial C_i}{\partial D_i}\right) = \left(1 + r_D^i\right) = 0. \tag{34.23}$$

In the social optimum banks do not fail as we assume that $\pi_i \left(1 + r_L\right) > 1 + r_D$ and hence the deposit rate will be equal to the risk-free rate, $r_D^i = r$. The first order condition in equation (34.23) then solves for

$$\frac{\partial C_i}{\partial D_i} = \frac{\pi_i \left(1 + r_L\right) - (1 + r)}{\pi_i \left(1 + r_L\right)}. \tag{34.24}$$

The amount of deposits can now be obtained from solving this condition and given the amount of equity the bank holds, the optimal leverage κ^* would emerge.

In perfect competition with full knowledge of the bank risk, the bank would maximize their profits

$$\Pi_B^i = \pi_i \left(\left(1 + r_L\right) L_i - \left(1 + r_D^i\right) D_i\right). \tag{34.25}$$

Here the bank would only repay their depositors if their loans have been repaid and depositors would take into account the possibility of the bank failing. Their expected return on deposits would have to equal the risk-free rate if we assume that banks extract all surplus from depositors. Hence we have

$$\pi_i \left(1 + r_D^i\right) = 1 + r, \tag{34.26}$$

which solves for the deposit rate to be set at

$$1 + r_D^i = \frac{1 + r}{\pi_i}. \tag{34.27}$$

Inserting this result into equation (34.25) recovers the social benefits from equation (34.22) when noting that there $r_D^i = r$. The leverage in the case of perfect information is identical to the social optimum, making regulation unnecessary.

Unknown risks Let us now assume that depositors do not know the risk of banks. In this case the condition for the deposit rate in equation (34.26) will change to take into account the unknown risk, π_i. Furthermore, we assume that the amount of deposits are also unknown as they would depend on the risk through the optimal leverage the bank would choose. Hence equation (34.26) changes to

$$\left(1 + r_D^i\right) \int_0^1 \pi_i D\left(\pi_i\right) dF\left(\pi_i\right) = (1 + r) \int_0^1 D\left(\pi_i\right) dF\left(\pi_i\right), \tag{34.28}$$

where $F(\cdot)$ denotes the probability function for the risk π_i. We have taken into account the unknown risk and unknown deposit size on the right-hand side and compare this to the same amount of deposits on the left-hand side to make the payments received from deposits and the risk-free asset comparable. Solving this equation we get

$$1 + r_D^i = \frac{1 + r}{\overline{\pi}}, \tag{34.29}$$

where $\overline{\pi} = \frac{\int_0^1 \pi_i D(\pi_i) dF(\pi_i)}{\int_0^1 D(\pi_i) dF(\pi_i)}$ denotes the average success probability, weighted by the amount of deposits. This deposit rate is identical for all banks as long as we assume that the distribution of risks is identical for all banks. We can now insert this deposit rate into equation (34.25) such that we obtain the bank profits as

$$\Pi_B^i = \pi_i \left((1 + r_L) \left(E_i + D_i - C_i\right) - \frac{1 + r}{\overline{\pi}} D_i \right). \tag{34.30}$$

The first order condition for the optimal amount of deposits, and hence the optimal leverage for the given amount of equity, is then given by

$$\frac{\partial \Pi_B}{\partial D_i} = \pi_i \left(1 + r_L\right) \left(1 - \frac{\partial C_i}{\partial E_i}\right) - \frac{\pi_i}{\overline{\pi}} (1 + r) = 0, \tag{34.31}$$

which solves for

$$\frac{\partial C_i}{\partial D_i} = \frac{\overline{\pi}\left(1 + r_L\right) - (1 + r)}{\overline{\pi}\left(1 + r_L\right)}. \tag{34.32}$$

Comparing this result with the social optimum in the case of perfect knowledge of the risk in equation (34.24), we see that the right hand side is larger in equation (34.32) if $\overline{\pi} \geq \pi_i$ and smaller if $\overline{\pi} < \pi_i$. Assuming that marginal costs of providing loans and attracting deposits are increasing in deposits, this implies that risky banks, those with a low repayment rate π_i borrow more deposits, thus have a higher leverage. The consequence is that if the bank risk is not known, risky banks will borrow more than is optimal and safer banks will borrow less than would be optimal. The optimal leverage increases with the (average) risk in equations (34.24) and (34.32), thus the more risky banks are, the more leveraged they are. It is not that highly leveraged banks seek to take more risky loans, here the higher risks lead to a higher leverage.

We can now consider to limit the leverage of these risky banks by imposing a maximum leverage ratio κ on banks. Firstly we show that $\overline{\pi}$ is increasing in κ. To this end consider that banks will seek the leverage either as implied by equation (34.32), which is decreasing in $\overline{\pi}$, or they set $D = \kappa E$, whatever is the smaller value. Thus banks will either choose their optimal leverage or if the capital requirements are stricter through a a regulator setting a lower leverage, the bank will choose the leverage of this restriction. Figure 34.2 shows the relationship between the leverage and the risk the bank takes; κ^{**} denotes the optimal leverage, representing the optimal level of deposits as obtained from solving equation (34.32), that the bank would choose if it was not restricted by regulation. We will compare the average repayment rates with two capital requirements, one will be represented by a less strict maximum leverage, κ_1 and the other by a more strict maximum leverage, $\kappa_2 < \kappa_1$.

Recounting that $\overline{\pi}$ is the (weighted) average of the actual failure rates π_i, we can easily see that in the area marked A in this graph, the average π_i, $\overline{\pi}_A$, is largest as all values for π_i are above those of the other areas. Area B will thus have a smaller average repayment rate, $\overline{\pi}_B < \overline{\pi}_A$, given all values of π_i are smaller. Furthermore, the average repayment rate in area C will be even smaller, $\overline{\pi}_C < \overline{pi}_B$, as it is that while the range of repayment values π_i is identical to that of area B, its large values have a lower weight, given the unconstrained leverage limits its extent. Hence we find that the average repayment rates in the three marked areas fulfill $\overline{\pi}_A > \overline{\pi}_B > \overline{\pi}_C$.

The average failure rate associated with κ_2, $\overline{\pi}_2$, will be a weighted average of ares A and B, thus as such we have $\overline{\pi}_A > \overline{\pi}_2 > \overline{\pi}_B$. Similarly, the average failure rate for κ_1, $\overline{\pi}_1$, will be the average of ares A and B, the average of which was which is $\overline{\pi}_2$, and area C, hence we find that $\overline{\pi}_2 > \overline{\pi}_1 > \overline{\pi}_C$ and as the maximum leverage κ increases we see that average repayment rate decreases. From equation (34.29) we know that the deposit rate depends on the average failure rate $\overline{\pi}$. As shown above, if the leverage κ increases, the average repayment rate increases, in turn implying that the deposit rate also increases.

Seeking to maximize social welfare, we know that it is maximized at D^*, corresponding to κ^* for the leverage, when solving equation (34.24). The derivation of equation (34.24) assumed, though, that the deposit rate was fixed. We know from the above analysis that as the leverage ratio is increased, the deposit rate will increase. With $1 + r_D = \frac{1+r}{\overline{\pi}}$ in equation (34.32), we know that with an increased deposit rate would reduce the amount of deposits that unconstrained banks would take. As these banks are unconstrained, they could take on more deposits and this would increase

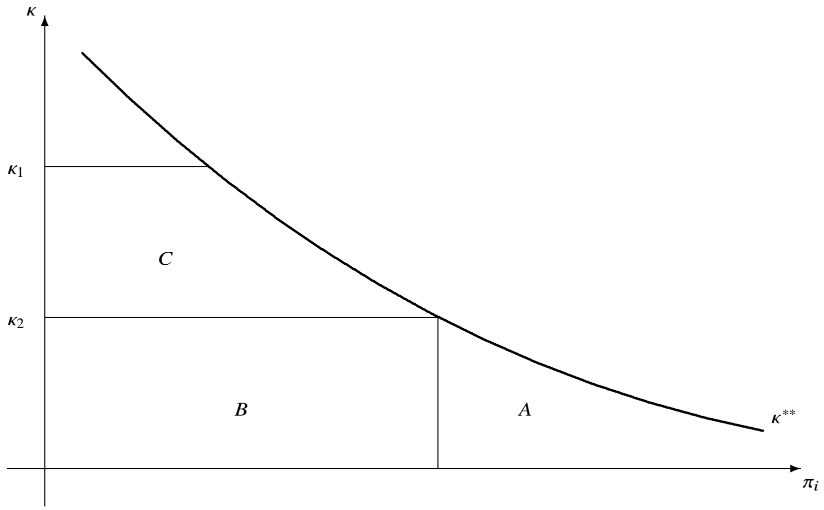


Fig. 34.2: Leverage and repayment rates

social welfare. Hence welfare is reduced as we increase the leverage ratio given that the deposit rate is increasing. We thus need to balance these two effects, on the one hand limiting the leverage of the high-risk banks and on the other hand inducing low-risk banks to increase their deposits. Figure 34.3 shows this trade-off and we see clearly that combining both effects results in a leverage ratio that is below the socially optimal level κ^* at $\kappa^{**} < \kappa^*$.

Even though a regulator has no informational advantage over depositors in this model, they can improve welfare by imposing a leverage constraint. This leverage constraint requires banks with low repayment rates, thus banks taking higher risks, to choose a lower leverage ratio than they otherwise choose. This Arises because the unlike in the social optimum deposits are not risk free as individual banks can fail. These additional costs to banks of higher deposit rates associated with higher leverage, will reduce the optimal leverage. Thus in a situation where deposit insurance makes the deposit rate unaffected by the risks the bank takes, here the higher leverage and high default rates on loans, capital requirements should be strengthened such that banks choosing to provide loans with high default rates that would choose a high leverage reduce their leverage and thus risk.

If the regulator has better information on the risk of the bank than the depositor, it is obvious that the inferred average repayment rate $\overline{\pi}$ will converge towards the actual repayment rate π_i as the information becomes more and more precise. Depositors could infer the information the regulator has from the leverage constraint imposed on the bank and the deposit rate adjusts such that the leverage ratio chosen by the bank increases towards the social optimum as the information becomes more precise.

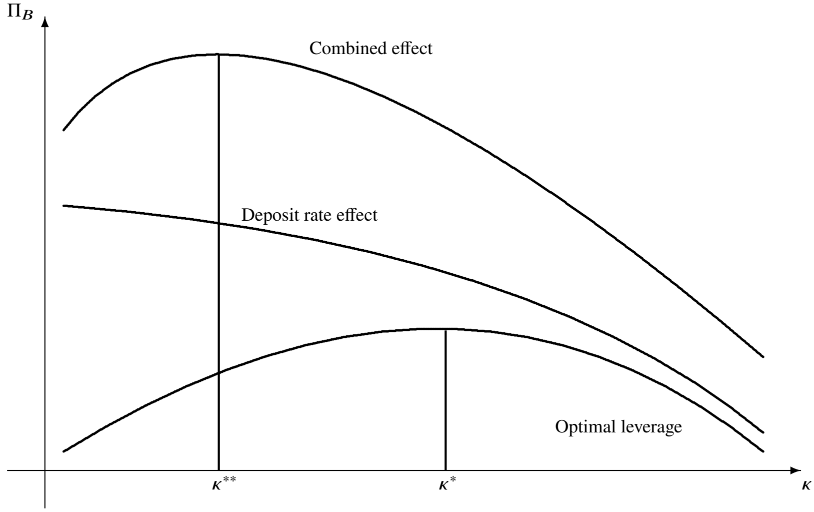


Fig. 34.3: Individually and socially optimal leverage

Summary We have seen that with regulators and depositors uncertain about the default of loans that banks provide, those banks that choose high default risks will select a leverage that is higher than socially optimal. Consequently, regulators would choose capital requirements higher than the capital requirements they would choose for a bank with the same expected, but known, default rate. This will not affect those banks who chose lower default and are thus less risky as they would optimally choose a lower leverage and hence these capital requirements would not be a restriction on their choice.

It is thus that in banking systems where regulators are less well informed about the risks banks take, capital requirements will be higher than in banking systems where regulators are better informed. These more stringent capital requirements will be binding only for banks taking high risks and serve to reduce the risk from high leverage that to added to the high default risk these banks are exposed to.

Reading Ding, Hill, & Perez-Reyna (2021)

Résumé

We have seen how the imposition of capital requirements can lead to a distortion in the provision of loans towards more risky loans if the risk weights do not adequately reflect risks, which can lead to the overall risks the bank is taking actually increasing with stricter capital requirements; this would depend on the preferences of the bank regarding the risk-return relationship. A similar effect might be observed if banks react to stricter capital requirements by reducing loans rather than increasing equity.

Finally, if the regulator is uncertain about the risks a bank will take, its capital requirements will be stricter than for banks whose risks are better known, assuming the expected risks are identical.

It has become apparent that the imposition of capital requirements does not necessarily reduce the risks banks are taking. being required to operate with a lower leverage and thus less profit opportunities, banks will have incentives to increase the risks they are taking in order to increase the profitability of the loan they are providing. When setting capital requirements, this effect needs to be considered.

34.2 Credit risk management

Capital requirements in itself reduce the risks banks pose by making it less likely that a bank fails as it has a larger cushion against losses and if a bank fails the social costs will be lower as the amount of deposits lost in a bank failure will be smaller. However, capital requirements are often used as a tool to direct banks towards granting less risky loans and as such reduce the risks of them failing even more. We will here consider how capital requirements are used to implement the optimal level of risk-taking by banks.

Let us assume that banks choose a leverage ratio of κ such that with equity E deposits D are determined as $D = \kappa E$ and the amount of loans is then given by $L = E + D = (1 + \kappa) D$. Banks face costs of screening and monitoring borrowers that are increasing in the lending, but also the repayment rate π such that these costs are given by $C = \frac{1}{2} c\pi^2 L = \frac{1}{2} (1 + \kappa) \pi D$. The higher the repayment rate the higher the monitoring costs will be; this is because ensuring that loans repaid becomes increasingly difficult as risks are already low and reducing them further will become increasingly difficult.

With loan rates r_L and deposit rates r_D, the profits of banks are given as

$$\Pi_B = \pi \left((1 + r_L) L - (1 + r_D) D \right) - C - E \tag{34.33}$$

Banks maximize their profits by choosing the optimal level of risk, π. The first order condition then becomes

$$\frac{\partial \Pi_B}{\partial \pi} = \left((1 + r_L) (1 + \kappa) - (1 + r_D) \kappa - c (1 + \kappa) \pi \right) E = 0, \tag{34.34}$$

where we have inserted for the costs and the amount of loans and deposits from the above assumptions. This first order condition solves for

$$\pi^* = \frac{1 + r_L}{c} - \frac{1 + r_D}{c} \frac{\kappa}{1 + \kappa}. \tag{34.35}$$

A regulator would take into account the social costs of bank failure, which includes the costs arising from the lost deposits of a failing bank. These costs are proportional to the amount of loans the bank has provided, and will include the financing of reimbursing any depositors for the losses they made, thus $\hat{C} = \hat{c} L = \hat{c} (1 + \kappa) E$.

Hence the objective function of the regulator becomes

$$\begin{aligned}\Pi_R = {} & \pi\left(\left(1+r_L\right)L-\left(1+r_D\right)D\right)-C-E \\ & -\left(1-\pi\right)\left(1+r_D\right)D-\hat{C},\end{aligned} \tag{34.36}$$

consisting of the profits of the bank, the costs of the deposits, and the financing costs from their reimbursement. The optimal level of risk for the regulator is then given from the first order condition, which becomes

$$\frac{\partial \Pi_R}{\partial \pi} = \left(\left(1+r_L\right)\left(1+\kappa\right)-c\left(1+\kappa\right)\pi\right)E = 0, \tag{34.37}$$

where we have inserted for the costs and the amount of loans and deposits from the above assumptions. This first order condition solves for

$$\hat{\pi}^* = \frac{1+r_L}{c} > \pi^*. \tag{34.38}$$

We see that the socially optimal repayment rate, $\hat{\pi}^*$, always exceeds the optimal repayment rate of the bank, π^*, regardless of any capital requirements imposed. It is thus not possible to induce the bank to reduce the risks they are taking by imposing capital requirements.

Given the regulator costs of $\hat{C}$, a regulator would only intervene and reimburse depositors if it is beneficial for them to do co. Let us assume that if the regulator intervenes, he can implement the repayment rate $\hat{\pi}^*$, while without intervention the bank could freely choose the repayment rate π that is best for them, but the regulator does not incur costs their costs of $\hat{C}$. Hence the regulator intervenes if from (34.36)

$$\begin{aligned}&\hat{\pi}^*\left(\left(1+r_L\right)\left(1+\kappa\right)-\frac{1}{2}c\left(1+\kappa\right)\hat{\pi}^*\right)-\left(1+\kappa\right)\hat{c} \\ &\geq \pi\left(\left(1+r_L\right)\left(1+\kappa\right)-\frac{1}{2}c\left(1+\kappa\right)\pi\right),\end{aligned} \tag{34.39}$$

where the right-hand side denotes the social welfare as relevant to the regulator intervening by imposing its optimal repayment rate and the left-hand side the bank profits, which are the social benefits without regulatory intervention. For convenience we have excluded the common element of equity and deposits of the bank as they appear on both sides. Inserting for $\hat{\pi}^*$ from equation (34.38) this condition transforms into

$$\left(\pi-\frac{1+r_L}{c}\right)^2 \geq 2\frac{\hat{c}}{c},$$

which then solves for

$$\pi \leq \hat{\pi}^{**} = \frac{1+r_L}{c}-\sqrt{\frac{2\hat{c}}{c}} = \hat{\pi}^*-\sqrt{\frac{2\hat{c}}{c}} < \hat{\pi}^*. \tag{34.40}$$

If we assume that $(1+r_D)^2 > 2c\hat{c}(1+\kappa)$, then we easily can verify that

$$\hat{\pi}^* > \hat{\pi}^{**} > \pi^*. \tag{34.41}$$

While the regulator cannot induce banks to choose the socially optimal repayment rate $\hat{\pi}^*$, the first best solution, it can induce it to choose $\hat{\pi}^{**}$, the intervention level. Assume that if banks choose $\pi < \hat{\pi}^{**}$, then the regulator intervenes such that the bank owners lose all value due to not meeting regulatory requirements. To ensure banks make choices that comply with this interventional level, it must be more beneficial for them to choose $\hat{\pi}^{**}$ than to choose π^*; thus we require that $\hat{\Pi}_B^{**} = \Pi_B(\hat{\pi}^{**}) \geq 0$, where the inequality indicates that if the regulator intervenes, the bank owners receive no payments. Using equation (34.33) this condition easily solves for

$$\kappa \leq \kappa^* = \frac{1}{\hat{\pi}^*}\frac{\frac{1}{2}c(\hat{\pi}^*)^2 - \hat{\pi}^*(1+r_L)+1}{r_L - r_D - \frac{1}{2}c\hat{\pi}^*}, \tag{34.42}$$

where we assume that $r_L - r_D - \frac{1}{2}c\hat{\pi}^* < 0$ and $\frac{1}{2}c(\hat{\pi}^*)^2 - \hat{\pi}^*(1+r_L)+1 < 0$.

Thus the second best solution, the intervention level of the regulator can be implemented by requiring a leverage ratio that does not exceed κ^*. It is that rather than setting a risk limit, which will be difficult to monitor as internal bank assessment may not be easily verified, regulators will use capital requirements, here the maximum leverage ratio κ^*, to induce banks to take risks that are not higher than what regulators are finding optimal.

We have seen that regulators determine the risk banks are taking by determining capital requirements. Banks reacting optimally to these capital requirements, will then implement the best possible risk level when providing loans. However, the socially optimal risk level may not be achievable as the costs for regulators would be higher than allowing banks to take higher risks. It is therefore that the capital requirements can be interpreted as a compromise between the interests of the banks and the regulator, representing the socially optimal risks.

Reading Carletti, Dell'Ariccia, & Marquez (2021)

34.3 Interactions with other regulations

Capital requirements based on the risk a bank is exposed to are not the only regulations that affect the leverage a bank will take. Banks are subject to a variety of other regulations that we will see can supplement or enhance the choices banks make with respect to their leverage or choice of loans. We will specifically look at the effect the addition of a maximum leverage ratio to risk-based capital requirements has on the choice of risks in chapter 34.3.1. Deposit rates may be limited to induce banks to provide low-risk loans and be used as a more effective substitute for capital requirements as we will show in chapter 34.3.2. Finally, in chapter 34.3.3 we will look at how deposit insurance can replace the imposition of capital requirements.

34.3.1 Leverage ratio

Capital requirements are commonly based on the risk a bank is exposed to rather than a maximum leverage ration that is unrelated to risks; it is frequently observed, however, that banks are also subject to a leverage ration that is set as being equal for all banks, regardless of the risks they are taking. We will consider here why a leverage ratio might be beneficial for companies to disclose the risk they are taking truthfully to a regulator.

Let us assume that the social costs of a bank failing are proportional to its deposits D, such that the total costs are given by cD. With banks failing if the loans the bank has provided are not repaid, which happens with probability $1 - \pi_i$, the expected costs of failure are $(1 - \pi_i)\, cD$. These costs need to be balanced against the costs of equity E_i, which is r_E; equity can be used to reduce the costs of failure as some of the losses will be borne by the bank owners rather than depositors. With capital requirements set through a capital ratio κ_i such that $E_i = \kappa_i^* D$, we thus have the optimal capital requirements given when setting the costs of the bank failing and its cost of equity equal, $(1 - \pi_i)\, cD = r_E \kappa_i^* D$. This solves for

$$\kappa_i^* = \frac{c\,(1 - \pi_i)}{r_E}. \tag{34.43}$$

We can now investigate how banks would choose their optimal capital and then subsequently how regulators could seek achieve the social optimum.

No capital requirements The bank profits are given by

$$\begin{aligned} \Pi_B^i &= \pi_i \left(\left(1 + r_L^i\right) L_i - (1 + r_D)\, D\right) - (1 + r_E)\, E_i \\ &= \left(\pi \left(\left(1 + r_L^i\right)(1 + \kappa_i) - (1 + r_D)\right) - (1 + r_E)\, \kappa_i\right) D, \end{aligned} \tag{34.44}$$

where the first term denotes the revenue from the loan, $L_i = D + E_i$ with its loan rate r_L^i, less the payments to depositors including interest r_D, provided the loan is repaid, and the final term the cost of equity with we deduct the obtain the economic profits of the bank; the second line is obtained by inserting that $E_i = \kappa_i D$. If $1 + r_E > \pi_i \left(1 + r_L^i\right)$ such that ther costs of equity exceed the expected return on loans, we get

$$\frac{\partial \Pi_B^i}{\partial \kappa_i} = \pi_i \left(1 + r_L^i\right) - (1 + r_E) < 0. \tag{34.45}$$

Consequently it would be optimal for banks to hold the lowest possible capital ratio κ_i, equivalent to the highest possible leverage $\frac{1}{\kappa_i}$. Bank will therefore not hold any equity and the costs of bank failure will exceed the cost of equity, making the social optimum not achievable without regulation.

With capital requirements We assume there are two types of loans available that banks can choose; one loan has a high repayment rate π_H and the other loan a low repayment rate $\pi_L < \pi_H$. the expected repayment of the low risk loan is higher than that of high-risk loan, thus we have $\pi_H \left(1 + r_L^H\right) > \pi_L \left(1 + r_L^L\right)$.

A regulator imposes capital requirements κ_i^*, depending on whether the bank chooses low-risk loans or high-risk loans. In this case, given equation (34.45), and if we reasonably suppose that the capital requirements for high-risk loans are higher than for low-risk loans, $\kappa_L^* > \kappa_H^*$, banks would want to report that they have provided low-risk loans as this would impose lower capital requirements. Hence, if the regulator cannot verify the risks of the loans, all banks would only be subject to the capital requirements of low-risk banks, κ_H. We therefore need to find a mechanism to ensure that banks are revealing the risks they are taking truthfully.

Assume that banks revealing the risks they are exposed to not truthfully are identified by the regulator with probability p and fined a fraction γ of their profits Π_B^i. Obviously, the bank providing low-risk loans would always identify itself as being low-risk, as this results in a lower capital requirements and cannot lead to a fine. We are thus only concerned about the bank providing the high-risk loans and want to provide incentives for them to identify themselves as such. The profits of the bank when disclosing their high risks truthfully are given by equation (34.44) as

$$\Pi_B^{LL} = \left(\pi_L \left(\left(1 + r_L^L\right)\left(1 + \kappa_L^*\right) - \left(1 + r_D\right)\right) - \left(1 + r_E\right)\kappa_L^*\right) D. \quad (34.46)$$

If the bank would disclose that it taking low risks while in fact taking high risks, the profits are given by

$$\begin{aligned} \Pi_B^{LH} &= p\left(1-\gamma\right)\left(\pi_L\left(1 + r_L^L\right)\left(1 + \kappa_H^*\right)\right. \\ &\quad \left. - \left(1 + r_D\right) - \left(1 + r_E\right)\kappa_H^*\right) D \\ &\quad + \left(1-p\right)\left(\pi_L\left(1 + r_L^L\right)\left(1 + \kappa_H^*\right)\right. \\ &\quad \left. - \left(1 + r_D\right) - \left(1 + r_E\right)\kappa_H^*\right) D \\ &= \left(1-\gamma p\right)\left(\pi_L\left(1 + r_L^L\right)\left(1 + \kappa_H^*\right)\right. \\ &\quad \left. - \left(1 + r_D\right) - \left(1 + r_E\right)\kappa_H^*\right) D, \end{aligned} \quad (34.47)$$

where the first term denotes the case the regulator identifies that the bank has not been truthful and therefore the bank retains only a fraction $1 - \gamma$ of their profits and the second term denotes the case in which the bank is not identified by the regulator and thus benefits from the lower capital requirements.

The bank will disclose the risks they are taking truthfully if it is more profitable to do so, $\Pi_B^{LL} \geq \Pi_B^{LH}$. This conditions can be solved as

$$\gamma p \geq \frac{\left(\pi_L\left(1 + r_L^L\right) - \left(1 + r_E\right)\right)\left(\kappa_L^* - \kappa_H^*\right)}{\left(\pi_L\left(1 + r_L^L\right) - \left(1 + r_E\right)\right)\kappa_H^* + \left(\pi_L\left(1 + r_L^L\right) - \left(1 + r_D\right)\right)}. \quad (34.48)$$

Thus if the regulator is able to identify banks not disclosing risks truthfully well, p, or the fines for making a wrong declaration are high, γ, banks will be truthfully declare that they have provided high-risk loans. However, if the regulator is weak in that γp is sufficiently low, the regulator would not be able to implement the higher capital requirements for banks taking higher risks, leaving them with capital that is too low.

We can now continue by introducing a leverage ratio to overcome the limitations of a weak regulator.

Leverage ratio The regulator introduces a leverage ratio in addition to the risk-based capital requirements of κ_i^*. This leverage ratio $\hat{\kappa}$ would not be binding for low-risk banks but binding for high-risk banks as $\kappa_H^* \hat{\kappa} > \kappa_L^*$. We can then rewrite the profits of the high-risk bank not revealing the risks truthfully, equation (34.47), as

$$\hat{\Pi}_B^{LH} = (1 - \gamma p) \left(\pi_L \left(\left(1 + r_L^L \right) (1 + \hat{\kappa}) - (1 + r_D) \right) - (1 + r_E) \hat{\kappa} \right) D. \qquad (34.49)$$

The profits of a bank disclosing their high risks truthfully remain unchanged. In order for a bank to truthfully reveal their high risks, the profits of doing so need to exceed those of disclosing wrongly that they have provided low-risk loans, hence $\Pi_B^{LL} \geq$
$widehatPi_B^{LH}$. This condition can be solved for

$$\hat{\kappa} \geq \frac{\kappa_L^* \left(\pi_L \left(1 + r_L^L \right) - (1 + r_E) \right) + \left(\pi_L \left(1 + r_L^L \right) - (1 + r_D) \right)}{(1 - \gamma p) \left(\pi_L \left(1 + r_L^L \right) - (1 + r_E) \right)}. \qquad (34.50)$$

Thus by setting a leverage ratio sufficiently high, the regulator can ensure that banks identify themselves as high-risk and comply with the higher capital requirements for their type of bank. The regulator could set the lowest value for the leverage that meets this condition, such that equation (34.50) becomes an equality. With a stronger supervisor, one that can impose higher fines γ or is more able to detect those that do not disclose their risks completely, p, the constraint becomes less restrictive and hence the leverage limit becomes less onerous.

Summary We have seen that by adding a leverage constraint to the capital requirements based on the risk of the bank will induce high-risk banks to disclose their risks truthfully while low-risk banks remain unaffected. It is through this mechanism that if the regulator is not sufficiently strong in enforcing the disclosure of the risks banks take, the optimal capital requirements, which are based on the risks the banks actually take, can be implemented; this was not possible without a leverage constraint.

Reading Blum (2008)

34.3.2 Deposit rate control

It is often desirable that banks choose to provide loans that are not too risky, however, the incentives of banks due to limited liability are often such that they prefer to provide loans with higher risks and thus higher returns if they are repaid. We will discuss here how deposit rates can be used in addition to capital requirements to achieve this aim where capital requirements alone would be ineffective.

Let us assume that banks can invest into two types of loans, one with success rate π_H and interest rate r_L^H and one with success rate $\pi_L < \pi_H$ and loan rate r_L^L such that $\pi_H \left(1 + r_L^H\right) \geq \pi_L \left(1 + r_L^L\right)$ and hence the low-risk loan would be preferable to the high-risk loan, which a regulator will seek to implement. Loans L are financed by deposits D requiring an interest rate r_D and equity E with cost of equity r_E, where $1 + r_E > \pi_H \left(1 + r_L^H\right)$. deposits and equity are linked through capital requirements κ such that $E = \kappa D$; this capital requirements κ determined the leverage of the bank, which is $\frac{1}{\kappa}$. The expected profits of banks are then given by

$$\begin{aligned} \Pi_B^i &= \pi_i \left(\left(1 + r_L^i\right) L - (1 + r_D) D\right) - (1 + r_E) E \\ &= \left(\pi_i \left(\left(1 + r_L^i\right)(1 + \kappa) - (1 + r_D)\right) - (1 + r_E)\kappa\right) D. \end{aligned} \tag{34.51}$$

The low-risk loan is chosen if it is more profitable than the high-risk loan, thus $\Pi_B^H \geq \Pi_B^L$, which requires

$$1 + r_D \leq 1 + r_D^* = \frac{\pi_H \left(1 + r_L^H\right) - \pi_L \left(1 + r_L^L\right)}{\pi_H - \pi_L} (1 + \kappa). \tag{34.52}$$

Markets are not perfectly competitive and we define $\varepsilon = \frac{\partial D}{\partial(1+r_D)} \frac{1+r_D}{D}$ as the elasticity of deposits with respect to the deposit rate. It is thus that the deposits a bank attracts will depend on the deposit rate it offers. The optimal deposit rate maximizes the profits of the bank and the first-order condition becomes

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial (1 + r_D)} &= -\pi_i D + \pi_i \left(\left(1 + r_L^i\right)(1 + \kappa) - (1 + r_D)\right) \frac{\partial D}{\partial (1 + r_D)} \\ &= \pi_i \left(\left(1 + r_L^i\right)(1 + \kappa) \frac{\varepsilon}{1 + r_D} - (1 + \varepsilon)\right) D = 0. \end{aligned} \tag{34.53}$$

From this first-order we obtain the optimal the optimal deposit rate as

$$1 + r_D^{**} = \frac{\varepsilon}{1 + \varepsilon} \left(1 + r_L^i\right)(1 + \kappa) \tag{34.54}$$

Inserting the optimal deposit rate back into the bank profits of equation (34.51) we obtain

$$\Pi_B^i = \left(\frac{1 + \kappa}{1 + \varepsilon} \pi_i \left(1 + r_L^i\right) - (1 + r_E)\kappa\right) D. \tag{34.55}$$

The optimal capital ratio, κ would be derived from its first order condition

$$\frac{\partial \Pi_B^i}{\partial \kappa} = \frac{\pi_i \left(1 + r_L^i\right)}{1 + \varepsilon} - (1 + r_E) < 0, \tag{34.56}$$

where the inequality arises from our assumption that $1 + r_E > \pi_H \left(1 + r_L^H\right) \leq \pi_L \left(1 + r_L^L\right)$ and $\varepsilon > 0$. Hence without regulation, no capital would be held by banks. Furthermore, replacing in constraint (34.52) for the bank choosing the low-risk looan the deposit rate by the optimal deposit rate from equstion (34.54), we get that the low risk loans is chosen only if

$$\varepsilon < \frac{\pi_H \left(1 + r_L^H\right) - \pi_L \left(1 + r_L^L\right)}{\pi_L \left(r_L^L - r_L^H\right)}. \tag{34.57}$$

Thus the low-risk loan is selected on if the elasticity of demand for deposits is sufficiently low; this might be the case if competition between banks for deposits is not very strong. This result is independent of any capital requirements and thus a regulator could not ensure that banks choose the low-risk loan. Hence to ensure the condition in equation (34.52) is fulfilled and low-risk loans are chosen, a regulator would have to complement capital requirements by a limit on deposit rates. Maintaining the profits of banks such that they are indifferent between capital requirements and limits on the deposit rate, we get by totally differentiating equation (34.55) that

$$d\Pi_B^H = \left(\pi_H \left(1 + r_L^H\right) - (1 + r_E)\right) D d\kappa - \pi_H D d\left(1 + r_D\right) = 0, \tag{34.58}$$

from which we easily obtain that

$$\frac{\partial \left(1 + r_D\right)}{\partial \kappa} = \frac{\pi_H \left(1 + r_L^H\right) - (1 + r_E)}{\pi_H} < 0. \tag{34.59}$$

The negative sign of this relationship indicates that capital requirements and restrictions on deposit rates are substitutes.

Thus capital requirements may not be sufficient to ensure that banks choose to provide low-risk loans but have to be complemented by limits on the deposit rate that a bank is allowed to pay. This limit on the deposit rate will depend on the capital requirements and hence banks can increase the capital ratio required, reducing the leverage, and in turn allow banks to pay higher deposit rates to benefit depositors. Without a limit on the deposit rate, the bank will choose the low-risk loan only if the elasticity of demand for deposits is sufficiently low; this cannot be controlled by the regulator and for this reason in times where this elasticity is high, limits on deposit rates might become necessary. Rather than imposing an explicit limit on deposit rates, regulators might allow, implicitly or explicitly, banks to collude in setting lower deposit rates to achieve the same aim.

Reading Hellmann, Murdock, & Stiglitz (2000)

34.3.3 The impact of deposit insurance

A main concern that leads to the introduction of capital requirements is that banks have an incentive to take on higher risks than is socially optimal as they do not have to cover the full losses to depositors if they fail; capital requirements are then supposed to align the risks of the bank with the social optimum. We will here see how an adequately priced deposit insurance would lead to banks adopting a capital ratio that is socially optimal and how, if this cannot be achieved as deposit insurance is not priced, socially optimal capital requirements can be determined.

Let us assume the probability π of loans L being repaid with interest r_L is not known, but we know that it follows a log-normal distribution; we can thus value the deposit insurance for deposits D, including interest r_D, as a put option. Denoting the variance of the repayment rate π by σ and $\Phi(\cdot)$ the standard normal distribution, we get as the value of deposit insurance that

$$\begin{aligned} P &= (1 + r_D) D\Phi(d_2) - \pi (1 + r_L) L\Phi(d_1), \\ d_1 &= \frac{1}{\sigma} \ln \frac{(1 + r_D) D}{(1 + r_L) L} - \frac{1}{2}\sigma, \\ d_2 &= d_1 + \sigma, \end{aligned} \tag{34.60}$$

as shown in chapter 18.1.2. With a capital ratio κ we have $E = \kappa D$ and $L = D + E = (1 + \kappa) D$. Using capital costs r_E sufficiently high to exceed the expected returns on the provision of loans, $1 + r_E > \pi(1 + r_L)$, we have the profits for banks paying the deposit insurance premium as in equation (34.60), given by

$$\begin{aligned} \Pi_B &= \pi ((1 + r_L) L - (1 + r_D) D) - (1 + r_E) E - P \\ &= (\pi ((1 + r_L)(1 + \kappa) - (1 + r_D)) - (1 + r_E)\kappa) D - P. \end{aligned} \tag{34.61}$$

The optimal capital ratio is then obtained by maximizing the profits of banks, the first order condition of which requires

$$\frac{\partial \Pi_B}{\partial \kappa} = (\pi (1 + r_L) - (1 + r_E)) D - \frac{\partial P}{\partial \kappa} = 0. \tag{34.62}$$

The first term of this expression is negative from our assumption that $1 + r_E > \pi(1 + r_L)$ and the second term is negative as in chapter 18.1.2 it was shown that the premium increases in the leverage, thus reduces with higher capital ratios. Therefore an internal solution for the optimal capital ratio can be found.

If the deposit insurance is provided for free or by a premium not responsive to the capital ratio, then the final term in equation (34.62) would vanish and the bank would prefer a capital ration of zero and hence not provide any equity. As in most cases deposit insurance is either not priced or not adequately priced to take into account all risks, a regulator could use the value of deposit insurance to implement the resulting capital ratio that would require banks to internalise the costs of their deposit insurance by solving equation (34.62) for the optimal capital ratio κ.

If deposit insurance were priced appropriately taking into account the risks of the bank, including its leverage $\frac{1}{\kappa}$, the bank would internalise the costs imposed on the general public from the loss of deposits and no additional capital requirements were needed. In the absence of such deposit insurance, a regulator could impose adequate capital requirements to achieve the same risk profile of banks.

Reading Freixas & Gabillon (1999)

Résumé

We have seen that capital requirements at times might not be effective in achieving the desired outcome, especially if the intention of these requirements is to direct the banks to the provision of less risky loans. It is therefore important to note that while capital requirements can be used to some extend to align the optimal choice of banks with the social optimum, this might not be the case, or only in certain circumstances, in some instances and we need to supplement capital requirements by other regulatory measures. this adds to the complexity of banking regulation and imposes additional restrictions on the decisions of banks.

34.4 Regulatory regimes

Capital requirements and other restrictions imposed on banks are not the only tool for regulators to achieve the desired outcome of a banking system where banks take low risks. We will look here on the wider regulatory regime that is applied by regulators. Specifically, in chapter 34.4.1 we will look at how the licensing of banks can affect capital requirements and chapter 34.4.2 will determine how the use us risk-based capital regulation reduces capital requirements relative to a non-risk based leverage ratio.

34.4.1 Regulated market entry

Capital requirements are often used to limit the risks that banks are taking, however, regulators complement such requirements with other measures. Most notably is that banks require a licence to operate and this licence is only available to banks that can show to have sufficient knowledge and skills, besides sufficient capital. We will here look at the impact such a regulation of market entry can have on the capital requirements of those banks that are granted a licence.

Let us consider an economy in which N individuals, have funds of D_i each. Each individual can open a bank, using D_i as equity, or deposit into other banks. A fraction λ of these individuals have the ability to monitor loans they give at some cost c; this monitoring improves the success rate of loans from π_L to $\pi_H > \pi_L$, such that with r_L denoting the loan rate we assume that monitoring is efficient such that

$$(\pi_H - \pi_L)(1 + r_L) > c. \tag{34.63}$$

The benefits of the repayment rate of the loan increasing from π_L to π_H is exceeding the monitoring costs.

We define the capital ratio as $\kappa = \frac{E}{D}$, where E denotes equity and D the total deposits of a bank. Of course, we have $E = D_i$ as this was the initial; funds available to the owner of a bank. The total loans a bank can give are given by $L = D + E = (1 + \kappa) D$. Deposits are obtained from all individuals that do not act as a bank. With M of such banks, we have the total deposits in the banking system given by $\hat{D} = (N - M) D_i$ and hence the deposits of each bank are determined as $D = \frac{\hat{D}}{M} = \frac{N-M}{M} D_i$. From $D_i = E = \kappa D$, we get can write the deposits per bank as

$$D = \frac{N - M}{M} \kappa D$$

and hence $\frac{N-M}{M} = \kappa$ or

$$\begin{aligned} M &= \frac{\kappa}{\kappa + 1} N, \\ D &= \frac{1}{\kappa} D_i \end{aligned} \tag{34.64}$$

for the number of banks and the amount of deposits, respectively.

As monitoring of bank loans reduces the risks of banks, we will seek to establish a banking system in which banks are only operated by those able and willing to monitor. We will first look at the case where banks are entirely unregulated and anyone can open a bank.

Unregulated market entry With monitoring desirable, we would want those who are able to monitor loans to do so, hence with deposit rate r_D we need that the profits of doing so and achieving a high repayment rate π_H, at cost cL, exceeds the profits if not monitoring and only achieving the low repayment rate π_L. We thus require

$$\pi_H \left((1 + r_L) L - (1 + r_D) D \right) - cL \geq \pi_L \left((1 + r_L) L - (1 + r_D) D \right),$$

which solves for

$$1 + r_D \leq 1 + r_D^* = (1 + \kappa)(1 + r_L) - \frac{c (1 + \kappa)}{\pi_H - \pi_L} \tag{34.65}$$

after inserting for $L = D + E$ and noting the expressions for D and E from above.

Similarly, banks need to be willing to take deposits rather than only investing their own funds directly into loans. Thus the profits of a bank monitoring its loans must be higher than the bank only investing its own funds D_i directly into loans and monitoring these; this requires

$$\pi_H \left((1 + r_L) L - (1 + r_D) D \right) - cL \geq \pi_H (1 + r_L) D_i - cD_i,$$

from which we easily obtain that

$$1 + r_D \leq 1 + r_D^{**} = (1 + r_L) - \frac{c}{\pi_H}. \tag{34.66}$$

So far we have established conditions under which banks are willing to accept deposits and willing to monitor loans, but we also need establish that banks exist that are willing to conduct this monitoring. An individual who is not able to monitor should use his funds to provide a deposit to a bank and not open a bank himself. If this individual were to open a bank, the number of banks, M, would increase to $M + 1$ and hence deposits per bank become

$$D^* = \frac{N - (M - 1)}{M + 1} D_i = \frac{N - (\kappa + 1)}{\kappa N + (\kappa + 1)} D_i, \tag{34.67}$$
$$L^* = \frac{M}{M + 1} L = \frac{\kappa N}{\kappa N + (\kappa + 1)} L = \frac{(\kappa +) N}{\kappa N + (\kappa + 1)} D_i.$$

The profits of operating a bank with these loans and deposits would need to be less attractive than providing a deposit to another bank. As the individual has no ability to monitor loans, he will operate a bank with a low repayment rate π_L; if he were to provide a deposit to another bank he would be repaid his deposits if that bank does not fail, which happens with probability π_H as the bank will monitor its loans. We thus need

$$\pi_L \left((1 + r_L) L^* - (1 + r_D) D^* \right) \leq \pi_H (1 + r_D) D_i.$$

This condition solves for

$$1 + r_D \geq 1 + r_D^{***} = \frac{(1 + \kappa) \pi_L (1 + r_L) N}{(1 + \kappa) (\pi_H - \pi_L) + N (\pi_L + \kappa \pi_H)}. \tag{34.68}$$

Finally, for individuals not able to monitor, we require that bank deposits are more attractive than giving the loan directly to ensure deposits do exist. Noting that due to its inability to monitor loans, the loan provided directly would have the low repayment rate π_L, we require

$$\pi_H (1 + r_D) D_i \geq \pi_L (1 + r_L) D_i, \tag{34.69}$$

hence

$$1 + r_D \geq 1 + r_D^{****} = \frac{\pi_L}{\pi_H} (1 + r_L). \tag{34.70}$$

We can now show easily using equation (34.63) that $1 + r_D^{**} > 1 + r_D^{****}$ and furthermore

$$\frac{\partial \left(1 + r_D^{***}\right)}{\partial \kappa} = -\frac{(\pi_H - \pi_L) \pi_L (1 + r_L) N^2}{\left((\kappa + 1) (\pi_H - \pi_L) + N (\pi_L \kappa \pi_H)\right)^2} \tag{34.71}$$
$$< 0$$

and

$$\lim_{\kappa \to +\infty} 1 + r_D^{***} = \frac{N}{(\pi_H - \pi_L) + \pi_H N} \pi_L (1 + rL) < 1 + r_D^{****}. \qquad (34.72)$$

Figure 34.4 illustrates these constraints and area denoted $\mathfrak{A}$ denotes the possible capital ratios that banks can adopt such that only banks are only operated that are can monitor, are actually willing to monitor, accept deposits, and individuals without monitoring ability provide deposits. As long as $1 + r_L > 1 + r_D$ banks will seek to obtain the minimal capital ratio that is feasible as we can easily see that profits are reducing as the capital ratio reduces, leading to a capital ratio of banks at κ^* as indicated. Of course, the other parameters must such that a feasible region remains. For example, if the monitoring costs c are too high, the constraint $1 + r_D^*$ moves downwards and no feasible area remains if it crosses $1 + r_D^{***}$ below $1 + r_D^{****}$; in this case banks would not exist that monitor loans.

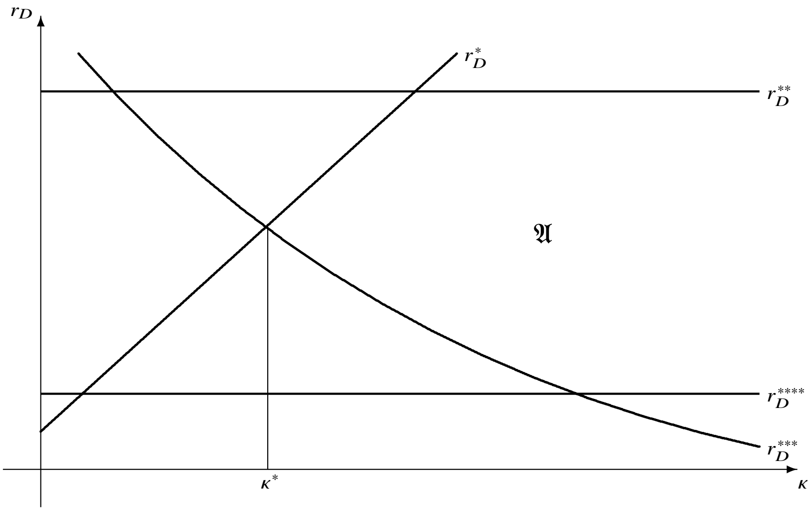


Fig. 34.4: Capital ratio of unregulated banks

In order to ensure that only banks able and willing to monitor loans are operating, and thus keeping risks low, we would require a capital ratio of at least κ^*. Any lower capital ratio would lead to a banking system that is not meeting these requirements. We could thus impose capital requirements of at least κ^* to ensure the banking system has the desired properties; higher capital requirements do not bring any benefits.

Regulators often do more than merely impose capital requirements but will also licence banks and only those deemed to have the requisite ability are allowed to operate a bank. We will look at the impact such screening of banks on market entry has on the capital requirements of banks.

Screening of banks We now assume that the regulator imposes minimum capital requirements in form of a minimum capital ratio κ, then the incentives for the banks in equations (34.65) and (34.66) remain unchanged, thus the willing ness of banks to monitor and their willingness to accept deposits are unaffected. For individuals without the ability to monitor loans, the profits from opening a bank are

$$\begin{aligned} \Pi_B &= \pi_L \left(\left(1 + r_L\right) L - \left(1 + r_D\right) D \right) \\ &= \pi_L \left(\left(1 + \kappa\right) \left(1 + r_L\right) - \left(1 + r_D\right) \frac{1}{\kappa} \right) D_i \end{aligned} \tag{34.73}$$

With a minimum capital ratio of κ and fixed capital D_i, deposits might have to be rationed as banks may not be able to accommodate all deposits that are available. The total deposit demand would be from all those individuals that have not opened a bank, $(N - M)D_i$, and we assume that a fraction λ of individuals are opening a bank, $M = \lambda N$. The total deposits that can be accommodated by the banks is MD, where $D = \frac{D_i}{\kappa}$, given that $E = D_i$. Hence the fraction of deposits that banks can accept are given by

$$\gamma = \frac{MD}{\left(N - M\right) D_i} = \frac{\lambda}{1 - \lambda} \frac{1}{\kappa}, \tag{34.74}$$

where the numerator denotes the deposits banks are able to accept and the denominator the amount of deposits available. Therefore, the deposits that can be deposited into banks are a fraction γ of the funds available to each depositor, $D_i^* = \gamma D_i$.

Let us now propose that the regulator in addition can identify all individuals able to monitor loans and grant them a licence, such that there are $M = \lambda N$ able banks only, representing exactly the fraction of individuals that have the ability to monitor. This regulator can identify individuals perfectly with probability p, but with probability $1 - p$ the regulator cannot identify individuals able to monitor loans and will grant $M = \lambda N$ bank licences randomly. In this case the depositor makes returns of

$$\begin{aligned} \Pi_D = {} & p\pi_H \left(1 + r_D\right) D_i^* \\ & + \left(1 - p\right) \left(\lambda \pi_H \left(1 + r_D\right) D_i^* \right. \\ & \left. + \left(1 - \lambda\right) \pi_L \left(1 + r_D\right) D_i^* \right) - D_i^*, \end{aligned} \tag{34.75}$$

where the first term denotes the regulator identifying able individuals, who then produce repayment rates of π_H, and the second term regulators not identifying them, such that depositors face a fraction λ of banks able to monitor loans and hence having a high repayment rate and a fraction $1 - \lambda$ of banks not able to monitor loans and facing a low repayment rate π_L. To ensure that in this case only banks exist that are able to monitor loans, we need that the profits of being depositor exceeds that of operating a bank without the ability to monitor loans, $\Pi_D \leq \Pi_B$, which using equations (34.73) and (34.75) becomes

$$1 + r_D \geq 1 + \hat{r}_D^{***} = \frac{\left(1 + \kappa\right) \pi_L \left(1 + r_L\right) + \frac{\lambda}{1-\lambda}}{1 + \frac{\lambda}{1-\lambda} \left(p\pi_H + \left(1 - p\right) \left(\lambda \pi_H \left(1 - \lambda\right) \pi_L\right)\right)}. \tag{34.76}$$

We see that the ability of the regulator to identify individuals able to monitor loans, p affects the attractiveness of deposits and the more able the regulator is in identifying them, the less strict this constraint becomes.

As before, the bank deposit must be more attractive to an individual unable to monitor the loan than giving the loan directly. Here the bank deposit will be rationed to D_i^* and the depositor will not earn interest on his entire funds, but would be able to provide a loan for the full funds. Thus we require

$$\pi_H (1 + r_D) D_i^* \geq \pi_L (1 + r_L) D_i,$$

or

$$1 + r_D \geq \hat{r}_D^{*****} = \frac{\pi_L}{\pi_H} \frac{1 - \lambda}{\lambda} \kappa (1 + r_L). \tag{34.77}$$

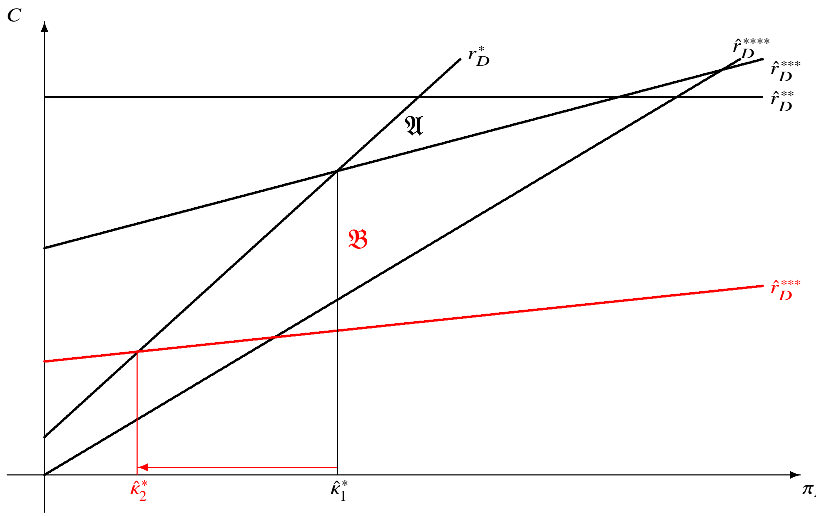


Fig. 34.5: Minimum capital requirements

Figure 34.5 illustrates these constraints. We see that in contrast to the case of unregulated market entry, corresponding to $p = 0$ and $\gamma = 1$ and hence $\lambda = \frac{\kappa}{1+\kappa}$, the conditions $\hat{r}_D^* * * *$ and $\hat{r}_D^* * * **$ are now having a positive slope, rather than the slope being negative or zero, respectively. This implies that in general the capital ratio that will ensure a banking system in which loans are monitored have to less strict; banks require less equity or can have a higher leverage. This is because now the regulator can identify those individuals able to monitor loans and exclude those from operating a bank that do not have this ability, the reliance on incentives to ensure that only that able to monitor loan are operating a bank is reduced. This allows the capital requirements to be loosened by allowing for a lower capital ratio or a higher leverage of banks. The feasible capital ratios are given in area $\mathfrak{A}$; it is now not possible for a regulator to impose high capital requirements as the regulatory constraint would not allow all deposits to be accepted and in this case individuals might prefer to provide loans directly, which due to their inability to monitor it, is not desirable.

We see, as illustrated in figure 34.5 that as p increases and hence the regulator being more likely to identify individuals able to monitor loans, the slope of the restriction $1 + \hat{r}_D^{***}$ reduces and is shifted downwards, which can easily be confirmed from equation (34.76) with the denominator increasing in p. Consequently the capital requirements become less onerous as indicated by κ_2^*. This is as the reliance on incentives are becoming less and important given the identification of such individuals by the regulator; the area of feasible capital ratios is extended from area $\mathfrak{A}$ to include also area $\mathfrak{B}$.

Summary We have seen that regulating market entry by only allowing banks to operate that have the requisite skills to maintain low risks when granting loan reduces the capital requirements on banks. By excluding banks that are not able to reduce risks, regulators have to rely less on capital requirements to exclude these banks, but they do so by not granting a licence; this allows for lower capital requirements. There is a trade-off between low regulatory inference in terms of granting a bank licence, corresponding to a low p, and a high capital requirements on the one hand and strict requirements for a bank licence combined with low capital requirements. We can interpret that capital requirements and the stringency of the rules to obtain a banking licence are substitutes.

Reading Morrison & White (2005)

34.4.2 Risk-based capital requirements

Capital requirements are commonly based on the risks that banks are exposed to, but regulators also make use of a generic leverage ratio, which is not based on the risks the bank is exposed to. We will now investigate whether it is more beneficial for banks and depositors to have a risk-based capital requirement or whether a generic leverage ratio that does not consider risks is preferable.

Let us assume that a loan L is repaid with probability π , including interest r_L, and with probability $1 - \pi$ it yields only a fraction $\hat{\lambda} < 1$ of the original amount lent; this would leave the bank unable to repay depositors their funds D including the applicable interest r_D. If a loan is not repaid in full, the repayment $\hat{\lambda}$ can either be $\hat{\lambda} = \lambda$ with probability p or $\hat{\lambda} = 0$ with probability $1 - p$. In the former case the bank manages to recover some of the loan as the borrower defaults and in the latter case banks cannot recover anything. We assume that $\pi(1 + r_D) < 1$ and hence without recovery depositors would not provide funds to the bank as they cannot expect to receive the required repayment.

Loan are financed by deposits D and equity E such that $L = D + E$; we define a capital ratio κ such that $E = \kappa D$. If the loan is not repaid, depositors obtain the recovered amount $\hat{\lambda}L$ and if we assume that the value of $\hat{\lambda}$ is known to depositors, we get in the only relevant case of $\hat{\lambda} = \lambda$ that depositors break even if

$$\pi (1 + r_D) D + (1 - \pi) \lambda L \geq D. \tag{34.78}$$

If the loan is repaid to the bank, depositors are repaid fully, including interest, and if the loan is not repaid to the bank, depositors will be paid the amount recovered from the loan; This payment to depositors must exceed the initial investment to be attractive to depositors. Inserting for $L = E + D + (1 + \kappa) D$, we can solve this for the capital ratio required as

$$\kappa \geq \kappa^* = \frac{1 - \pi (1 + r_D)}{\lambda (1 - \pi)} - 1. \tag{34.79}$$

Thus deposits will be available to banks who are able to recover some of the defaulted loan if the capital requirements are sufficiently high.

Let us now assume that we do not know the loan recovery $\hat{\lambda}$, but receive a signal s about its value. This signal has precision ρ such that

$$Prob \left(s = \lambda | \hat{\lambda} = \lambda\right) = Prob \left(s = 0 | \hat{\lambda} = 0\right) = \rho > \frac{1}{2}. \tag{34.80}$$

We can interpret ρ as probability of the signal being correct, thus its accuracy; at $\rho = \frac{1}{2}$ the signal does not provide any additional information and at $\rho = 1$ it provides perfect information. Using Bayes' Rule we easily get

$$\begin{aligned} p_\lambda &= Prob \left(\hat{\lambda} = \lambda | s = \lambda\right) = \frac{\rho p}{\rho p + (1 - \rho)(1 - p)}, \\ p_0 &= Prob \left(\hat{\lambda} = \lambda | s = 0\right) = \frac{(1 - \rho) p}{p (1 - \rho) + \rho (1 - p)} \end{aligned} \tag{34.81}$$

where we can easily verify that $p_\lambda > p_0$.

We can now determine the optimal capital requirements if we assume that the regulator does not take into account information on the recovery of defaulting loans.

Leverage ratio A regulator could now decide to ignore any signal and decide on a common capital requirement for all banks, regardless of their ability to recover defaulting loans. With depositors having no access to the signal on loan recoveries, they can only use the expected recovery rate $p\lambda$ to assess whether they seek to provide funds to the bank; thus replacing λ in equation (34.78) with $p\lambda$, the capital requirements become

$$\kappa \geq \kappa_p^* = \frac{1 - \pi_H (1 + r_D)}{p\lambda (1 - \pi_H)} - 1. \tag{34.82}$$

As the capital requirements are not based on the actual risk of the bank, represented through the recovery rate on defaulting loans, they are akin to a general leverage ratio $\frac{1}{\kappa_p^*}$ imposed in banks. We assume that banks are subjected to the lowest possible capital requirements and hence condition (34.82) is fulfilled with equality.

The expected shortfall to depositors if the bank defaults, is then given by

$$\Pi_D = p\lambda L - (1 + r_D)\, D = \left(p\lambda \frac{1 + \kappa_p^*}{\kappa_p^*} - \frac{1 + r_D}{\kappa_p^*} \right) E. \tag{34.83}$$

The first term denotes the expected payments the depositors can expect from the bank and the second term represents the payment depositors were expecting to obtain. We express this shortfall in terms of the equity of the bank.

We can now compare the use of a common leverage ratio with that of capital requirements which are based on the risk of the bank.

Risk-based capital requirements If a regulator sets capital requirements based on the signal they receive on the ability of banks to recover loans, depositors can infer this signal from these capital requirements. Hence for capital requirements consistent with signal $s = \lambda$ ($s = 0$), depositors will replace λ in equation (34.78) with $p_\lambda \lambda$ ($p_0 \lambda$). Here $p_\lambda \lambda$ ($p_0 \lambda$) represent the probability of the recovery rate being $\hat{\lambda} = \lambda$ ($\hat{\lambda} = 0$), given the signal the bank has received; these probabilities were derived in equation (34.81). We thus get the respective capital requirements as

$$\begin{aligned}
\kappa \geq \kappa_\lambda^* &= \frac{1 - \pi_H (1 + r_D)}{p_\lambda \lambda (1 - \pi_H)} - 1 \\
&= \frac{Prob\,(s - \lambda)\,(1 - \pi_H\,(1 + r_D))}{\rho p \lambda\,(1 - \pi_H)} - 1, \\
\kappa \geq \kappa_0^* &= \frac{1 - \pi_H (1 + r_D)}{p_0 \lambda (1 - \pi_H)} - 1 \\
&= \frac{Prob\,(s = 0)\,(1 - \pi_H\,(1 + r_D))}{(1 - \rho)\, p \lambda\,(1 - \pi_H)} - 1,
\end{aligned} \tag{34.84}$$

where $Prob\,(s = \lambda) = \rho p + (1 - \rho)(1 - p)$ and $Prob\,(s = 0) = \rho\,(1 - p) + p\,(1 - \rho)$ denote the probabilities of observing the respevtive signals. We can easily verify that $Prob\,(s = \lambda) > Prob\,(s = 0)$ if $\rho > \frac{1}{2}$. Using these results, it is straightforward to show that

$$\kappa_\lambda^* \leq \kappa_p^* \leq \kappa_0^* \tag{34.85}$$

and

$$\begin{aligned}
1 + \kappa_\lambda^* &= \frac{Prop\,(s = \lambda)}{\rho} \left(1 + \kappa_p^*\right), \\
1 + \kappa_0^* &= \frac{Prob\,(s = 0)}{1 - \rho} \left(1 + \kappa_p^*\right), \\
\lim_{\rho \to \frac{1}{2}} \left(1 + \kappa_\lambda^*\right) &= \lim_{\rho \to \frac{1}{2}} \left(1 + \kappa_0^*\right) = 1 + \kappa_p^*, \\
\lim_{\rho \to 1} \left(1 + \kappa_\lambda^*\right) &= p \left(1 + \kappa_p^*\right), \\
\lim_{\rho \to 1} \left(1 + \kappa_0^*\right) &= +\infty.
\end{aligned} \tag{34.86}$$

The expected shortfall for depositors if the bank fails is then given by

$$\begin{aligned}
\hat{\Pi}_D &= Prob\left(s=\lambda\right)\left(p_\lambda \lambda L - \left(1+r_D\right)D\right) \\
&\quad + Prob\left(s=0\right)\left(-\left(1+r_D\right)D\right) \\
&= \left(\rho p \lambda \frac{1+\kappa_\lambda^*}{\kappa_\lambda^* - \left(\frac{Prob(s=\lambda)}{\kappa_\lambda^*} + \frac{Prob(s=0)}{\kappa_0^*}\right)\left(1+r_D\right)}\right) E.
\end{aligned} \tag{34.87}$$

The first term captures the case where the signal suggest that the bank can recover the loan and the second term the case where this is not the case. Using the result in equation (34.86) and comparing with equation (34.83) we obtain that

$$\begin{aligned}
\lim_{\rho \to \frac{1}{2}} \hat{\Pi}_D &= \left(\frac{1}{2} p \lambda \frac{1+\kappa_p^*}{\kappa_p^*} - \frac{1+r_D}{\kappa_p^*}\right) E \\
&< \Pi_D \\
\lim_{\rho \to 1} \hat{\Pi}_D &= \left(p \lambda \frac{p\left(1+\kappa_p^*\right)}{p\left(1+\kappa_p^*\right)-1} - \frac{p}{p\left(1+\kappa_p^*\right)-1\left(1+r_D\right)}\right) E \\
&> \Pi_D
\end{aligned} \tag{34.88}$$

as we can easily verify that $\frac{p\left(1+\kappa_p^*\right)}{p\left(1+\kappa_p^*\right)-1} > \frac{\kappa_p^*-1}{\kappa_p^*}$ and $\frac{p}{p\left(1+\kappa_p^*\right)-1} < \frac{1}{\kappa_p^*}$. From these two equations we see that there must exist some $\frac{1}{2} < \rho^* < 1$ such that for $\rho < \rho^*$ the losses to depositors are smaller if a regulator chooses the leverage ratio κ_p^* and ignores the risk of the banks; only for more accurate signals of $\rho > \rho^*$ should they base their capital requirements on risks as this reduces the losses to depositors. The reason for this result is that in the case the signal has low accuracy, a regulator would impose high capital requirements on banks who are able to recover payments from a defaulting loan, even though their confidence that they identified the bank type correctly is low.

Using that $Prob\left(s=\lambda\right) < \rho$ and $Prob\left(s=0\right) < 1-\rho$, we can determine the average capital requirements of banks if the regulator uses the signal on the bank type; we obtain

$$\begin{aligned}
& Prob\left(s=\lambda\right)\left(1+\kappa_\lambda^*\right) + Prob\left(s=0\right)\left(1+\kappa_p^*\right) \\
&= \frac{Prob\left(s=\lambda\right)^2}{\rho}\left(1+\kappa_p^*\right) + \frac{Prob\left(s=0\right)^2}{1-\rho}\left(1+\kappa_p^*\right) \\
&< \rho\left(1+\kappa_p^*\right) + \left(1-p\right)\left(1+\kappa_p^*\right) \\
&= 1+\kappa_p^*
\end{aligned} \tag{34.89}$$

We see that if a regulator uses its signal to determine the capital requirements of banks, the average capital requirements will be lower. This is because banks able to recover defaulting loans will no longer be subjected to excessive capital requirements; this weighs more than banks not able to recover defaulting loans having too lenient capital requirements imposed. The losses to depositors will also be lower with such risk-based capital requirements as long as the signal the regulator receives is sufficiently accurate.

Summary We have seen that risk-based capital requirements lead to lower capital requirements overall than using generic limits on the leverage of banks that is not founded on the risks that banks take. Such an approach is also beneficial to depositors as their losses from a failing bank are lower as long as the information the regulator has is sufficiently precise. It is therefore that capital requirements that are taking the risks of banks into account, such as using risk-weighted assets, will be beneficial to banks and depositors alike. Apart from being seen as 'fair' to require banks taking more risks to have higher capital requirements, it is reducing the overall costs from capital requirements in the banking system and reduces losses to depositors if a bank fails. For this reason risk-based capital requirements will be socially preferable to a leverage ratio.

Reading Ahnert, Chapman, & Wilkins (2021)

Résumé

We have seen that the licensing of banks can reduce the capital requirements of banks to ensure they pursue a low-risk lending policy. Similarly, basing capital requirements on the actual risks banks have taken rather than imposing a generic restriction identical for all banks, will reduce the capital requirements overall, although not for all banks. Imposing such additional constraints on banks such as licensing or requiring them to provide sufficient information to regulators such that risk-based capital requirements can be applied can be beneficial to banks in that they reduce their capital requirements. These measures will, however, increase the complexity of the regulatory burden and while costs may reduce due to lower capital requirements, they will most likely increase elsewhere to ensure compliance with the additional requirements such regulations impose on banks.

34.5 Bailing out banks

Banks are often not outright failing, but they are bailed out, either by other banks, or by governments. Such bailouts will obviously affect the incentives of banks to provide risky loans as the losses banks incur from such loans not being repaid will be smaller with a bailout than without a bailout. These changed incentives to provide loans with higher risks will then need to be addressed by regulators through capital requirements. Therefore, in chapter 34.5.1 we will investigate the implications of such bailouts for capital requirements and chapter 34.5.2 will then look at the influence capital requirement can have on the contagion of bank failures.

34.5.1 Risk reduction in bailed out banks

Banks commonly expect to be bailed out, either by being bought up by other banks or through government interventions; this is mainly done to protect depositors from losses. This expectation of a bail-out will therefore affect the deposit rate banks are having to pay; with a bail-out such deposit rates should be reduced as depositors face less risks. As deposit rates are no longer responding to the risks bank are taking, banks have an incentive to increase risks. We will investigate here how capital requirements need to be adjusted such that banks do not take higher risks id depositors anticipate that they will be bailed out.

Let us assume that loans L are financed by deposits D and equity E such that $L = D + E$ and the capital ratio κ is defined such that $E = \kappa D$ and hence $L = (1 + \kappa) D$. Loans are repaid with probability π, including interest r_L or with probability $1 - \pi$ the bank can only recover a fraction λ of the original loan and will subsequently fail. We will be interested to determine the capital requirements that need to be imposed such that banks that are bailed out do not take on more risks than banks that are not bailed out. To this end we first determine the optimal level of risk a bank that is not bailed out would take.

No bailouts With a deposit rate of r_D we get the profits of depositors as

$$\Pi_D = \pi (1 + r_D) D + (1 - \pi) \lambda L - D, \tag{34.90}$$

where the first term denotes the case where the bank's loan is repaid and hence depositors paid in full. The second term covers the case where loans are not repaid and depositors only receive the amount recovered by the bank; from this we have to subtract their initial investment. If we assume that banks can exyract all benefits from depositors such that $\Pi_D = 0$ and using that $L = (1 + \kappa) D$, we get the deposit rate as

$$1 + r_D = \frac{1 - (1 - \pi) \lambda (1 + \kappa)}{\pi}. \tag{34.91}$$

The profits of banks are determined by

$$\begin{aligned} \Pi_B &= \pi ((1 + r_L) L - (1 + r_D) D) \\ &= ((1 + \kappa) (\pi (1 + r_L) + (1 - \pi) \lambda) - 1) D. \end{aligned} \tag{34.92}$$

We now assume that the probability of loans being repaid is decreasing in the amount lent and hence for a given amount of equity we have $\frac{\partial \pi}{\partial \kappa} = \frac{\partial \pi}{\partial L} \frac{\partial L}{\partial \kappa} = \frac{\partial \pi}{\partial L} E < 0$. We can justify this assumption by the assertion that as a bank provides more and more loans it will become more and more difficult to find borrowers of low risk and thus the bank has to include more and more risky borrowers. We furthermore assume that $\frac{\partial^2 \pi}{\partial \kappa^2} < 0$ and this effect is increasing as lending increases as banks find it ever more difficult to identify low-risk borrowers.

Banks will maximize their profits by selecting their optimal capital ratio, leading to the first order condition

$$\frac{\partial \Pi_B}{\partial \kappa} = \left(\pi (1 + r_L) + (1 - \pi) \lambda + (1 + \kappa) (1 + r_L - \lambda) \frac{\partial \pi}{\partial \kappa}\right) D \quad (34.93)$$
$$= 0,$$

which solves for

$$\frac{\partial \pi}{\partial \kappa} = -\frac{\pi (1 + r_L) + (1 - \pi) \lambda}{(1 + \kappa) (1 + r_L - \lambda)}. \tag{34.94}$$

Solving this expression for the capital ratio κ will give us the capital ratio that is optimal for banks. It similarly would gives us the optimal level of risk, π for a given capital ratio. hence there is a clear relationship between the capital ratio and the risk a bank takes and we can thus associate the same value of $\frac{\partial \pi}{\partial \kappa}$ for a given capital ration κ with the same risk π.

Thus far we considered the case where failing banks are not bailed out to use as a benchmark that bailed out banks should match. We will thus now turn our attention to the case where such a bailout occurs.

Bailed-out banks If failing banks are bailed out such that depositors are fully compensated, they cannot incur any losses and hence their return will be the risk-free rate, which we for simplicity assume to be zero here, thus $r_D = 0$. Using this results the bank profits from equation (34.92) become

$$\Pi_B = \pi \left((1 + r_L) (1 + \kappa) - 1\right) D \tag{34.95}$$

and the optimal capital ratio is derived from the first order condition

$$\frac{\partial \Pi_B}{\partial \kappa} = \left(\pi (1 + r_L) + \frac{\partial \pi}{\partial \kappa} \left((1 + r_L) (1 + \kappa) - 1\right)\right) D = 0, \tag{34.96}$$

which can be solved for

$$\frac{\partial \pi}{\partial \kappa} = -\frac{\pi (1 + r_L)}{(1 + \kappa) (1 + r_L) - 1}. \tag{34.97}$$

Comparing equations (34.94) and (34.97), we see that if the right-hand side in equation (34.97) is smaller (in absolute terms) than in equation (34.94), the capital ratio with bailouts is smaller. This is because the marginal product on the left-hand side is decreasing as the capital ratio increases; a smaller value thus corresponds to a lower capital ratio. This condition is fulfilled if

$$\lambda < \lambda^* = \frac{\pi (1 + r_L) + (1 - \pi) \lambda}{(1 + r_L) (1 + \kappa)}. \tag{34.98}$$

Hence for sufficiently large losses when recovering defaulting loans, the capital ratio is lower than for a bank that is not bailed out. The fact that deposit rates are not increasing as the risk the bank takes increases, will induce the bank to seek higher risks, which manifests itself in a lower value for the repayment rate π, and thus a higher default risk, for a given capital ratio or a lower capital ratio for a given risk. In either interpretation, the risk of the bank increases.

A regulator might now seek to implement the same capital ratio as it was established without a bailout in equation (34.94); the same capital ration would then imply that it is optimal to choose the same risks as a bank that has no prospect of baing bailed out. This can be achieved if the right-hand sides of both expressions are identical. This then solves for

$$1+\kappa^* = \frac{\pi\left(1+r_L\right)+\lambda\left(1-\pi\right)}{\lambda\left(1+r_L\right)}. \tag{34.99}$$

In order to achieve the same risks, thus capital ratio, as in the case of no bailout being available, regulators need to impose a tighter (higher) capital requirement in the form of κ^*.

The bailout to banks in order to compensate depositors for any losses is costly to the general public, but imposes no costs on banks. On the contrary, banks benefit from the bailout through lower deposit rates. We will now see how internalising these costs will affect the capital ratio chosen by banks.

Internalized bailouts The losses from compensating depositors are given by $D - \lambda L = \left(1-\lambda\left(1+\kappa\right)\right)D$, representing the deposits they have provided, and on which no interest is payable due to them being risk-free, less the amount recovered from defaulting loans. If there are costs $c > 1$ arising from financing this compensation and any regulator would want to internalize the costs it would charge these costs to banks, reducing their profits accordingly. Thus, banks obtain their profits if loans are repaid, but if loans are defaulting, they have to cover the losses arising from the bailout. We thus have bank profits given by

$$\Pi_B = \pi\left(\left(1+r_L\right)\left(1+\kappa\right)-1\right)D - \left(1-\pi\right)c\left(1-\lambda\left(1+\kappa\right)\right)D, \tag{34.100}$$

yielding the first order condition for the optimal capital ratio as

$$\begin{aligned}\frac{\partial\Pi_R}{\kappa} = &\left(\frac{\partial\pi}{\partial\kappa}\left(1+r_L\right)\left(1+\kappa\right)+\pi\left(1+r_L\right)\right.\\ &\left.+\frac{\partial\pi}{\partial\kappa}c\left(1-\lambda\left(1+\kappa\right)\right)+\left(1-\pi\right)\lambda c\right)D = 0.\end{aligned} \tag{34.101}$$

This gives us the optimal capital ratio fulfilling the equation

$$\frac{\partial\pi}{\partial\kappa} = -\frac{\pi\left(1+r_L\right)+\left(1-\pi\right)\lambda c}{\left(1+r_L\right)\left(1+\kappa\right)+c\left(1-\lambda\left(1+\kappa\right)\right)}. \tag{34.102}$$

The regulator would now again seek to set incentives such that with the bailout the bank takes the optimal capital ratio from the case without a bailout, and hence takes the same risk when providing loans. This is achieved by equalling equations (34.97) and (34.102), which solves for

$$1 + \kappa^{**} = \frac{c}{c-1} \frac{\pi (1 + r_L) + \lambda (1 - \pi)}{\lambda (1 + r_L)} = \frac{c}{c-1} (1 + \kappa^*), \tag{34.103}$$

where the last equality emerges when inserting from equation (34.99). As we assumed that $c > 1$, we see that the capital requirements, κ^{**}, have increased. Thus if we wanted bank to take into account the costs of a bailout, capital requirements would need to be increased to reduce the risk-taking by banks further.

We can now see that this capital ratio is higher if the expression in (34.102) is smaller than that in equation (34.94). This in the case if

$$\lambda < \lambda^{**} = \frac{c}{c-1}\lambda^*, \tag{34.104}$$

where λ^* is defined in equation (34.98). With $c > 1$ this constraint is less binding than λ^* for bailed out banks that do not internalise the costs they impose on depositors. Hence if losses from defaulting loans are sufficiently large, the regulator would have to impose higher capital requirements than when not internalising the costs of bailouts. The same argument as before applies in that the bank would take higher risks than in cases where a bail-out is not anticipated.

Summary We have thus seen that banks who are bailed out have incentives to increase the risks they are taking by reducing the capital ratio, equivalent to increasing their leverage. Were regulators seeking to impose a capital ratio on the bank that induces them to take the same risks as when not bailed out, capital requirements would be stricter than is optimal for bank without the prospects of being bailed out; regulators seeking to internalise the costs of bailouts would have to impose even higher capital requirements. It is thus that the prospect of bailouts increases capital requirements if regulators wanted to avoid banks taking higher risks.

If depositors expect banks to be bailed out, then they will not be sensitive to the risks that banks are taking and this allows banks to increase risks through a higher leverage, which will increase their profitability. As this increases risks, a regulator would want to reduce these risks through the imposition of more strict capital requirements, which limits the amount of risks banks take to the optimal risk level that does not involve a bail-out. Hence bailing banks out requires more stringent capital requirements in order to limit risks.

Reading Ma & Nguyen (2021)

34.5.2 Bailouts and contagion

Bailouts are often conducted to avoid contagion, in addition to benefitting depositors. We will here investigate how capital requirements can be used to reduce the incentives of banks to increase their risks if they are faced with contagion, but also can benefit from being bailed out.

We assume that if bank j fails, bank i fails with probability p; this simplifying approach takes into account the possibility of contagion within the banking system.

Such contagion may arise from the exposure to the other bank arising from interbank loans or the lower valuation of assets due to the failing bank liquidating similar assets. Hence, for a bank to survive, the bank's own loans need to be repaid and either the other bank needs to have their loans being repaid as well or if they are failing, they have to be bailed out.

We will first establish the optimal risk taking by banks if bailouts are not occurring and can then compare this result with the case of failing banks being bailed out.

No bailouts Banks finance their loans L_i through deposits D, identical for all banks, and equity E_i such that $L_i = D + E_i$. We also introduce a capital ratio κ_i, which is defined through $E_i = \kappa_i D$ and hence $L_i = (1 + \kappa_i) D$. The loan L_i, is repaid including interest r_L with probability π_i, but the bank will not obtain this revenue if the other bank fails, $1 - \pi_j$ and contagion occurs, p. From this revenue the bank repays its depositors D, including interest r_D, and to focus on economic profits, we also deduct the equity of the bank, E_i, including its cost of equity r_E. In addition banks have to exert effort to increase the probability of the loan being repaid, which we assume to be increasing in the size of the loan and the repayment rate. Thus bank profits are given by

$$\begin{aligned} \Pi_B^i &= \pi_i \left(1 - p \left(1 - \pi_j\right)\right) \left(\left(1 + r_L\right) L_i - \left(1 + r_D\right) D\right) \\ &\quad - \left(1 + r_E\right) E_i - \frac{1}{2} c \pi_i^2 L \\ &= \left(\pi_i \left(1 - p \left(1 - \pi_j\right)\right) \left(\left(1 + r_L\right) \left(1 + \kappa_i\right) - \left(1 + r_D\right)\right)\right. \\ &\quad \left. - \left(1 + r_E\right) \kappa_i - \frac{1}{2} c \left(1 + \kappa_i\right) \pi_i^2 \right) D, \end{aligned} \tag{34.105}$$

where the second equality has been obtained by inserting for L_i and E_i. The bank will choose the repayment rate that maximizes their profits, giving us the first order condition as

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial \pi_i} &= \left(\left(1 - p \left(1 - \pi_j\right)\right) \left(\left(1 + r_L\right) \left(1 + \kappa_i\right)\right.\right. \\ &\quad \left.\left. - \left(1 + r_D\right)\right) - c \left(1 + \kappa_i\right) \pi_i \right) D = 0, \end{aligned} \tag{34.106}$$

which can be solved for

$$\pi_i = \frac{1 - p \left(1 - \pi_j\right)}{c \left(1 + \kappa_i\right)} \left(\left(1 + r_L\right) \left(1 + \kappa_i\right) - \left(1 + r_D\right)\right). \tag{34.107}$$

As all banks are otherwise identical having the same amount of deposits, we focus on symmetric equilibria that would require $\pi_i = \pi_j = \pi$ and $\kappa_i = \kappa_j = \kappa$. Thus equation (34.107) solves for

$$\pi^* = \frac{(1 - p) \left(\left(1 + r_L\right) (1 + \kappa) - \left(1 + r_D\right)\right)}{c (1 + \kappa) - p \left(\left(1 + r_L\right) (1 + \kappa) - \left(1 + r_D\right)\right)} \tag{34.108}$$

where we assume that $c(1+\kappa) > p((1+r_L)(1+\kappa)-(1+r_D))$ to obtain a feasible solution.

As expected, if the capital ratio is increased the risk the bank takes is reduced through banks choosing a higher repayment rate for loans. Formally we see this relationship through the equation

$$\frac{\partial \pi^*}{\partial \kappa} = \frac{(1-p)(1+r_D)c}{(c(1+\kappa)-p((1+r_L)(1+\kappa)-(1+r_D)))^2} > 0. \tag{34.109}$$

The risks of the loan, π^*, and the contagion risk, p, are substitutes as they both cause the bank to fail. We find that

$$\frac{\partial \pi^*}{\partial p} = -\frac{(1+rL)(1+\kappa)-(1+r_D)}{(1+\kappa)-p((1+r_L)(1+\kappa)-(1+r_D))} < 0. \tag{34.110}$$

Possible bailouts Let us now consider the probability that a bank is bailed out if it fails due loans not being repaid; we do not consider a bailout if the origin of the failure is contagion, but only if the bank itself fails due to its loans defaulting. The bank is bailed out with probability $\hat{p}$ and in this case retains a value of bD, the remainder is taken by the regulator who conducts the bailout. Accounting for this bailout and the fact that contagion can only happen if the other bank is not bailed out, the bank profits from equation (34.105) changes to

$$\begin{aligned} \hat{\Pi}_B^i &= \left(\pi_i\left(1-p\left(1-\pi_j\right)(1=\hat{p})\right)((1+r_L)(1+\kappa_i)-(1+r_D))\right. \\ &= \left.+(1-\pi_i)\hat{p}b-(1+r_E)\kappa_i-\frac{1}{2}c(1+\kappa_i)\pi_i^2\right)D. \end{aligned} \tag{34.111}$$

We have to add to the previous case that the bank does not obtain the revenue from the loan repayments if the other bank is not bailed out $\hat{p}$ and if the bank itself fails due to loans not being repaid$1-\pi_i$ obtains the value bD if it is bailed out, $\hat{p}$. In this case the optimal risk-taking by the bank is determined from the first order condition

$$\begin{aligned} \frac{\partial \hat{\Pi}_B^i}{\partial \pi_i} &= \left(\left(1-p\left(1-\pi_j\right)(1-\hat{p})\right)((1+r_L)(1+\kappa_i)-(1+r_D))\right. \\ &\quad \left.-\hat{p}b-c(1+\kappa_i)\pi_i\right)D=0 \end{aligned} \tag{34.112}$$

such that we obtain the optimal repayment rate as

$$\pi_i = \frac{\left\{\begin{array}{c}\left(1-p\left(1-\pi_j\right)(1-\hat{p})\right)((1+r_L)(1+\kappa_i)-(1+r_D)) \\ -\hat{p}b-c(1+\kappa_i)\pi_i\end{array}\right\}}{c(1+\kappa_i)}. \tag{34.113}$$

If we again only consider symmetric equilibria such that $\pi_i = \pi_j = \pi$ and $\kappa_i = \kappa_j = \kappa$, we obtain

$$\hat{\pi}^* = \frac{(1-p(1-\hat{p}))((1+r_L)(1+\kappa)-(1+r_R))-\hat{p}b}{c(1+\kappa)-p(1-\hat{p})((1+r_L)(1+\kappa)-(1+r_D))}. \tag{34.114}$$

In the presence of bailouts, the banks choose riskier loans than in the case without bailouts if $\hat{\pi}^* < \pi^*$, which when using equations (34.107) and (34.114) requires

$$b \geq b^* = \frac{\left\{ \begin{array}{c} p\left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right) \\ \times \left(c\left(1+\kappa\right)-\left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right)\right) \end{array} \right\}}{c\left(1+\kappa\right)-p\left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right)}. \tag{34.115}$$

Thus, if the size of the bailout in the sense of the value of the bank retained by the bank owners is sufficiently high, the bank will choose more risky loans. We see clearly that a sufficiently large bailout incentivises banks to increase their risks as even if the loans they provide default, they will retain some value, bD, while without bail-outs banks would lose their entire value.

If bailouts become more likely, the risks banks take increases if $\frac{\partial \hat{\pi}^*}{\partial \hat{p}} < 0$, which requires that

$$p \leq p^* = \frac{bc\left(1+\kappa\right)}{\left\{ \begin{array}{c} \left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right) \\ \times \left(c\left(1+\kappa\right)-\left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right)+b\right) \end{array} \right\}}. \tag{34.116}$$

Hence if contagion is sufficiently unlikely, then if bailouts become more likely, the risk-taking of banks increases. It is only in the case of failing due to contagion that banks would not be bailed out and therefore they know that if they fail, they may impose the failure on another bank, who in turn could impose a failure on them, leading a to a situation where banks are not bailed out. If this reciprocal contagion is sufficiently unlikely, banks will increase their risks as bailouts become more likely.

The size of the bailout, b, increases risks as we have shown that $\frac{\partial \hat{\pi}^*}{\partial b} < 0$ in equation (34.114), higher capital requirements reduces risks due to $\frac{\partial \hat{\pi}^*}{\partial \kappa} > 0$ as can easily be verified from equation (34.109); thus capital requirements have another effect as we can obtain that

$$\frac{\partial p^*}{\partial \kappa} = bc\frac{\zeta_0}{\zeta_1 \zeta_2}, \tag{34.117}$$

where

$$\begin{aligned} \zeta_0 &= b\left(1+r_D\right) \\ &\quad + \left(2\left(1+r_L\right)^2\left(1+\kappa\right)^2-\left(1+r_D\right)^2-c\left(1+\kappa\right)^2\left(1+r_L\right)\right), \\ \zeta_1 &= \left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right)^2, \\ \zeta_2 &= \left(c\left(1+\kappa\right)-\left(\left(1+r_L\right)\left(1+\kappa\right)-\left(1+r_D\right)\right)-b\right)^2. \end{aligned}$$

This expression is negative and an increase in capital requirements reduces the threshold at which the probability of bailouts increases risks, provided

$$b < \frac{c\,(1+\kappa)^2\,(1+r_L) + (1+r_D)^2 - 2\,(1+r_L)^2\,(1+\kappa)^2}{1+r_D}, \tag{34.118}$$

We thus see that if the size of the bailout, b is sufficiently small, increasing capital requirements κ^* will lead to a lower threshold for contagion to increase risk-taking as we make bailouts more likely; the banking system becomes more stable. Increasing the probability of a bailout, or the willingness of regulators for a bailout of failing banks, can thus be counteracted by an increase in capital requirements without increasing the risk-taking by banks.

Summary We have seen that bailouts induce banks to take higher risks if the size of the bailout is sufficiently high. However, regulators can counteract this develop if they limit the size of the bailout and increase capital requirements for banks. In this case, the increased capital requirements will induce banks to reduce risks for a wider range of contagion. It is therefore possible for capital requirements to have a positive effect on the risk taking of banks, but with contagion and bailouts, this is based on strict conditions regarding the size of the bailout and the likelihood of contagion; both cannot be too high. If these conditions are not fulfilled, increasing the expectations of a bailout will always increase risks, even if capital requirements are increased.

Reading Dell'Ariccia & Ratnovski (2019)

Résumé

We have seen that if depositors expect a failing bank to be bailed out, the bank has an incentive to increase the risk it is taking. This increased risk can be reduced by stricter capital requirements that a regulator imposes; such capital requirements directly limits the risks due to a lower leverage and will incentivise the bank to reduce the risks it is exposed to from lending. These capital requirements will be more strict then amount of capital the bank would hold if it was not being bailed out; this higher capital requirement compensates for the non-responsiveness of depositors to the risks banks are taking and align the risks in the two situations. Bailouts often do not happen in isolation at a single bank, but in particular at times of wider stress in the banking system. If contagion of failures between banks are a concern, then increasing capital requirements can reduce this risk, provided any bailouts are not too high, while higher bailouts would increase the likelihood of contagion.

We have therefore seen that capital requirements are also a useful tool to adjust for the incentive changes arising from bailouts, in addition to the its effect of being able to reduce contagion, provided bailouts are not too large. The impact of capital requirements are more than merely influencing the risks banks take, additional concerns such as the impact of bailouts and contagion between banks can also be addressed.

34.6 The use of bail-in bonds

Regulators do not only impose requirements on the minimum capital a bank must hold but there are also requirements for debt on which losses can be imposed without forcing the bank to be liquidated. Such bail-ins of bondholders, commonly through subordinated bonds, allows regulators to treat these as part of the capital requirements of a bank. We will here look at whether using such bonds is desired by banks and will then compare the choices of banks with that of a regulator that will take into account the social costs of a bank failing and depositors not being repaid.

A bank can obtain deposits D and subordinated debt S to finance the loans L they provide, in additional to holding equity E. Depositors are repaid with interest r_D as long as the funds available from repaid loans are sufficient. Loan repayments include interest r_L and these are repaid with probability π. If the funds from loan repayments are not sufficient to repay all depositors, the bank is liquidated and a fraction λ of the loan value realised, which is then distributed to depositors. Thus the profits of depositors are given by

$$\hat{\Pi}_D = \begin{cases} (1 + r_D) D - D & \text{if } (1 + r_D) D \leq \pi (1 + r_L) L \\ \lambda\pi (1 + r_L) L - D & \text{if } \pi (1 + r_L) L < (1 + r_D) D \end{cases} . \tag{34.119}$$

Subordinated debt is structured such that it can be bailed-in if the bank would otherwise fail as depositors and bondholders cannot be repaid both from the proceeds of the loans. If the funds available to the bank are not sufficient, depositors are repaid first and once they are fully repaid, the remainder is split between the bondholders, receiving a fraction γ, and the bank retaining a fraction $1 - \gamma$. If the proceeds the bank obtained is not sufficient to repay all depositors, bondholders will not receive any repayment of their bond. With the interest paid on the bond denoted by r_S, we obtain the profits of the bondholders as

$$\hat{\Pi}_S = \begin{cases} (1 + r_S) S - S \\ \quad \text{if } \; (1 + r_D) D + (1 + r_S) S \leq \pi (1 + r_L) L \\ \gamma (\pi (1 + r_L) L - (1 + r_D) D) - S \\ \quad \text{if } \; (1 + r_D) D \leq \pi (1 + r_L) L < (1 + r_D) D + (1 + r_S) S \\ -S \\ \quad \text{if } \; \pi (1 + r_L) L < (1 + r_D) D \end{cases} . \tag{34.120}$$

Banks generate profits from the repayment of the loans, less the repayments to depositors and bondholders, provided the bank does not fail. In addition we assume that the bank faces costs of providing loans that are increasing in the loan size. Such costs would cover the costs of assessing borrowers, monitoring loans, and other associated tasks. The profits of the bank are then given by

$$\hat{\Pi}_B = \begin{cases} \pi\left(1+r_L\right)L-\left(1+r_D\right)D-\left(1+r_S\right)S-E-\frac{1}{2}cL^2 \\ \quad \text{if} \quad \left(1+r_D\right)D+\left(1+r_S\right)S \leq \pi\left(1+r_L\right)L \\ \left(1-\gamma\right)\left(\pi\left(1+r_L\right)L-\left(1+r_D\right)D\right)-E-\frac{1}{2}cL^2 \\ \quad \text{if} \quad \left(1+r_D\right)D \leq \pi\left(1+r_L\right)L < \left(1+r_D\right)D+\left(1+r_S\right)S \\ -E-\frac{1}{2}cL^2 \\ \quad \text{if} \quad \pi\left(1+r_L\right)L < \left(1+r_D\right)D \end{cases}. \tag{34.121}$$

We now assume that the repayment rate of loans, π, is not known, but banks and depositors only know its distribution $F(\pi)$. Thus the expected profits of depositors, bondholders, and banks are given by

$$\begin{aligned} \Pi_D &= \int_0^{\frac{(1+r_D)D}{(1+r_L)L}} \left(\lambda\pi\left(1+r_L\right)L-D\right)dF\left(\pi\right) \\ &\quad + \int_{\frac{(1+r_D)D}{(1+r_L)L}}^{1} \left(\left(1+r_D\right)D-D\right)dF\left(\pi\right) \\ &= \lambda\left(1+r_L\right)L\int_0^{\frac{(1+r_D)D}{(1+r_L)L}} \pi dF\left(\pi\right) \\ &\quad + \left(1-F\left(\frac{\left(1+r_D\right)}{\left(1+r_L\right)L}\right)\right)\left(1+r_D\right)D-D, \end{aligned} \tag{34.122}$$

$$\begin{aligned} \Pi_S &= \int_0^{\frac{(1+r_D)D}{(1+r_L)L}} \left(-B\right)dF\left(\pi\right) \\ &\quad + \int_{\frac{(1+r_D)D}{(1+r_L)L}}^{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}} \left(\gamma\left(\pi\left(1+r_L\right)L\right.\right. \\ &\qquad \left.\left.-\left(1+r_D\right)D\right)-S\right)dF\left(\pi\right) \\ &\quad + \int_{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}}^{1} \left(\left(1+r_S\right)S-S\right)dF\left(\pi\right) \\ &= \gamma\left(1+r_L\right)L\int_{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}}^{1} \left(\left(1+r_S\right)S-S\right)\pi dF\left(\pi\right) \\ &\quad -\gamma\left(1+r_D\right)D\left(F\left(\frac{\left(1+r_D\right)D+\left(1+r_S\right)S}{\left(1+r_L\right)L}\right)\right. \\ &\qquad \left.-F\left(\frac{\left(1+r_D\right)D}{\left(1+r_L\right)L}\right)\right) \\ &\quad + \left(1+r_S\right)S\left(1-F\left(\frac{\left(1+r_D\right)D}{\left(1+r_L\right)L}\right)\right)-S, \end{aligned} \tag{34.123}$$

$$\begin{aligned}
\Pi_B = &\int_0^{\frac{(1+r_D)D}{(1+r_L)L}} \left(-E - \frac{1}{2}cL^2\right) F(\pi) \\
&+ \int_{\frac{(1+r_D)D}{(1+r_L)L}}^{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}} \left((1-\gamma)\left(\pi\left(1+r_L\right)L\right.\right. \\
&\left.\left. - \left(1+r_D\right)D\right) - E - \frac{1}{2}cL^2\right) dF(\pi) \\
&+ \int_{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}}^{1} \left(\pi\left(1+r_L\right)L - \left(1+r_D\right)D\right. \\
&\left. - \left(1+r_S\right)S - E - \frac{1}{2}cL^2\right) dF(\pi) \\
= &\left(1+r_L\right)L\left((1-\gamma)\int_{\frac{(1+r_D)D}{(1+r_L)L}}^{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}} \pi dF(\pi)\right. \\
&\left. + \int_{\frac{(1+r_D)D+(1+r_S)S}{(1+r_L)L}}^{1} \pi dF(\pi)\right) \\
&- (1-\gamma)\left(1+r_D\right)D\left(F\left(\frac{\left(1+r_D\right)D + \left(1+r_S\right)S}{\left(1+r_L\right)L}\right)\right. \\
&\left. - F\left(\frac{\left(1+r_D\right)D}{\left(1+r_L\right)L}\right)\right) \\
&- \left(1+r_D\right)D\left(1 - F\left(\frac{\left(1+r_D\right)D}{\left(1+r_L\right)L}\right)\right) - E - \frac{1}{2}cL^2.
\end{aligned} \tag{34.124}$$

If deposit and bond markets are competitive such that depositors and bondholders make no profits, $\Pi_D = \Pi_S = 0$, banks can select the combination of deposits and bonds, D and S, that maximizes their profits. The deposit and bond rates are then determined from our assumption that $\Pi_D = \Pi_S = 0$. We can then determine the Lagrangian to maximize their profits as

$$\mathcal{L}_B = \Pi_B + \xi_1 \Pi_D + \xi_2 \Pi_S, \tag{34.125}$$

with ξ_i denoting the Lagrange multipliers; we also use that $L = D+S+E$. Conducting the maximization, we will obtain the optimal amount of deposits and subordinated bonds.

A regulator would want to take into account the losses to depositors; the losses to the holders of subordinated bonds would not be taken into account as these have been contracted for and bondholders expected to be bailed in. The losses to depositors are given by

$$\Pi_D^* = \int_0^{\frac{(1+r_D)D}{(1+r_L)L}} (1-\lambda)\,\pi\left(1+r_L\right)L dF(\pi) \tag{34.126}$$

and we can easily get that

$$\frac{\partial \Pi_D^*}{\partial D} = (1-\lambda)(1+r_L)\left(\int_0^{\frac{(1+r_D)D}{(1+r_L)L}} \pi dF(\pi) + \left(\frac{1+r_D}{1+r_L}\right)^2 \frac{D(L-D)}{L^2} f\left(\frac{(1+r_D)D}{(1+r_L)L}\right)\right) > 0, \tag{34.127}$$

$$\frac{\partial \Pi_D^*}{\partial S} = (1-\lambda)(1+r_L)\left(\int_0^{\frac{(1+r_D)D}{(1+r_L)L}} \pi dF(\pi) - \left(\frac{1+r_D}{1+r_L}\right)^2 \frac{D^2}{L^2} f\left(\frac{(1+r_D)D}{(1+r_L)L}\right)\right) > 0,$$

with $f(\cdot)$ denoting the density function.

We can now easily see that $\frac{\partial \Pi_D^*}{\partial D} > 0$ and while the sign of $\frac{\partial \Pi_D^*}{\partial B}$ can take positive or negative values, however it is most likely to be positive. This will be the case, for example, if the distribution of the repayment rate is uniform; thus we would find $\frac{\partial \Pi_D^*}{\partial S} > 0$.

The Lagrangian for a regulator will then become $\mathcal{L}_R = \Pi_B - \Pi_D^* + \xi_1 \Pi_D + \xi_2 \Pi_S = \mathcal{L}_B - \Pi_D^*$. We can now compare the optimal solution of the bank with that of the regulator. Assume the bank has chosen the optimal amount of deposits and subordinated bonds such that $\frac{\partial \mathcal{L}_B}{\partial D} = \frac{\partial \mathcal{L}_B}{\partial S} = 0$, in which case we obtain for the optimal solution of the regulator that

$$\begin{aligned} \frac{\partial \mathcal{L}_R}{\partial D} &= \frac{\partial \mathcal{L}_B}{\partial D} - \frac{\partial \Pi_D^*}{\partial D} = -\frac{\partial \Pi_D^*}{\partial D} < 0, \\ \frac{\partial \mathcal{L}_R}{\partial S} &= \frac{\partial \mathcal{L}_B}{\partial S} - \frac{\partial \Pi_D^*}{\partial S} = -\frac{\partial \Pi_D^*}{\partial S} < 0. \end{aligned} \tag{34.128}$$

With both derivatives negative, we see that the regulator would find it optimal for banks to hold less deposits and less subordinated bonds, thus have a lower overall leverage than is optimal for the bank. Thus regulator might want to impose capital requirements on banks to reduce their leverage.

Furthermore, we can determine the optimal combination of deposits and subordinated bonds. Define the ratio of deposits and subordinated bonds by η such that $\eta = \frac{D}{S}$ and using that $L = D + S + E$ we get $D = \frac{\eta}{1+\eta}(L-E)$. As for the optimal choice of deposits and subordinated bonds by the bank, their ratio η would also be optimal such that $\frac{\partial \mathcal{L}_B}{\partial \eta} = 0$ and with

$$\frac{\partial \Pi_D^*}{\partial \eta} = (1-\lambda)\frac{(1+r_D)^2}{1+r_L}\frac{(L-E)^2}{L}\frac{\eta}{(1+\eta)^2} f\left(\frac{(1+r_D)D}{(1+r_L)L}\right) > 0 \tag{34.129}$$

we easily see that

$$\frac{\partial \mathcal{L}_R}{\partial \eta} = \frac{\partial \mathcal{L}_B}{\partial \eta} - \frac{\partial \Pi_D^*}{\partial \eta} = -\frac{\partial \Pi_D^*}{\partial \eta} < 0. \tag{34.130}$$

It is therefore that the regulator would prefer a lower ratio of deposits to subordinated bonds than is optimal for the bank. It is therefore that the regulator could supplement their capital requirements with requirements for subordinated bonds that can be bailed in. Combining these two regulatory requirements would then give banks a financing structure that takes into account the social costs of imposing losses on depositors and is socially preferable.

We have seen that it is optimal for banks to finance their loan with a combination of deposits and subordinated bonds that can be bailed-in to prevent a bank to fail. It was found, however, that banks choose too many deposits and bonds alike, giving it a too high leverage compared to the social optimum and the allocation was biased too much in favour of deposits rather than subordinated bonds. A regulator might therefore want to impose not only capital requirements to limit the leverage of banks, but complement this by requirements for bonds that can be bailed-in.

Reading Clayton & Schaab (2025)

34.7 Political influences

Banks are affected by regulations and will seek to influence them in their favour. At the same time, politicians may want to influence banks into providing loans for companies or projects that are advantageous to them, but that normally would not be financed by banks. We will look at these interactions between regulators and banks as well as the way regulations can be used to make it viable for banks to provide loans that benefit politicians. We commence in chapter 34.7.1 with exploring how regulations can be adjusted to enable banks to provide loans that politicians would like banks to provide, before in chapter 34.7.2 we will look at how banks can affect regulations, specifically, capital requirements, through lobbying politicians.

34.7.1 Preferred loans

It is common that governments seek to entice banks to provide loans to borrowers that support their policies. banks might be reluctant to provide such loans as the risks involved might be too high or they cannot obtain a loan rate that covers their costs. We will briefly discuss here how banks can be enticed to provide loans in such circumstances.

As in most instances regulators seeking to support government policies cannot direct banks to provide loans to specific borrowers at specific terms, they have to rely on incentives. Banks are subject to capital requirements and thus have a limited amount of loans they can provide. Assume that banks could either lend to a borrower the government, and hence the regulator, prefers or provide a loan to an alternative borrower. This alternative borrower will repay the L loan with probability $\hat{\pi}$ and pay a loan rate of $\hat{r}_L$ such that the expected repayment is that the regulator can direct banks to route loans away from their usual borrowers towards such preferred borrowers. They cannot, however, dictate that banks have to provide loans at all. Therefore assume that banks have an alternative safe investment that yields them a return of θ on the loan amount of L. Such an alternative might be the provision of interbank loans or investment into fixed assets such as property. We assume that regulators can prevent banks from giving other loans unless the preferred borrower has obtained a loan, making the alterantive investment the only other option for the bank.

Loans provided to the preferred borrower are repaid with probability π, including interest r_L, and financed by deposits D, requiring interest r_D and equity E such that $L = D + E$. Banks are subjected to a capital requirement in the form of a maximum capital ratio κ such that $E = \kappa D$. For banks to give a loan this needs to be more profitable than the alternative investment, hence we require that

$$\pi \left((1 + r_L) L - (1 + r_D) D \right) \geq \theta L, \tag{34.131}$$

which solves for

$$1 + \kappa \geq 1 + \kappa^* = \frac{\pi (1 + r_D)}{\pi (1 + r_L) - \theta}. \tag{34.132}$$

after inserting for $L = (1 + \kappa) D$ and assuming that the alternative investment is less profitable than the loan to the preferred borrower, $\theta < \pi (1 + r_L)$. We thus see that banks would provide the loan if the capital ratio is sufficiently high.

If we now increase the risk of a loan, that is reduce the repayment rate π, we have due to

$$\frac{\partial (1 + \kappa^*)}{\partial \pi} = -\frac{\theta (1 + r_D)}{(\pi (1 + r_L) - \theta)^2} < 0 \tag{34.133}$$

that the capital requirements are increasing even more to ensure that banks are choosing the preferred borrower over alternative investments. Hence if we interpret (34.132) as the minimal capital requirements that the regulator needs to apply such that banks choose the borrower preferred by the government.

Similarly, a less profitable loan with a lower interest rate r_L requires higher capital requirements as we easily see that

$$\frac{\partial (1 + \kappa^*)}{\partial (1 + r_L)} = -\frac{\theta (1 + r_D)(1 + r_L)}{(\pi (1 + r_L) - \theta)^2} < 0. \tag{34.134}$$

Hence we see that in order to entice banks to provide loans to borrowers the government sees as favourable, but who are seen as either too risky or not sufficient profitable due to a low loan rate, banks need to be subjected to higher capital requirements. The resulting lower leverage will provide incentives to banks to forego the alternative investment that yield a low return and despite the reservations they have about the preferred borrower provide the loan as requested by the government.

Reading Thakor (2021)

34.7.2 Lobbying

Banks are not only subjected to capital requirements by a regulator setting minimum capital ratios or maximum leverage, they will seek to influence such decisions to their own benefit. Thus banks may want to lobby regulators to impose less stringent capital requirements. We will now analyse how banks can conduct such lobbying and what the impact on the capital requirements set by the regulator will be.

Let us assume that banks use a fraction γ of their revenue to lobby the regulator; these funds may be used to provide benefits to regulators, such as inviting to industry events at no costs to them, often in attractive locations. With loans L being repaid with probability π_i, including interest r_L, the revenue banks obtain is given by $\pi_i (1 + r_L) L$ such that the amount used for lobbying activity is given by $G_i = \gamma \pi_i (1 + r_L) L$.

We can now analyse how the bank optimally select their capital ratio and the amount of revenue spent on lobbying.

Bank decision We assume that the repayment rate of loans, π_i, is either high, π_H with probability p or low with $\pi_L < \pi_H$ with probability $1 - p$. These possible repayment rates may reflect economic conditions with a high repayment rate during times of economic expansion and low repayment rates during recessions. Loans are financed using deposits D, yielding interest r_D. and equity E, with a return of r_E^i. We define a capital ratio κ such that $E = \kappa D$ and we have $L = E + D = (1 + \kappa) D$. The return of equity can be determined as the as the profits the bank generates after having deducted the lobbying costs, divided by the amount of equity the bank holds, thus

$$
\begin{aligned}
1 + r_E^i &= \frac{\pi_i (1 + r_L) L - (1 + r_D) D - G_i}{E} \\
&= \frac{(1 - \gamma) \pi_i (1 + r_L) (1 + \kappa) - (1 + r_D)}{\kappa}.
\end{aligned}
\tag{34.135}
$$

If we assume In order to have deposits and equity co-existing, they need to provide the same return; we further assume that deposit insurance guarantees depositors the repayment of their deposits including interest. The return on equity will depend on the repayment rates of the loans, whether the economy is in a state to provide high or low repayment rates. We thus require

$$1 + r_D = p \left(1 + r_E^H\right) + (1 - p) \left(1 + r_E^L\right). \tag{34.136}$$

Inserting from equation (34.135), we obtain the deposit rate as

$$1 + r_D = (1 - \gamma) (p \pi_H + (1 - p) \pi_L) (1 + r_L), \tag{34.137}$$

provided the bank does not default. This additional assumption is made as in the case of a default the bank owners will receive no payments. In order to banks to avoid default, they need to generate profits; thus we require that

$$\begin{aligned}\Pi_B^i &= \pi_i\,(1+r_L)\,L-(1+r_D)\,D-G_i \\ &= ((1-\gamma)\,\pi_i\,(1+r_L)\,(1+\kappa)-(1+r_D))\,D\geq 0,\end{aligned} \tag{34.138}$$

where we inserted for the lobbying amount. If we want to assume that the bank cannot default, this condition needs to be fulfilled for the low repayment rate π_L. Inserting for the deposit rate from equation (34.137), the condition becomes

$$\kappa \geq \kappa^* = (1-\gamma)\,p\frac{\pi_H-\pi_L}{\pi_L}. \tag{34.139}$$

Thus to avoid a bank failing, it requires a sufficiently high capital ratio. This will limit its leverage and the losses from the high default rate of loans do not cause losses overall, only once the capital ratio is smaller is the amount of loans and hence the accumulated losses, sufficient to cause the bank to fail.

If the capital ratio is too low to prevent the bank from failing if the low repayment rate is realised, $\kappa<\kappa^*$, the returns on equity from equation (34.136) changes to

$$1+r_D = p\left(1+r_E^H\right) \tag{34.140}$$

due to limited liability as then $1+r_E^L=0$ due to the bank failing. Again, using equation (34.135) we obtain the deposit rate as

$$1+r_D=\frac{1-\gamma}{p+\kappa}p\pi_H\,(1+r_L)\,(1+\kappa)\,. \tag{34.141}$$

Similarly to the returns on deposit and equity having to be equal, the expected return on loans has to equal the returns of deposits to avoid a dominance of direct lending, hence we require that

$$(1+r_D)\,D = p\pi_H\,(1+r_L)+(1-p)\,\pi_L\,(1+r_L)\,L, \tag{34.142}$$

where the right-hand side gives the return on the loans in the two possible states of high and low repayment rates, respectively. Inserting for the deposit rate from equation (34.141) for potentially failing banks, gives us

$$\gamma=\frac{p\pi_H-(p\pi_H+(1-p)\,\pi_L)\,(p+\kappa)}{p\pi_H}. \tag{34.143}$$

We have thus obtained the optimal amount of lobbying a bank should conduct. We see that $\frac{\partial\gamma}{\partial\kappa}=-\frac{p\pi_H+(1-p)\pi_L}{p\pi_L}<0$ and hence if the lobbying effort increases, the capital requirements reduces. This is because a higher capital requirement reduces the profits of the bank as the bank can give less loans; as the amount of lobbying is a fraction of the revenue generated from loans, the amount generated for lobbying will reduce.

Having determined the optimal decision by banks, we will now continue with the decision by regulators to set capital requirements.

Regulator decision We now assume that regulators or share a fraction λ of the total wealth available, in addition to the lobbying revenue. With W_i denoting the wealth available for consumption, we have

$$W_i + G_i = \pi_i (1 + r_L) L, \tag{34.144}$$

such that the total value produced, $\pi_i (1 + r_l) L$, is split between lobbying, G_i, and the wealth remaining for consumption, W_i . We then have for regulators and non-regulators, respectively

$$\begin{aligned} \hat{W}_i &= \lambda W_i + G_i, \\ \hat{\hat{W}}_i &= (1 - \lambda) W_i. \end{aligned} \tag{34.145}$$

Regulators obtain a fraction λ of the wealth generated, in addition to the lobbying expenses they benefit from, while all other market participants share the remaining wealth. Using equation (34.144) and the definition of G_i, we easily obtain $W_i = (1 - \gamma) \pi_I (1 + r_L) (1 + \kappa) D$, and hence

$$\hat{W}_i = (\gamma + \lambda (1 - \gamma)) \pi_i (1 + r_L) (1 + \kappa) D. \tag{34.146}$$

Regulators will seek to maximize their benefits over the two possible states of the economy. We thus have

$$\begin{aligned} \Pi_R &= p\hat{W}_H + (1 - p) \hat{W}_L \\ &= (\lambda + \gamma (1 - \lambda)) (p\pi_H + (1 - p) \pi_L) (1 + r_L) (1 + \kappa) D. \end{aligned} \tag{34.147}$$

The capital ratio optimal for regulators can be obtained from solving the first order condition

$$\begin{aligned} \frac{\partial \Pi_R}{\partial \kappa} &= ((\lambda + \gamma (1 - \lambda)) (p\pi_H + (1 - p) \pi_L) (1 + r_L) \\ &\quad + (1 - \lambda) (p\pi_H + (1 - p) \pi_L) () (1 + \kappa) \frac{\partial \gamma}{\partial \kappa} \Big) D \\ &= 0, \end{aligned} \tag{34.148}$$

which using equation (34.143) solves for

$$\kappa^* = \frac{p (\lambda - p (1 - \lambda)) \pi_H - (1 - \lambda) \left(1 - p^2\right) \Pi_L}{2 (1 - \lambda) (p\pi_H + (1 - p) \pi_L)}. \tag{34.149}$$

Inserting this result back into equation (34.143) would give the optimal lobbying effort by banks if we assume that regulators will impose such a requirement.

We can now investigate the impact the importance of the regulator has on the capital requirements. Taking λ as a measure of the importance of the regulator as this measures the share of the non-lobbying revenue it receives, we have

$$\frac{\partial \kappa}{\partial \lambda} = \frac{p\pi_H}{2(1-\lambda)^2 (p\pi_H + (1-p)\pi_L)} > 0. \tag{34.150}$$

Thus the capital ratio would increase if the revenue banks generate have to be shared more widely; with given deposits, the higher equity required would increase the total amount of lending and hence the revenue that can be shared with the regulator. However, from equation (34.143) we see that

$$\frac{\partial \gamma}{\partial \lambda} = -\frac{p\pi_H + (1-p)\pi_L}{p\pi_H} \frac{\partial \kappa}{\partial \lambda} = -\frac{1}{2(1-\lambda)^2} < 0 \tag{34.151}$$

and the an increase in the importance of the regulator reduces the lobbying effort. This will counteract the increase in the revenue from higher capital ratios by reducing the benefits of lobbying.

Combining these two results we obtain that $\frac{\partial \kappa}{\partial \gamma} = \frac{\partial \kappa}{\partial \lambda}\frac{\partial \lambda}{\partial \gamma} < 0$ and in increase in lobbying will reduce the capital requirements.

Summary We have seen that banks can use lobbying to reduce capital requirements. They do so by ensuring that regulators benefit from the increased revenue a higher leverage generates to banks by providing them with a fraction of this revenue through lobbying expenses that benefits the regulator. Their concern for the overall welfare will limit how low capital requirements are set as the costs of bank failures, measured by less revenue available from consumption; setting capital requirements too low will generate too little revenue affecting the regulator as well as the wider economy and hence a minimum level of capital requirements are retained.

Reading Gersbach & Papageorgiou (2019)

Résumé

We have seen that capital requirements can be used by regulators to entice banks to provide loans they believe should be given for the benefits of society. By negating banks the use of funds for other loans than the preferred loans, sufficiently low capital requirements will enable banks to provide loans they would otherwise not consider. This might be achieved in practice by allocating a lower risk weight to such loans than to other loans of comparable risks. We might thus see capital weights and hence capital requirements distorted by political interference.

Banks will also seek to reduce any capital requirements that will be imposed on them through lobbying. By letting the regulatory authority participate in the benefits of lower capital requirements, for example if they provide more benefits to regulators as they become subject to less tight capital requirements, the regulator will take these effects of their capital requirements into account, resulting in lower capital requirements. It is thus that the cost of lobbying will be offset by the benefits of lower capital requirements, benefitting the banks, too.

Conclusions

The setting of capital requirements is one of the most important regulatory constraints put on banks. We have seen here that imposing such restrictions can have many different effects on the behaviour of banks, but also that capital requirements can be used to not only address their risk-taking, but also affect alleviating the effect of possible bailouts of banks. Increased capital requirements can give incentives to banks to increase the risks of those loans they take, such that instead of many relatively low-risk loans, with higher capital requirements banks will provide fewer loans, but at a higher risk. This compensates them for the loss of profits if the capital ratio is to be increased.

Even though capital requirements may in some instances be not very effective in reducing the risks banks take, or may even increase the risks banks take, in most cases they will achieve their aims if properly designed. The effect of capital requirements goes beyond merely reducing the risks that banks are taking by also addressing a moral hazard issue if banks are expected to be bailed out by other banks, the central bank, or governments. Imposing higher capital requirements can reduce this moral hazard and the risk-taking of banks will reduce.

Capital requirements are not always able to reduce the risks banks take as intended and may need to be complemented by other measures, such as a leverage ratio which is not based on the risks of the banks. In other cases capital requirements can be substituted for by other measures, for example a limit on deposit rates. It is, however, generally preferred that capital requirements are based in the risks of banks rather than a fixed leverage ratio unrelated to their risks.

Given the importance of capital requirements for banks and the restrictions it imposes on them, it is not surprising to see that banks will seek to obtain concessions from regulators for less stringent capital requirements. Lobbying by banks can be expected and will in general be successful in reducing the restrictions on banks if they ensure that the regulators benefit from the freedoms they provide banks with.

Overall we have seen that the use of capital requirements has multiple effects, some positive in the sense that they reduce the risks banks are taking to the social optimum, others negative in that they might distort the provision of loans towards more risky loans and the use of resources for lobbying regulators. It has to be noted that not in all cases will capital requirements be effective in achieving their goal of reducing the risks banks are taking, but they might have to be supplemented by other measures, or other measures alone would be as effective. It is therefore that changing capital requirements will have many effects on bank behaviour that will have to be weighed against each other to ensure they have the desired effect and are cost-effective.

Chapter 35

Liquidity regulation

The second key element of baking regulation, aside from capital requirements, is the imposition of liquidity requirements on banks where banks are typically required to hold a certain amount of cash reserves. The aim of this requirement is to ensure that banks can meet the demand for liquidity from either deposits being withdrawn or interbank loans not being extended. With banks being able to withstand the withdrawal of funds, the banking system as a whole is deemed to be more stable as a bank run is less likely to occur and if it occurs, less likely to spread. Similarly, by allowing interbank loans to be withdrawn, a liquidity shortage in the bank or banks withdrawing these, will not easily spread to other banks and hence the possibility of a systemic banking crisis is reduced.

In this chapter we will look at the impact of liquidity requirements on the incentives of banks to increase risks as in chapter 35.1 and chapter 35.2 will then look at the impact of diversification on the liquidity requirements of banks.

35.1 Risk-taking incentives

Liquidity requirements are imposed on banks with the aim to reduce the risks of banks not being able to meet depositor withdrawals and thus the risk of a bank run. However, liquidity requirements also bind funds of the bank that have to invested at no or very low interest rates rather than being lent out at much more attractive loan rates. We will here explore how banks will react to such liquidity requirements and explore the effect on their risk-taking in lending.

Let us assume that banks provide loans that are repaid with probability π and the loan rate r_L is reflecting the default rate such that $\frac{\partial(1+r_L)}{\partial\pi} < 0$ and $\frac{\partial^2(1+r_L)}{\partial\rho^2} < 0$. Cash holdings are a fraction ρ of deposits D such that $C = \rho D$ and all remaining funds are lent out, implying that $L = D - C = (1 - \rho) D$. With a deposit rate of r_D, the bank profits are then given by

$$\begin{aligned}\Pi_B &= \pi\left(\left(1+r_L\right)L-\left(1+r_D\right)D+C\right) \\ &= \pi\left(\left(1+r_L\right)\left(1-\rho\right)+\rho-\left(1+r_D\right)\right)D.\end{aligned} \tag{35.1}$$

The bank obtains the repayment of the loans and in turn repays depositors; in addition it retains the cash it had held. As we find that

$$\frac{\partial \Pi_B}{\partial \rho} = -\pi r_L D < 0, \tag{35.2}$$

we see that banks seek to minimize their cash holdings. Therefore any liquidity requirements would be binding.

The banks would choose a level of risk that maximizes their profits, giving us the first order condition

$$\begin{aligned}\frac{\partial \Pi_B}{\partial \pi} &= \left(\left(1+r_L\right)\left(1-\rho\right)+\rho-\left(1+r_D\right)\right. \\ &\quad \left.+\pi\left(1-\rho\right)\frac{\partial\left(1+r_L\right)}{\partial \rho}\right)D \\ &= \left(\left(1+r_L\right)+\pi\frac{\partial\left(1+r_L\right)}{\partial \pi}+\frac{\rho-\left(1+r_D\right)}{1-\rho}\right) \\ &= 0.\end{aligned} \tag{35.3}$$

Taking the total differential of this equilibrium condition, we obtain

$$\left(2\frac{\partial\left(1+r_L\right)}{\partial \pi}+\pi\frac{\partial^2\left(1+r+_L\right)}{\partial \pi^2}\right)d\pi-\frac{r_D}{\left(1-\rho\right)^2}d\rho = 0, \tag{35.4}$$

which transforms into

$$\frac{\partial \pi}{\partial \rho} = \frac{r_D}{\left(1-\rho\right)^2}\frac{1}{2\frac{\partial(1+r_L)}{\partial \pi}+\pi\frac{\partial^2(1+r_L)}{\partial \pi^2}} < 0, \tag{35.5}$$

with the sign emerging from our assumption that $\frac{\partial(1+r_D)}{\partial \pi} < 0$ and $\frac{\partial^2(1+rL)}{\partial \pi^2} < 0$. Therefore imposing tighter liquidity restrictions would induce the bank to take more risk. If liquidity requirements are increased, banks will be able to provide less loans as cash reserves need to be held back, which reduces the profits of the bank. Banks now seek to recover some of these lost profits by providing more risky loans at a higher loan rate.

We have thus seen that if banks are subjected to liquidity requirements, they will take higher risks. Thus the risks from a liquidity shortage of a bank might reduce, thus making a bank run less likely, but this increases the risks of the bank failing due to defaults on the loans they have provided.

Reading Glocker (2021)

35.2 Benefits of diversification

Diversification of loan portfolios should reduce the idiosyncratic risks from defaults that are particularly difficult to assess by depositors and regulators. We will evaluate how risks becoming more predictable in the sense of banks being only exposed to systematic risk, will affect the liquidity requirements of banks.

Let us assume that the repayment rate of loans consists of two elements, the first element $\pi \in \left[\underline{\pi}; \overline{\pi}\right]$ is common to all banks and represents the general economic conditions affecting the ability of borrowers to repay loans. This common repayment rate has a distribution $F(\cdot)$ and density $f(\cdot)$. Loans are given for two time periods and the common element π of the repayment rate will become public knowledge at after only one time period, before loans are due to be repaid. The second element, $p_i \in [0; 1 - \overline{\pi}]$ is different for each bank and not known to depositors, but banks will receive this information. The total repayment rate of loans is then given by $\pi_i = \pi + p_i$.

Banks are holding a fraction ρ of their deposits D as cash reserves, $C = \rho D$, and the loans given are $L = D - C = (1 - \rho) D$. When, after one time period, banks learn the repayment rate of their loans, they are assumed to be able to sell their loans and make an investment that yields them a certain return of $\lambda \pi_I (1 + r_L) L$, but leaves the depositors with no repayments. Such an investment will happen if

$$\pi (1 + r_L) L + C - (1 + r_D) D \leq \lambda \pi_I (1 + r_L) L,$$

where r_L and r_D are the loan and deposit rates, respectively. The left-hand side represents the profits the bank would make from retaining the loans they have provided, including the cash reserves they hold, while the right-hand side represents the alternative investment banks can make. Inserting for the loan amount L and cash reserves C we obtain that this condition is fulfilled if

$$\pi_i \leq \pi^* = \frac{1 + r_D - \rho}{(1 - \lambda)(1 - \rho)(1 + r_L)} \tag{35.6}$$

Assume that depositors are not willing to tolerate any losses; thus banks could not take the alternative investment and the mere possibility of banks choosing it, given that repayment rates are not perfectly known to depositors and therefore the validity of the above condition cannot be verified in advance, would leave banks exposed to a bank run. As depositors only know the common element of the repayment rate, π, but not the true repayment rate π_i they would consider the worst case scenario of $p_i = 0$ and hence a bank run would occur if $\pi < \pi^*$. In this case the bank would liquidate all loans and we assume that the revenue generated is not sufficient to cover the deposits fully, leaving the bank with no funds. Hence ex-ante, the profits of the bank are given by

$$\Pi_B = \int_{\pi^*}^{\overline{\pi}} \pi_i \left((1 + r_L) L + C - (1 + r_D) D\right) dF(\pi_i), \tag{35.7}$$

where we only consider the uncertainty on π, rather than π_i, as this is the relevant information depositor will have. Noting that from equation (35.6) that we have

$$\frac{\partial \pi^*}{\partial \rho} = -\frac{r_D}{(1-\gamma)(1-\rho)^2(1+r_L)} < 0, \tag{35.8}$$

we can define further the net assets of the bank, the value of its equity, by $E = (1+r_L)L + C - (1+r_D)D = ((1-\rho)(1+r_L)+\rho-(1+r_D))D$ and this gives us

$$\frac{\partial E}{\partial \rho} = -r_L < 0. \tag{35.9}$$

Banks will choose cash reserves that maximize their profits, leading to the first order condition

$$\frac{\partial \Pi_B}{\partial \rho} = \frac{\partial E}{\partial \rho}\int_{\pi^*}^{\overline{\pi}} \pi_i dF(\pi_i) - E\frac{\partial \pi^*}{\partial \rho}(E[p_i]+\pi^*)f(\pi^*) = 0, \tag{35.10}$$

which can be solved for the optimal cash ratio ρ.

We now assume that a bank is well diversified and the idiosyncratic risk is eliminated. In this case π_i is known to depositors as $p_i = E[p_i]$ and we do not need to use the worst-case scenario with $p_i = 0$, but use the certain value expected value instead. Hence equation (35.6) changes to become

$$\pi_i = \pi + E[p_i] \leq \pi^*, \tag{35.11}$$

and thus $\pi \leq \hat{\pi}^* = \pi^* - E[p_i]$. All other considerations remain unchanged and such that the first order condition for the optimal cash reserves in equation (35.10) becomes

$$\frac{\partial \hat{\Pi}_B}{\partial \rho} = \frac{\partial E}{\partial \rho}\int_{\pi^*}^{\overline{\pi}} \pi_i dF(\pi_i) - E\frac{\partial \Pi^*}{\partial \rho}(E[p_i]-\hat{\pi}^*)f(\hat{\pi}^*) = 0, \tag{35.12}$$

noting that E does not depend on $\hat{\pi}^*$ and $\frac{\partial \hat{pi}^*}{\partial \rho} = \frac{\partial^*}{\partial \rho}$. With the assumption that $f(\pi^*) = f(\hat{\pi}^*)$, which is the case in a uniform distribution, for example, we can now get the difference in the two first order conditions as

$$\frac{\partial \hat{\Pi}_B}{\partial \rho} - \frac{\partial \Pi_B}{\partial \rho} = \frac{\partial E}{\partial \rho}\int_{\hat{\pi}^*}^{\pi^*} \pi_i dF(\pi_i) + E\frac{\partial \pi^*}{\partial \rho}E[p_i]f(\pi^*) < 0. \tag{35.13}$$

The negative sign of this difference arises as $\frac{\partial E}{\partial \rho} < 0$ from equation (35.9) and the fact that $\hat{\pi}^* \leq \pi^*$ as long as $E[p_i] > 0$, making the first term negative. With $\frac{\partial \pi^*}{\partial \rho} < 0$ from equation (35.8), the second term is also negative. Hence if for the undiversified bank the cash ratio is optimal and $\frac{\partial \Pi_B}{\partial \rho} = 0$, we get $\frac{\partial \hat{\pi}_B}{\partial \rho} < 0$ for the diversified bank and the optimal cash holding of the diversified bank is lower.

It is therefore that banks holding a diversified portfolio of loans whose risk is predictable to depositors will have to hold less cash reserves to avoid a bank run.

This is the case because the risk banks face will be lower in a diversified portfolio and this lower risk allows depositors to avoid a loss to their funds for sure, even if cash reserves are lower.

Reading Gorton & Huang (2006)

Conclusions

We have seen that banks hold cash reserves to avoid failing in the face of a liquidity shock, provided this liquidity shock was not too large. While we found that a bailout reduces the incentives to hold such cash reserves, any regulator could impose liquidity requirements that are higher than what banks would hold and thus align the liquidity requirements with the cash reserves that banks would hold in the absence of bailouts, making the reliance on bailouts less common. However, by requiring banks to hold cash reserves beyond what is optimal for them, we have also seen that banks will increase the risks they are taking. Therefore, by increasing liquidity requirements bailouts may become less likely, but bank failures due to increased risk may also rise, counteracting at least some of the positive effects liquidity requirements have.

Chapter 36

Bank resolution

Bank regulation mainly focusses on capital requirements and to a lesser extent on liquidity requirements that banks need to meet. While the aim of such requirements is to reduce or even prevent the failure of banks, it is important to consider the way banks are liquidated if they are failing. We will here look at the incentives any regulator has to close a bank, even before it officially fails. Such a closure of a bank does not necessarily mean that the bank will be liquidated, it often involves the bank being taken over by another bank without any direct losses to depositors or the general public.

Before investigating the conditions under which banks should be closed in chapter 36.2, we will in chapter 36.1 look at conditions in which banks actually disclose that they are in distress; such an identification is essential to even consider the closure of a bank. An alternative to closing down a bank might be to introduce a so-called bad bank that purchases non-performing loans and allows banks to continue lending; this option will be discussed in chapter 36.3. We will then finally in chapter 36.4 not only investigate under which conditions a bank should be closed by a regulator, but also which type of regulator should make the closure decision.

36.1 Identification of distressed banks

If we wanted to consider the closure of a bank, it is essential to identify whether a bank is in distress or not. Banks will have more accurate information about their situation than a regulator or the general public. It is therefore possible that banks may hide the fact they are in distress in the hope of a subsequent recovery and thereby avoiding being closed. We will here look at incentives for banks to disclose the true state of their situation, allowing regulators to identify those banks that are in distress.

Let us assume that banks can hide information on their distress by not disclosing how many loans have defaulted. A bank having a high success rate π_H cannot pretend to have a low success rate as loans cannot be declared as default if they have not defaulted. A low success rate $\pi_L < \pi_H$ can, though, incentive banks not to disclose this fact; they may do so by extending loans that otherwise would default in a so-called evergreening, discussed in chapter 11.2.3.

Banks with a disclosed low success rate are liquidated with probability γ_L and those with a high success rate with a probability γ_H, this may take into account other factors that may lead to liquidation, for example assessments of the future prospects of the bank or general market conditions. We further assume that a high success rate is associated with additional costs C, which may arise from higher level of monitoring borrowers by banks. If $\pi_L (1 + r_L) L < (1 + r_D) D \leq \pi_H (1 + r_L) L$, with r_L denoting the loan rate, L the loan amount and D deposits, banks with low repayment rates are failing as they cannot repay their deposits in full and banks with high repayment rates would not fail.

Banks will disclose their low success rate if

$$\begin{aligned} &\gamma_L \pi_L (1 + r_L) L - (1 + r_D) D \\ &\quad \geq \gamma_H (\pi_L (1 + r_L) L + (\pi_H - \pi_L) \pi_L (1 + r_L) L) - (1 + r_D) D. \end{aligned} \tag{36.1}$$

The left-hand side shows the profits of the bank if the low repayment rate is revealed, taking into account the possible liquidation, γ_L and the right-hand side the profits if the bank pretends to have a high repayment rate, even though the repayment rate is actually low. In this case it will hold back a fraction $\pi_H - \pi_L$ of the loans and extend them, which we assume gives again a success rate of π_L in the next time period. Holding back this amount makes the bank look like its loans have a high repayment rate. We can solve (36.1) such that

$$\gamma_L \geq \gamma_L^* = (1 + (\pi_H - \pi_L)) \gamma_H. \tag{36.2}$$

We see that disclosure happens if it is sufficiently more likely that the bank is liquidated if the low repayment rate is disclosed than if the high repayment rate is applied.

The bank will choose the high repayment rate if it is more profitable than choosing the low repayment rate. The profits when choosing the high repayment rate is given by the proceeds of these loans, provided the bank is not liquidated, less the repayment of deposits and the costs of providing loans with high repayments. When choosing loans with low repayment rates, these costs are not incurred and hence the bank profits are the proceeds from these loans if the bank is not liquidated less the repayment of deposits. For high repayment rates to be chosen we thus need

$$\gamma_H \pi_H (1 + r_L) - (1 + r_D) D - C \geq \gamma_L \pi_L (1 + r_L) - (1 + r_D) D, \tag{36.3}$$

which solves for

$$\gamma_L \leq \gamma_L^{**} = \frac{\pi_H}{\pi_L} \gamma_H - \frac{C}{\pi_H (1 + r_L) L}. \tag{36.4}$$

We see that the conditions on γ_L in equations (36.2) and (36.4) are compatible if

$$\gamma_H (\pi_H - \pi_L) (1 - \pi_L) (1 + r_L) L \geq C \tag{36.5}$$

and we assume that the costs are sufficiently small such that we can find a range of γ_L such that banks would disclose that they are providing loans with low repayments rates, but are actually choosing low-risk loans.

If the condition in equation (36.2) is violated, then the bank does not disclose their low repayment rate the condition to select the loan with the low repayment rate becomes

$$\begin{aligned}&\gamma_i \left(\pi_L \left(1 + r_L\right) L + \left(\pi_H - \pi_L\right) \pi_L \left(1 + r_L\right) L\right) - \left(1 + r_D\right) D \\ &\quad \geq \gamma_i \pi_H \left(1 + r_L\right) L - \left(1 + r_D\right) D - C,\end{aligned} \tag{36.6}$$

which would require

$$\pi_L \left(1 - \pi_L\right) \left(1 + r_L\right) L + C \geq 0, \tag{36.7}$$

which is clearly fulfilled and hence banks would always choose the low repayment rate.

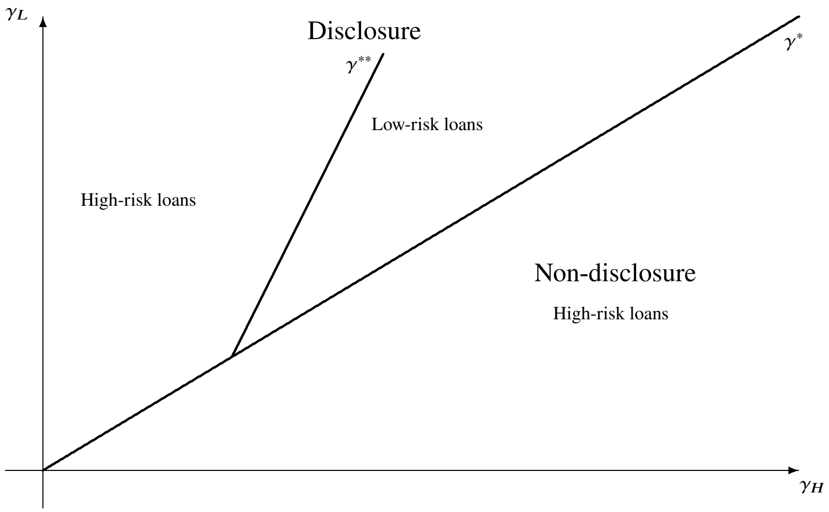


Fig. 36.1: Default disclosures

Figure 36.1 illustrates these results. We clearly see that only if the likelihood of the bank with high repayment rates being liquidated is sufficiently low, relative to the liquidation if the repayment rate is low, do banks disclose if they have chosen loans with high repayment rates. Here the benefits for disclosing this fact allows banks to gain benefits if the choose high repayment rates as in this case they are benefitting from the lower probability of liquidation. A higher liquidation rate for low-risk banks, those with high repayment rates, reduces these benefits and banks are not willing to disclose this information.

Readings Mitchell (2001), Freixas & Rochet (2008b, Ch. 9.5.2)

36.2 Bank closure decisions

A common assumption is that banks are only closed once they have failed. In many models such closures have been assumed that banks are liquidated, however, banks might be saved from liquidation if they are taken over other banks or a government fund. Once a bank is liquidated it has little value and the losses from any deposit insurance or other bailout arrangements can be substantial. Given that banks usually require a licence to operate, a regulator could step in at an earlier stage and decide whether a bank should be closed. If the regulator assesses a bank to be too risky to continue operating, it could close it down before the bank actually fails. This closing down might involve the sale at an attractive price to another bank. Regulators will make decisions on the closure of such banks strategically by comparing the costs of closing down a bank compared to the bank remaining open, but potentially accumulating larger losses. Such strategic considerations are discussed in chapter 36.2.1. Regulators may also use the threat of closing down banks to give incentives to banks for the exertion of effort to reduce their risks from lending. We will see in chapter 34.1.1 how such incentives can lead to distorted decision-making by regulators.

36.2.1 Strategic bank closures

It is common to assume that regulators close a bank in response to the bank not being able to repay its depositors. However, as deposits are rarely actually withdrawn, unless the bank is subject to a bank run, it could continue to operate, even if the value of the loans it has provided is below the liabilities arising from deposits. Therefore, regulators could keep the bank open in the hope of the bank generating profits in the future that will either avoid the regulator having to close the bank at all or at least to reduce the losses it may have to cover. We will look at such a situation here and analyse under which conditions regulators will close down a bank.

Banks can provide loans with a probability of it being repaid of $\pi_i > \frac{1}{2}$. The interest on a loan is denoted r_L^i and the loan amount is L. These loans are fully financed by deposits D, on which interest r_D is payable. There are two types of loans such that $\pi_H > \pi_L$ and it is socially preferable to provide loan the loan with the high repayment rate as we assume that $\pi_H \left(1 + r_L^H\right) > \pi_L \left(1 + r_L^L\right)$. However as we also assume that $\pi \left(\left(1 + r_L^H\right) L - (1 + r_D) D\right) < \pi_L \left(\left(1 + r_L^L\right) L - (1 + r_D) D\right)$, banks would prefer to provide the more risky loan, that is the loan with the lower repayment rate.

Banks and regulators have a time horizon of two time periods, with loans given for a single time period, such that banks could provide loans of different types in each time period. A regulator can close a bank after the first time period at cost C. In addition, the regulator has to compensate the depositors if the bank cannot meet its obligations. Alternatively the regulator could keep the bank open and if after the second time period depositors of both time periods can be repaid, the bank does not need to be liquidated, otherwise it has to be liquidated at cost C and depositors of both time periods will need to be compensated.

We initially determine the choice of loans by banks before than introducing a regulator to decide whether to close a bank.

Loan choice We assume that $\left(1+r_L^L\right)L < 2\left(1+r_D\right)D$ and $\left(1+r_L^H\right))L < 2\left(1+r_D\right)D$, such that the loan not being repaid in either period will cause the bank to fail; if the loan only from one time period is repaid, it cannot be used to repay depositors from both time periods as the loan rate is not high enough. The bank's profits are such that it receives its revenue $\left(1+r_L^i\right)L-\left(1+r_D\right)$ from each time period if it succeeds in both time periods only, hence for a bank investing in loan i in time period 1 and j in time period 2, its profits Π_B^{ij} are given by

$$\begin{aligned}
\Pi_B^{HH} &= 2\pi_H^2\left(\left(1+r_L^H\right)L-\left(1+r_D\right)D\right), \\
\Pi_B^{HL} &= \pi_H\pi_L\left(\left(1+r_L^H\right)L+\left(1+r_L^L\right)L-2\left(1+r_D\right)D\right), \\
\Pi_B^{LH} &= \pi_H\pi_L\left(\left(1+r_L^H\right)L+\left(1+r_L^L\right)L-2\left(1+r_D\right)D\right), \\
\Pi_B^{LL} &= 2\pi_L^2\left(\left(1+r_L^L\right)L-\left(1+r_D\right)D\right).
\end{aligned} \tag{36.8}$$

A bank has to receive loan repayments in both time periods in order to be able to repay its depositors and to receive the profits as the bank would fail if any of the two loans is not repaid.

If the bank has given loan a low risk loan in time period 1, then it will provide a low-risk loan in time period 2 if it is more profitable to do so, $\Pi_B^{HH} \geq \Pi_B^{HL}$. This condition solves for

$$\pi_L \leq \pi_L^* = 2\pi_H\frac{\left(1+r_L^H\right)L-\left(1+r_D\right)D}{\left(2+r_L^H+r_L^L\right)L-2\left(1+r_D\right)D}. \tag{36.9}$$

Similarly, if the bank has given a high risk loan in time period 1, then it will provide a high-risk loan in time period 2 if it is more profitable to do so, $\Pi_B^{LL} \geq \Pi_B^{LH}$, thus

$$\pi_L \geq \pi_L^{**} = \frac{1}{2}\pi_H\frac{\left(2+r_L^H+r_L^L\right)L-2\left(1+r_D\right)D}{\left(1+r_L\right)L-\left(1+r_D\right)D}. \tag{36.10}$$

It is easy to show that $\pi_L^{**} \geq \pi_L^*$. In the case that $\pi_L^*\pi_L\pi_L^{**}$, the bank would switch from providing loan a high-risk loan (low-risk loan) to providing loan low-risk loan (high-risk) in time period 2.

Knowing the choice of the bank in time period 2, let us now consider the choice in time period 1. If $\pi_L \leq \pi_L^* \leq \pi_L^{**}$, the bank will always choose the low-risk loan in time period 2, thus in order for the bank to choose the low-risk loan in time period 1 we need $\Pi_B^{HH} \leq \Pi_B^{HL} = \Pi_N^{LH}$, which gives us the same condition as in equation (36.9) and hence for $\pi_L \leq \pi_L^*$ the bank would choose low-risk loans in both time periods.

If $\pi_L^* < \pi_L^{**} \leq \pi_L$, then the bank will always choose the high-risk loan in the second time period. In order to choose high-risk loans also in the first time period, we need that $\Pi_B^{LL} > \Pi_B^{LH} = \Pi_B^{HL}$, yielding again equation (36.10). Thus for $\pi_L > \pi_L^{**}$ the bank would choose high-risk loans in both time periods.

In the intermediate case that $\pi_L^* < \pi_L < \pi_L^{**}$ the bank would switch the type of loan provided. As $\Pi_B^{LH} = \Pi_B^{HL}$, the bank is ex-ante indifferent between either combination. In order to resolve the choices banks would make in this scenario, we introduce a regulator who may close the bank.

Regulator intervention If the regulator closes a bank in the first time period, it incurs costs C and paying out the depositors if the bank fails, hence its costs for banks choosing low-risk loans and high-risk loans, respectively, are given by

$$\begin{aligned} \Pi_R^H &= C + (1 - \pi_H)(1 + r_D) D, \\ \Pi_R^L &= C + (1 - \pi_L)(1 + r_D) D. \end{aligned} \tag{36.11}$$

If the bank is not closed, the closure costs C are only borne if the bank does not succeed in both time periods, which happens with probability $1 - \pi_i \pi_j$. In addition, the regulator has to compensate both sets of depositors if the bank fails in both time periods, the probability being $(1 - \pi_i)^2$. If the bank only fails in one time period, the probability of this is $\pi_i \left(1 - \pi_j\right)$ and $\pi_j (1 = \pi_i)$, respectively, they can recover the profits of the bank from the successful loan. Thus the costs of regulators for the different loan combinations banks may take are given by

$$\begin{aligned}
\Pi_R^{HH} &= \left(1-\pi_H^2\right)C + (1-\pi_H)^2\,2\,(1+r_D)\,D \\
&\quad +2\pi_H\,(1-\pi_H)\left((1+r_D)\,D - \left(1+r_L^H\right)L - (1+r_D)\,D\right) \\
&= \left(1-\pi_H^2\right)C + 2\left(1-\pi_H^2\right)(1+r_D)\,D \\
&\quad -2\pi_H\,(1-\pi_H)\left(1+r_L^H\right)L, \\
\Pi_R^{LH} &= \Pi_R^{HL} = (1-\pi_H\pi_L)\,C + (1-\pi_H)\,(1-\pi_L)\,2\,(1+r_D)\,D \\
&\quad +\pi_H\,(1-\pi_L)\left((1+r_D) - \left(\left(1+r_L^H\right)L - (1+r_D)\,D\right)\right) \\
&\quad +\pi_L\,(1-\pi_H)\left((1+r_D)\,D - \left(\left(1+r_L^L\right)L - (1+r_D)\,D\right)\right) \\
&= (1-\pi_H\pi_L)\,C + 2\,(1-\pi_H\pi_L)\,(1+r_D)\,D \\
&\quad -\pi_H\,(1=\pi_L)\left(1+r_L^H\right)L - \pi_L\,(1-\pi_H)\left(1+r_L^L\right)L, \\
\Pi_R^{LL} &= \left(1-\pi_L^2\right)C + (1=\pi_L)^2\,2\,(2+r_D)\,D \\
&\quad +2\pi_L\,(1-\pi_L)\left((1+r_D)\,D\left(\left(1+r_L^L\right)L - (1+r_D)\,D\right)\right) \\
&= \left(1-\pi_L^2\right)C + 2\left(1-\pi_L^2\right)(1+r_D)\,D \\
&\quad -2\pi_L\,(1-\pi_L)\,(1+r_L)\,L.
\end{aligned} \tag{36.12}$$

Note that if the loan is not repaid in only one time period, the depositors in the other time period are repaid and thus do not get compensated by the regulator.

If $\pi_L < \pi_L^*$ and the bank chooses low-risk loans in both time periods, then the regulator would close the bank in time period 1 if its losses are lower, $\Pi_R^H < \Pi_R^{HH}$. This condition solves for

$$\begin{aligned}
C &< \frac{1+\pi_H-2\pi_H^2}{\pi_H^2}\,(1+r_D)\,2 - \frac{1-\pi_H}{\pi_H}\left(1+r_L^H\right)L \\
&< \frac{1-\pi_H}{\pi_H^2}\,(1+r_D)\,D
\end{aligned} \tag{36.13}$$

where the second inequality arises from an assumption that $\left(1+r_L^H\right)L > (1+r_D)\,D$ to ensure banks are profitable. Similarly for $\pi_L > \pi_L^{**}$ the bank chooses high-risk loans in both time periods and the regulator closes it early if $\Pi_R^L < \Pi_R^{LL}$, hence with the same procedure as above.

$$C < \frac{1-\pi_L}{\pi_L}\,(1+r_D)\,D \tag{36.14}$$

If we now assume that π_L is sufficiently small such that this condition is never fulfilled, then as $\pi_H > \pi_L$, condition (36.13) is also never fulfilled. Hence in these cases banks are never closed early. We thus see that banks choosing high-risk loan in both time periods or low-risk loans in both time periods will not be closed early.

In the case that a bank chooses the low-risk loan followed by a high-risk loan, $\pi_L^* < \pi_L < \pi_L^{**}$, the bank is closed early if $\Pi_R^H < \Pi_R^{HL}$. This requires

$$C < C^* = \frac{(2\pi_L - 1)(1 - \pi_H)}{\pi_H \pi_L}(1 + r_D) D, \tag{36.15}$$

where we assume that $\pi_i \left(1 + r_L^i\right) = (1 + r_D) D$ and hence competitive loan pricing that gives banks zero expected profits.

In the same way, if the bank chooses the high-risk loan followed by the low-risk loan, the bank would be closed early if $\Pi_R^L < \Pi_R^{LH}$ and hence with competitive loan pricing the condition becomes

$$C < C^{**} = \frac{(2\pi_H - 1)(1 = \pi_L)}{\pi_H \pi_L}(1 + r_+ D) D. \tag{36.16}$$

We can easily see that $C^* < C^{**}$. Hence, if $C < C^*$, the bank would be closed early and for $C > C^{**}$ it would not be closed early. The case that $C^* < C < C^{**}$ corresponds to a scenario in which the bank would be closed early if the bank chooses loan the low-risk loan first, followed by the high-risk loan. Knowing the bank will be closed if the low-risk loan is chosen first, the bank's profits are given by

$$\hat{\Pi}_B^H = \pi_H \left(\left(1 + r_L^H\right) L - (1 + r_D) D\right). \tag{36.17}$$

If $\hat{\Pi}_B^H \geq \Pi_B^{LH}$, the bank would choose the low-risk loan first, even if it is closed early. This condition gives us

$$\pi_L \leq \frac{\left(1 + r_L^H\right) L - (1 + r_D) D}{\left(2 + r_L^L + r_L^H\right) L - 2(1 + r_D) D} = \frac{\pi_L^*}{2\pi_H}. \tag{36.18}$$

With our assumption that $\pi_L > \frac{1}{2}$, this threshold is below π_L^* and as $\pi_L^* < \pi_L < \pi_L^{**}$, the bank would not choose the low-risk loan in time period 1. Consequently, the bank remains open for $C^* < C < C^{**}$ and commences with the high-risk loan, followed by a low-risk loan.

Figure 36.2 summarizes our results. If π_L is sufficiently low, then given the high default rate of high-risk loans, choosing the low-risk loan is more attractive and furthermore the regulator will face little risk when keeping the bank open. For large π_L, the higher returns and relatively low risks of the high-risk loans makes its choice more attractive. The regulator does not close the bank as on the one hand the probability of losses, and hence of incurring any costs is small, and on the other hand he is likely to recover any losses made in time period 1 as defaults are unlikely. In the intermediate case, the costs of early closure are too high if $C > C^*$ and if $C < C^*$ they are so small that the regulator prefers to close the bank early. In the case of medium closing costs, knowing the bank would choose, the low-risk loan in the second time period induces the regulator to not close the bank in order to recover any losses that may have been incurred.

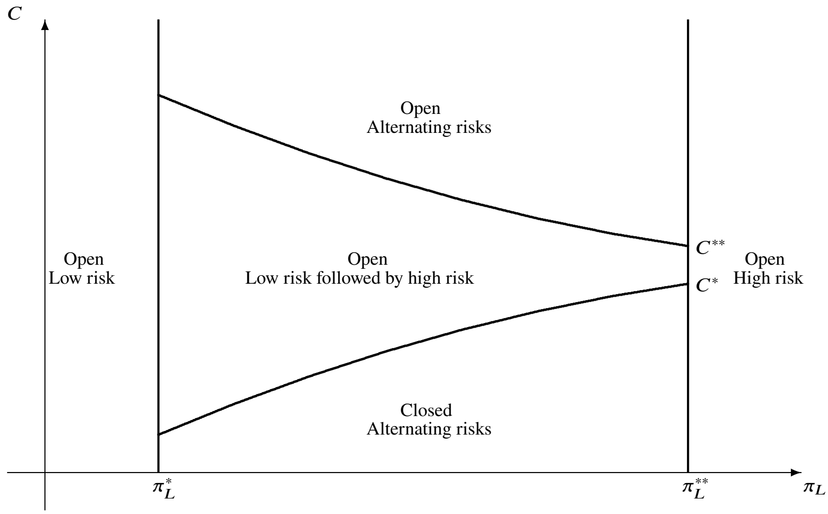


Fig. 36.2: Loan choice and bank closure

Summary We thus see that a regulator would decide to close a banks only if its own direct costs of doing so are relatively small. They would weigh their direct costs of closing a bank against the costs of having to reimburse depositors in case of a failure, but will also have to take into account the possibility that a bank may recover from losses, reducing the costs of a bank failure. Banks will also react to the risk of being closed down and in order to avoid such a possibility would ensure that loans are chosen such that regulators do not close the bank down in order to benefit from a recovery of previous losses. This adjustment of the behaviour of banks makes a closure of banks very unlikely.

Reading Mailath & Mester (1994)

36.2.2 Signals and incentives to exert effort

regulators can use the threat of closing a bank strategically as a tool in ensure banks reduce their risks. Through the possibility of banks being closed if they take on too high risks, banks will exert effort to reduce their risks and avoid the closure. We will explore here how these incentives may distort the closure decision to accommodate the need for risk reduction.

Let us assume that banks operate in two time periods and in each time period make profits of

$$\Pi_B^t = \pi_t \left(1 + r_L\right) L - \left(1 + r_D\right) D. \tag{36.19}$$

Here pi_t represents the unknown repayment rate of loans L. The loan rate is given as r_L and the deposit rate as r_D, payable on deposits D, which finance the loans the bank provides.

Optimal decision At the beginning of time period 2 a decision needs to be taken whether the bank should remain open or be closed down. At this point, the realization of π_1 is known and a signal about π_2 has been obtained, $\hat{\pi}_2$, such that we know the density of π_2, given this signal, denoted by $h_i(\pi_2|\hat{\pi}_2)$. Here i refers to the decision to close the bank, $i = C$, or the bank surviving, $i = S$. We assume that for the distribution function H_i we have

$$\frac{\partial H_C(\pi_2|\hat{\pi}_2)}{\partial \hat{\pi}_2} < \frac{\partial H_S(\pi_2|\hat{\pi}_2)}{\partial \hat{\pi}_2}, \tag{36.20}$$

implying that the higher the signal $\hat{\pi}_2$, the higher value of π_2 is likely to be. This result is driven by the assumption that banks exert effort to reduce the default of loans and a higher signal makes such effort more desirable.

Let us denote by $\Pi_B^{2,i}$ the profits of the banks in time period 2 if the regulator chooses action i, survival or closure of the bank. We can then determine the difference in the profits for banks between the regulator keeping the bank open and closing it,

$$\begin{aligned} \Delta\Pi_2 &= E\left[\Pi_B^{2,S}|\hat{\pi}_2\right] - E\left[\Pi_B^{2,C}|\hat{\pi}_2\right] \qquad (36.21)\\ &= (1 + r_L) L \int_0^1 \pi_2 \left(dH_S(\pi_2|\hat{\pi}_2) - dH_C(\pi_2|\hat{\pi}_2)\right) \\ &= (1 + r_L) L \int_0^1 \left(H_S(\pi_2|\hat{\pi}_2) - H_C(\pi_2|\hat{\pi}_2)\right), \end{aligned}$$

where the final equality is the result of integrating by parts. Note that closure here merely implies that a regulator, acting on behalf of depositors, would take control of the bank, it does not imply liquidation. The profits of the bank are given from the average repayment rates of loans as given by the distributions for survival and closure, respectively. The repayment of deposits are identical regardless of the repayment rate and hence are cancelling each other out for this difference.

We can now define a signal $\hat{\pi}_2^*$ such that $\Delta\Pi_2 = 0$. As the expression in the integral of equation (36.21) is increasing in s due to our assumption in equation (36.20), such a $\hat{\pi}_2^*$ would in general exist and for $\hat{\pi}_2 \leq \hat{\pi}_2^*$ the bank should survive and for $\hat{\pi}_2 < \hat{\pi}_2^*$ it should be closed.

Regulator decision As alluded to, banks need to exert effort to increase the repayment rate of loans and exerting such effort imposes a cost C on banks. Of course, effort is desirable and we will set the incentives such that banks will choose to exert this effort. The effort level will effect the probability of the loan being repaid in period 1, π_1, and we denote its density by $f_i(\pi_i)$. It will also affect the signal received at the beginning of period 2, $g_i(\hat{\pi}_2)$. We now assume that both these densities are increasing the higher the effort is, thus

$$\frac{\partial \frac{f_H(\pi_1)}{f_L(\pi_1)}}{\partial \pi_1} > 1, \qquad (36.22)$$
$$\frac{\partial \frac{g_H(\hat{\pi}_2)}{g_L(\hat{\pi}_2)}}{\partial \hat{\pi}_2} > 1,$$

where the subscript H indicates a high level of effort and the subscript L that no effort is exerted.

If p denotes the probability of the bank surviving, then the expected profits of exerting high effort must exceed those of no effort, hence we require that

$$\int_0^1 \int_0^1 p\left(\pi_1 (1 + r_L) L - (1 + r_D) D\right) f_H(\pi_1)\, gh(\hat{\pi}_2)\, d\pi_1 d\hat{\pi}_2 - C \qquad (36.23)$$
$$\geq \int_0^1 \int_0^1 p\left(\pi_1 (1 + r_L) L - (1 + r_D) D\right) f_L(\pi_1)\, g_L(\hat{\pi}_2)\, d\pi_1 d\hat{\pi}_2.$$

We here consider the profits of the bank to be averaged not only over all possible repayment rates in time period 1, but also the signals received about the repayment rate in time period 2. The bank only obtains its profits if the bank survives, and that decision will depend on the joint distribution of π_1, and s which for simplicity we assume to be independent, conditional on the effort level.

We can simplify equation (36.23) to

$$\int_0^1 \int_0^1 p\pi_1 (1 + r_L) \left(f_F(\pi_1)\, gh(\hat{\pi}_2) - f_L(\pi_1)\, g_L(\hat{\pi}_2)\right) d\pi_1 d\hat{\pi}_2 - C \geq 0 \qquad (36.24)$$

noting that the probability of the bank surviving, p, will depend on π_1 and $\hat{\pi}_2$. If we now seek to maximize the benefits of the bank surviving, given as the difference in profits between their survival and closure, $\Delta\Pi_2$ from equation (36.21), our objective function becomes

$$\Pi_B = \int_0^1 \int_0^1 p\Delta\Pi_2 f_H(\pi_i)\, g_H(\hat{\pi}_2)\, d\pi_1 d\hat{\pi}_2. \qquad (36.25)$$

These profits will be maximized subject to the bank having to exert effort, thus equation (36.23) serves as a constraint to this optimisation. Thus we can use the Lagrange multiplier ξ to define our objective function as

$$\mathcal{L} = \int_0^1 \int_0^1 p\left((\Delta\Pi_2 + \xi\pi_1) f_H(\pi_1)\, gh(\hat{\pi}_2) \right. \qquad (36.26)$$
$$\left. -\xi\pi_1 f_L(\pi_1)\, gL(\hat{\pi}_2)\right) d\pi_1 d\hat{\pi}_2 - \xi C.$$

We thus obtain the first order condition as

$$\frac{\partial \mathcal{L}}{\partial p} = \int_0^1 \int_0^1 ((\Delta\Pi_2 + \xi\pi_1) fh(\pi_1) g_H(\hat{\pi}_2) \\ -\xi\pi_1 f_L(\pi_1) g_H(\hat{\pi}_2)) \, d\pi_1 d\hat{\pi}_2. \tag{36.27}$$

With only corner solutions for p available, either the bank survives, $p = 1$ or it is closed, $p = 0$, we see that $p = 1$ is optimal and the bank survives, if the argument in equation (36.27) is positive as then the highest value will be chosen. This condition con be solved for

$$\frac{f_H(\pi_1)}{f_L(\pi_1)}\left(1 + \frac{\Delta\Pi_2}{\xi\pi_1}\right) \geq \frac{g_L(\hat{\pi}_2)}{g_H(\hat{\pi}_2)}. \tag{36.28}$$

Defining $\hat{\pi}_2^{**}$ as the signal for which equation (36.28) holds with equality, this $\hat{\pi}_2^{**}$ will be the tipping point at which the optimal solution switches from $p = 0$ to $p = 1$ and vice versa; we see that if $\hat{\pi}_2 \geq \hat{\pi}_2^{**}$ the bank survives and for $\hat{\pi}_2 < \hat{\pi}_2^{**}$ it is closed.

Taking the total differential of the equality in equation (36.28), we obtain

$$\left(\frac{f_H(\pi_1)}{f_L(\pi_1)} \frac{1}{\xi\pi_1} \frac{\partial\Delta\Pi_2}{\partial\hat{\pi}_2^{**}} + \left(\frac{g_H(\hat{\pi}_2)}{g_L(\hat{\pi}_2)}\right)^2\right) d\hat{\pi}_2^{**} \\ + \left(\frac{\partial \frac{f_H(\pi_1)}{f_L(\pi_1)}}{\partial\pi_1}\left(1 + \frac{\Delta\Pi_2}{\xi\pi_1}\right) - \frac{f_H(\pi_1)}{f_L(\pi_1)} \frac{\Delta\Pi_2}{\xi\pi_1^2}\right) d\pi_1 = 0 \tag{36.29}$$

As from equations (36.20) and (36.21) we have $\frac{\partial\Pi_2}{\partial s} > 0$, we see that the first term is positive. With the additional assumption that

$$\frac{\partial \frac{f_H(\pi_1)}{f_L(\pi_1)}}{\partial\pi_1} > \frac{f_H(\pi_1)}{f_L(\pi_L)} \frac{1}{\pi_1} \tag{36.30}$$

the second term is also positive implying that

$$\frac{\partial\hat{\pi}_2^{**}}{\partial\pi_1} < 0. \tag{36.31}$$

We have thus obtained the optimal threshold for the signal about the repayment rate in the second time period that decides whether a bank will survive or be closed. The bank exerts effort in order to reduce the likelihood of the bank being closed by the regulator. Thus the regulator making the decision to close a bank will employ a strategy, the threshold s^*, that maximizes the value of the bank itself. Figure 36.3 shows the relationship of $\hat{\pi}_2^{**}$ and $\hat{\pi}_2^*$. $\hat{\pi}_2^*$ represents the optimal closure decision once the signal is known and s^* the optimal closure decision ex-ante before this information is known; it thus represents the strategy of a regulator not having full information about the bank's future risks and thus profitability.

We see that over wide areas the closure decision of the regulator will be optimal, there are, however, also areas in which the regulator decides to close the banks when it should survives and areas where the bank survives but it should be closed. These

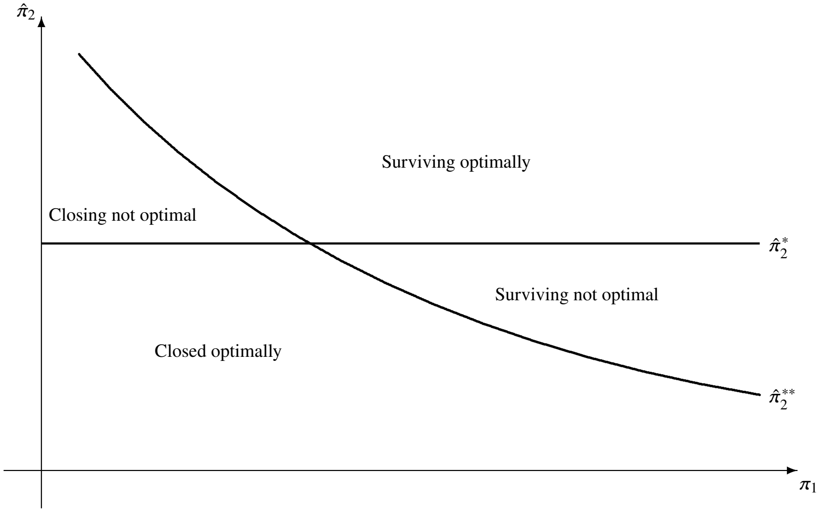


Fig. 36.3: Optimal and actual bank closure decisions

areas where there is the largest discrepancy between the repayment rate in time period 1 and the signal about the repayment rate in time period 2. If this signal is particularly high compared to the repayment rate in time period 1, the bank is closed even though it would be optimal for the bank to survive. Similarly if the signal is particularly low compared to the repayment rate in time period 1, the banks survives, even though it should be closed down. These deviations from the optimal decisions emerge because banks need to be incentivised to exert effort and if the signal is particularly low, the reduced threat of the bank being closed down increases the efforts of the bank to exert effort and reduce risks. On the other hand, if the signal is particularly high, no such incentives are needed and banks can be closed down more easily.itional safety causing the suboptimal closure (survival) of the bank.

Summary We have thus seen that regulators can use the threat of bank closure as an incentive for banks to exert effort in order to reduce the risks they face. Using the closure decision as an incentive device for banks to exert efforts distorts the closure decision, however. Banks might be closed too readily even if the signal received about the prospects of the bank is high; on the other hand banks might be kept open despite receiving a negative signal about their prospects. This distortion in closing decisions ensures that banks have optimal incentives to exert efforts and thus maximize the value of banks, but this comes at the price of making closure decisions that are suboptimal.

Reading Dewatripont & Tirole (1994, Ch. 7)

Résumé

We have seen that banks will be closed strategically only if the costs to regulators are sufficiently small. In all other cases, either the costs of closing a banks are too high for the regulator or the banks react to the threat of a bank closure by reducing the risks they are taking in the future, enticing the regulator to not close the bank. Thus the strategic behaviour of the regulator is complemented by the strategic behaviour of banks. It is not only that such strategic behaviour of banks reduce the threat of a bank closure to situations in which any additional costs to the regulator are small, they might also distort the decision-making. If the regulator uses the threat of closing the bank in order to influence the decisions of banks, it will be that sue to the reactio09n of banks, the closure decision is no longer be optimal. In order to have an effective threat, banks might be closed too readily on the one hand, but at times regulators might also have to be more lenient than is optimal as to not jeopardize the incentives the possible closure of the bank provides.

36.3 Bad banks

At times of stress to the banking system, government may set up a so-called bad bank. Banks can sell their non-performing loans to a bad bank and obtain cash reserves for these loans, freeing their balance sheet from risky assets to provide new loans without being as constrained by capital requirements as before the sale. The use of bad banks is often seen as an alternative to closing down a bank that has too many such non-performing loans. We will analyse the operation of such a bad bank and see when it is optimal to make use of such an arrangement.

Assume a bank has given loans L with interest r_L , that have a repayment rate π, the current value of these loans is given by $\pi(1+r_L)L$. In a solvent bank we have that the value of these assets, have to equal the liabilities, deposits D and equity $E=\kappa D$, where κ is the capital ratio. Using that $L=D+E=(1+\kappa)D$ we get that the value of deposits is given by

$$D=\frac{\pi(1+r_L)L}{1+\kappa}. \tag{36.32}$$

Let us for simplicity assume that the deposit rate is zero as depositors are protected by deposit insurance. In this case the bank profits when retaining these loans are

$$\Pi_B=\pi((1+r_L)L-D). \tag{36.33}$$

We can now assess the conditions under which the bank would be willing to sell their loans to a bad bank.

Loan sale Instead of holding on to the loans, the bank could sell these at a price L^* to the bad bank, who provides the bank with a bond that is risk-free and yields no interest. Freed of the risky loans, the bank can now attract new deposits and hand out new loans $\hat{L}$, that yield interest $\hat{r}_L$ and succeed with probability $\hat{\pi}>\pi$ and are thus less risky than the loans sold to the bad bank. As the government bond is risk-free, there are no capital requirements and the capital requirements would be $\hat{E} = \kappa \hat{L}$. The total deposits consist of the old deposits and the newly attracted deposits that are fully lent out $\hat{D} = D + \hat{L}$. The assets of the bank consist of the government bond L^* and the new loans given $\hat{L}$ such that $L^* + \hat{L} = \hat{D} + \hat{E}$.

To avoid instant bankruptcy we need to ensure that the funds received from the bad bank, L^* , cover the repayment of existing depositors, D. Thus $L^* \geq D$. From this we get that

$$L^* = \hat{D} + \hat{E} - \hat{L} \geq \hat{D} - (1 - \kappa) \hat{L} = D + \kappa \hat{L},$$

from which we obtain that

$$\hat{L} \leq \frac{L^* - D}{\kappa}. \tag{36.34}$$

In this case bank profits are given by

$$\hat{\Pi}_B = \hat{\pi} \left((1 + \hat{r}_L) \hat{L} + L^* - \hat{D} \right) + (1 - \hat{\pi}) \max \left\{ 0; L^* - \hat{D} \right\} - C, \tag{36.35}$$

where the first term denotes the successful repayment of the new loan, the value of the government bond, and the repayment of the deposits. The second term denotes the case where the loan is not repaid, but if the value of the government bond is sufficiently high, the bank may retain some of the value as depositors can be fully repaid. The final term C denotes the cost of selling to the bad bank; such costs may include a loss in reputation or higher future funding costs.

Noting that $\hat{D} = D + \hat{L}$, we get can maximize the profits of the bank over the optimal amount of new loans the bank provides. The first order condition is given by

$$\frac{\partial \hat{\pi}_B}{\partial \hat{L}} \geq \hat{\pi} \hat{r}_D - (1 - \hat{\pi}) = \hat{\pi} (1 + \hat{r}_L) - 1 > 0, \tag{36.36}$$

where the first inequality emerges from the fact that the second term in the profits of equation (36.35) might be zero if $L^* - D < 0$. Assuming new loans to be profitable, gives us a positive sign of this first order condition. Hence banks would choose the new loans $\hat{L}$ to be as large as possible and equation (36.34) is fulfilled with equality. Using this equality we then have

$$L^* - \hat{D} = \frac{1 - \kappa}{\kappa} (D - L^*) < 0, \tag{36.37}$$

after inserting for $\hat{D}$ and $\hat{L}$. Hence the second term in the profits of banks, equation (36.35), becomes zero such that we get

$$\hat{\Pi}_B = \hat{\pi} \left(\hat{r}_L \hat{L} + L^* - D \right) - C. \tag{36.38}$$

The bank will sell its loans to the bad bank if the profits of doing so exceed to of retaining the loans, $\hat{\Pi}_B \geq \Pi_B$. Inserting for all expressions from equations (36.38) and (36.33), this eventually becomes

$$L^* \geq \hat{L}^* = \frac{\kappa}{\hat{\pi}\left(\hat{r}_L + \kappa\right)} C + \frac{\kappa\left(1 - \pi + \kappa\right) + \hat{\pi}\left(\hat{r}_L + \kappa\right)}{\hat{\pi}\left(1 + \kappa\right)\left(\hat{r}_L + \kappa\right)} \pi\left(1 + r_L\right) L. \tag{36.39}$$

Not surprisingly, if the payment of the bad bank is sufficiently high, the bank will sell the bad loans.

Repurchase agreement Alternatively, rather than selling the loans to the bad bank, the bad bank could offer a repurchase agreement. In this case the bank re-buys the loans at the same price L^* at maturity. In this case we have four possible outcomes:

1. With probability $\pi\hat{\pi}$ both loans are repaid, the bank obtains $\left(1 + r_L\right) L + \left(1 + \hat{r}_L\right) \hat{L}$ and is repaying the deposits.
2. With probability $\left(1 - \pi\right) \hat{\pi}$ only the new loan is repaid and the bank can repay depositors if $\left(1 + \hat{r}_L\right) \hat{L} \geq \hat{D}$, or $\hat{L} \geq \frac{D}{\hat{r}_L}$.
3. With probability $\pi\left(1 - \hat{\pi}\right)$ only the original loan is repaid and the bank obtains $\left(1 + r_L\right) L$, staying solvent if $\left(1 + r_L\right) L > \hat{D}$, or $\hat{L} < \left(1 + r_L\right) L - D$.
4. With probability $\left(1 - \pi\right)\left(1 - \hat{\pi}\right)$ both loans are not repaid, leaving the bank insolvent.

We assume now that the bank remains solvent if only one of the loans is repaid by assuming that

$$\frac{D}{\hat{r}_L} < \hat{L} < \left(1 + r_L\right) L - D,$$

which can be solved for

$$1 + \hat{r}_L < \frac{1 + \kappa}{1 + \kappa - \pi} \tag{36.40}$$

As we will see below, $\hat{L}$ is again maximized such that with equation (36.34) to avoid instant bankruptcy we need

$$\hat{L} = \frac{L^* - D}{\kappa} > \frac{D}{\hat{r}_L},$$

which requires

$$L^* > \frac{\hat{r}_L + \kappa}{\hat{r}_L} D. \tag{36.41}$$

We can now show that

$$\begin{aligned}
\left(1 + r_L\right) L - \hat{D} &= \left(1 + r_L\right) L + \frac{1 - \kappa}{\kappa} D + \frac{1}{\kappa} L^* \\
&\leq \left(1 + r_L\right) L + \frac{1 - \kappa}{\kappa} D - \frac{1}{\kappa} \frac{\hat{r}_L + \kappa}{\hat{r}_L} D \\
&= \left(1 - \frac{\pi}{1 + \kappa} \frac{1 + \hat{r}_L}{\hat{r}_L}\right) \left(1 + r_L\right) L \\
&< 0,
\end{aligned} \tag{36.42}$$

where the final equality arises from the term in brackets to be negative for all reasonable values. Similarly we get

$$
\begin{aligned}
(1+\hat{r}_L)\,\hat{L}-\hat{D} &= \frac{\hat{r}_L}{\kappa}L^*-\frac{\hat{r}_L+\kappa}{\kappa}D \\
&= 0.
\end{aligned}
\tag{36.43}
$$

Using equations (36.42) and (36.43) we can obtain the bank profits as

$$
\begin{aligned}
\hat{\hat{\Pi}}_B &= \pi\hat{\pi}\left((1+r_L)\,L+(1+\hat{r})\,\hat{L}-\hat{D}\right) \\
&\quad +\pi\,(1-\hat{\pi})\max\left\{0;(1+r_L)\,L-\hat{D}\right\} \\
&\quad +(1-\pi)\,\hat{\pi}\max\left\{0;(1+\hat{r}_L)\,\hat{L}-\hat{D}\right\}-C \\
&= \pi\hat{\pi}\,(1+r_L)\,L+\hat{\pi}\left(\frac{\hat{r}_L}{\kappa}L^*-\frac{\hat{r}_L+\kappa}{\kappa}D\right)-C.
\end{aligned}
\tag{36.44}
$$

This captures the profits of the four possible outcomes discussed above, where we recognise that if both the existing and the new loan fail, the bank will fail and hence no profits are generated. Using the first expression of the bank profits, these are maximized over the optimal size of the loan the bank will provide, giving us the first order condition

$$
\frac{\partial\hat{\hat{\Pi}}_B}{\partial\hat{L}} \geq \pi\hat{\pi}\hat{r}_L-\pi\,(1-\hat{\pi})=\pi\,(\hat{\pi}\,(1+\hat{r}_L)-1)>0,
\tag{36.45}
$$

where we have eliminated the positive final term and kept the negative second term for simplicity. This result shows that banks choose again the largest possible value for the new loans $\hat{L}$.

To accept the repurchase agreement the profits this generates must exceed that of not selling the loan, $\hat{\hat{\Pi}}_B \geq \Pi_B$, which solves for

$$
L^* \geq \hat{\hat{L}}^* = \frac{\kappa}{\hat{\pi}\hat{r}_L}C+\frac{\kappa\,(1-\pi)+\kappa^2\,(1-\hat{\pi})+\hat{\pi}\hat{r}_L}{\hat{\pi}\hat{r}_L\,(1+\kappa)}\pi\,(1+r_L)\,L.
\tag{36.46}
$$

Comparing equations (36.38) and (36.44), we can see that the sale of loans to the bad bank is preferred over a repurchase agreement if $\hat{\Pi}_B > \hat{\hat{\Pi}}_B$, which solves for $L^* > \pi\,(1+r_L)\,L$. Thus if the bad bank pays more than the intrinsic value, the bank prefers the sale as $L^* > D > \pi\,(1+r_L)\,L$, this is always the case. Hence banks would not engage in a repurchase agreement as the bad bank has to pay more than the intrinsic value of the bonds to ensure that the bank remain solvent and can repay its depositors.

Regulator preferences If the loans are sold, the costs to the regulator are given by

$$
\hat{\Pi}_R = \pi\,(1+r_L)\,L-\hat{L}^*,
\tag{36.47}
$$

consisting of the value of the loans less the purchase price. For a repurchase agreement, the losses are arising only if the bank is insolvent as then the price L^* is not returned.

$$\hat{\hat{\Pi}}_R = -\left(\pi\left(1-\hat{\pi}\right)+\left(1-\pi\right)\left(1-\hat{\pi}\right)\right)\hat{\hat{L}}^* = -\left(1-\hat{\pi}\right)\hat{\hat{L}}^*. \tag{36.48}$$

The bank is insolvent if both loans fail or if the new loan is not repaid, in which case the bad bank will lose the amount it initially paid to the bank.

The regulator prefers an outright sale of the loans if this provides him with the lower losses $\hat{\Pi}_R \leq \hat{\hat{\Pi}}_R$. We can derive that

$$\frac{\partial\hat{\hat{\Pi}}_R}{\partial C} = -\frac{\left(1-\hat{\pi}\right)\kappa}{\hat{\pi}\hat{r}_L} > -\frac{\kappa}{\hat{\pi}\left(1+\hat{r}_L\right)} = \frac{\partial\hat{\Pi}_R}{\partial C}, \tag{36.49}$$

if $\hat{\pi} > \frac{\kappa}{\hat{r}_L\kappa}$, which is a reasonable assumption. At $C = 0$ we can see that for reasonable values of the parameters, we have $\hat{\hat{\Pi}}_R < \hat{\Pi}_R$. Hence as the profits are decreasing faster in $\hat{\Pi}_R$ than in $\hat{\hat{\Pi}}_R$, we will find a C^* such that for $C \leq C^*$ the regulator prefers an outright sale of the loans and for $C > C^*$ the repurchase agreement is preferred. To see this inference note that $\hat{\Pi}_R$ and $\hat{\hat{\Pi}}_R$ are negative. Thus if the cost of selling loans to a bad bank are low, these should be sold outright and otherwise a repurchase agreement is preferred by the regulator.

Summary We have seen that bad banks need to purchase loans above their fair value to ensure the bank stays solvent and can repay its depositors. It is for this reason attractive to banks, provided they are receiving a payment high enough to be better off by providing new loans than holding on to their existing loans and receiving compensation for any additional costs, such the loss of reputation, associated with the sale to a bad bank. As banks will have to receive a payment in excess of the fair value of the loans, they would prefer an outright sale over a repurchase agreement. However, the losses to regulators may be higher from an outright sale. This will depend on the size of these additional costs that banks are facing. If these costs are small, then the bad bank does not need to pay a price too much in excess of the fair value of the loans and their losses are smaller in an outright sale compared to a repurchase agreement. In a repurchase agreement, the bad bank needs to take into account a complete loss of the payment made to the bank if the bank becomes insolvent after investing the newly obtained funds into new loans and these loans defaulting; in this case the repurchase agreement cannot be honoured by the bank and the regulator will have a large loss. Only once the additional costs to banks are high and hence the bad bank needs to provide a substantial premium for the loan they purchase, will a repurchase agreement become more attractive as there any initial losses are recovered through the repurchase; the possible loss due to the insolvency of the bank will be smaller than the losses from paying a higher premium for the loans due to the higher costs of banks.

Reading Hauck, Neyer, & Vieten (2015)

36.4 Regulatory structure

Regulators make decisions on when a bank should be supported or closed based on their assessment of the costs involved if closing the banks compared to not closing the bank. The calculation of such costs will, however, depend on the type of regulator. There are in essence two possible regulators, one being the central bank who might provide support to banks through loans and will be concerned about the repayment of these loans, the other is a deposit insurance who would mostly be concerned be concerned about the payment that has to be made to depositors if the bank fails. We will here determine which of these regulators would provide support to a bank and which is the optimal choice from the perspective of social welfare.

We consider the optimal decision maker to support or close a bank facing a liquidity shortage. If a bank faces a liquidity shortage and is not supported, it will be closed down and this imposes costs of c, proportional to the size of deposits. If a bank is closed, its assets, the loans L, are sold and a fraction λ of its value is realized.

Social optimum As the bank is only facing a liquidity shortage, but is not insolvent due to loans defaulting, there is no impact of this liquidity shortage on social welfare. The bank might, however, fail if the loan it has given is not repaid. Loans are repaid with probability π, including interest r_L, and these proceeds used to repay deposits with their interest r_D. If the loan is not repaid, then the bank needs to be closed, incurring costs c. Thus if the bank is allowed to continue operating the welfare is given by

$$\Pi_W = \pi \left(1 + r_L\right) L - (1 - \pi)\, cD - \left(1 + r_D\right) D, \tag{36.50}$$

If, on the other hand, the bank is closed down, the welfare consists of the proceeds of the loan sale less the costs of closing down the bank and repaying deposits, giving us

$$\hat{\Pi}_W = \lambda L - cD - \left(1 + r_D\right) D. \tag{36.51}$$

If loans are fully funded by deposits such that $D = L$, the bank allowed to continue operating if the welfare of doing so is higher than the welfare if closing down the bank, $\Pi_B > \hat{\Pi}_B$. This condition solves for

$$\pi > \pi^* = \frac{\lambda}{1 + r_L + c}. \tag{36.52}$$

Thus if the repayment rate for the loans of the bank is sufficiently high, it should not be closed instantly. We can now compare this result with the decision a regulator would make.

Central bank Let us assume the central bank provides a loan to cover the liquidity shock which is a fraction ν of the deposits; thus the loan would be νD. If the bank loan is repaid, the bank will be able to repay the central bank loan, but otherwise it is lost to the central bank and the bank is closed. Hence the outcome of the central bank is given by

$$\Pi_{CB} = -(1-\pi)(\nu D + cD). \tag{36.53}$$

If the central bank does not provide the loan, the bank would need to be closed instantly. We assume that as the regulator the central bank again has to bear the costs of closing the bank. Thus the outcome in this case is

$$\hat{\Pi}_{CB} = -cD, \tag{36.54}$$

assuming closing costs are borne by the regulator. The central bank provides the loan if this generates the higher outcome, $\Pi_{CB} > \hat{\Pi}_{CB}$, which solves for

$$\pi > \pi^{**} = \frac{\nu}{\nu + c} \tag{36.55}$$

We see that the threshold in the repayment rate is higher for the central bank, $\pi^{**} > \pi^{*}$, if

$$\nu > \nu^{*} = \frac{\lambda c}{1 + r_L + c - \lambda}. \tag{36.56}$$

It is thus that if the liquidity shock and thus loan required is sufficiently large the central bank will be more restrictive than the social optimum as the losses of the loan outweigh the costs of closing down the bank. For small liquidity shocks and thus small loans, the concerns over liquidation costs lead to the bank being closed less often than is optimal.

Deposit insurance Alternatively, the regulator of bank closure could be the deposit insurer. If the bank fails, the loan the deposit insurance provides to the bank is not repaid and the deposit insurance has to pay for the remaining deposits $D - \nu D$ and will lose its loan, νD, such that the total outcome for the deposit insurance is a loss of size of the deposits D and the costs of closing the bank down. Hence when providing the loan the outcome for the deposit insurance is

$$\Pi_{DI} = -(1-\pi)(L + cD). \tag{36.57}$$

When not providing the loan and closing the bank, they obtain the revenue from selling the loans λL, have to pay out the depositors and face the closure costs. Thus

$$\hat{\Pi}_{DI} = \lambda L - D - cD. \tag{36.58}$$

The loan is provided by the deposit insurance if this generates the higher outcome, $\Pi_{DI} > \hat{\Pi}_{DI}$, for which we need

$$\pi > \pi^{***} = \frac{\lambda}{1 + c} > \pi^{*}. \tag{36.59}$$

We observe that the central bank is more restrictive in providing a loan to the bank, $\pi^{**} > \pi^{***}$ if

$$\nu > \nu^{**} = \frac{\lambda c}{1 + c - \lambda} > \nu^{*} \tag{36.60}$$

and hence for large liquidity shocks, central banks are less likely to provide a loan to prevent the bank failing from a liquidity shock than a deposit insurer.

Optimal regulator We see from figure 36.4 depicting the relationship between the size of the liquidity shock and the repayment rate of the bank loan that the optimal closure rate π^* is least deviated from if for $\nu^* < \nu < \nu^{**}$ the central bank makes the closure decision, while for $\nu > \nu^{***}$ it is the deposit insurance. In the case that $\nu < \nu^*$ the social costs of closing the bank too readily at π^{***} for deposit insurance or being too lenient at π^{**} with the central bank need to be balanced.

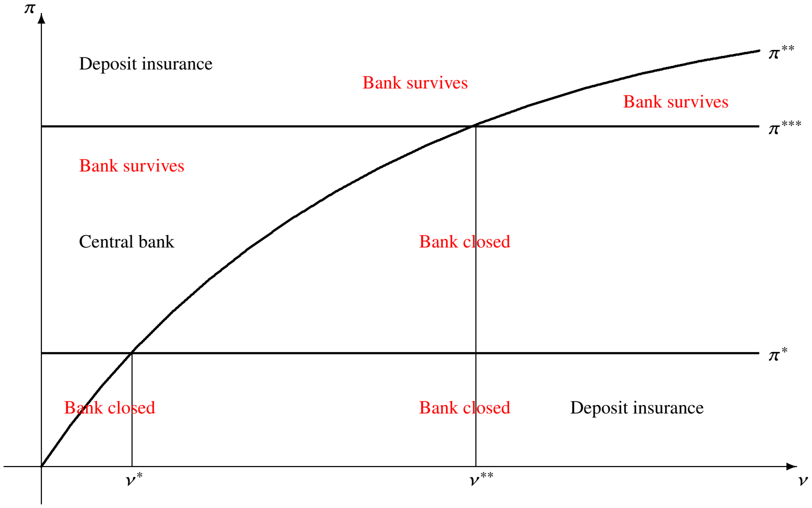


Fig. 36.4: Optimal regulator choice

Let us assume the repayment rate of loans, π, is known, but that the size of the liquidity shock and hence the loan size, ν, is not known when the decision is made who will act as the regulator and thus who has to provide the loan if the bank is not to be closed down. The social welfare of providing the loan is given by equation (36.50) and for closing the bank by equation (36.51). From equation (36.55) we obtain that the central bank closes the bank if $\pi < \frac{\nu}{\nu+c}$, or

$$\nu > \hat{\nu} = \frac{\pi c}{1-\pi}. \tag{36.61}$$

Hence define $p = Prob\,(\nu < \hat{\nu})$ as the probability that the liquidity shock is smaller than this threshold and hence the bank is not closed down, then the welfare from engaging the central bank as regulator is given by

$$\begin{aligned}\Pi_W^{CB} &= p\left(\pi\left(1+r_L\right)L-(1-\pi)\,cD-\left(1+r_D\right)D\right) \\ &\quad +(1-p)\left(\lambda L-cD-\left(1+r_D\right)D\right) \\ &= \left(p\pi\left(1+r_L+c\right)+(1-p)\,\lambda-c-\left(1+r_D\right)\right)L.\end{aligned} \tag{36.62}$$

For the deposit insurance we similarly get

$$\Pi_W^{DI} = \begin{cases} \left(\pi\left(1+r_L+c\right)-c-\left(1+r_D\right)\right)L & \text{if } \pi > \pi^{***} \\ \left(\lambda-c-\left(1+r_D\right)\right)L & \text{if } \pi \leq \pi^{***} \end{cases}. \tag{36.63}$$

Thus the central bank is preferred if the welfare it generates exceeds that of the deposit insurer, $\Pi_W^{CB} > \Pi_W^{DI}$, which requires

$$\begin{cases} \pi < \pi^* \text{ if } \pi > \pi^{***} \\ \pi \geq \pi^* \text{ if } \pi \leq \pi^{***} \end{cases}. \tag{36.64}$$

As $\pi < \pi^{***}$, the first condition cannot be met and for $\pi > \pi^{***}$ the deposit insurance is the preferred regulator. Using the second condition, we see that for $\pi^* < \pi < \pi^{***}$ the central bank is the preferred regulator and for $\pi < \pi^*$ it is again the deposit insurance.

If we optimally select the regulator, they will then make the decision about closing the bank that suits them best. As indicated in figure 36.4 this gives rise to banks always being closed down if their repayment rate is sufficiently low and if it is high, the bank will never be closed. In both of these cases the deposit insurance is the regulator. In an intermediate range, the central banks regulates the closure decision of banks and they will close a bank if the liquidity shock is too large.

Summary We can thus conclude that deposit insurance is best placed to decide the closure of banks in very safe and very risky banking systems, while the central bank is optimal for banking systems of medium risk. As a consequence, high risk banks are always closed and low risk banks can always obtain a loan to avoid failure due to a liquidity shock; banks with medium risk obtain such loans only for small liquidity shortages and are closed for larger liquidity shortages. It is therefore that the choice of regulator makes a difference when it comes to which banks are closed and the regulatory structure needs to be considered carefully for each banking system, depending on the risks that banks typically take.

Reading Repullo (2000)

Conclusions

While regulation seeks to prevent the failure of banks, it is only prudent to consider the possibility of banks failing despite any measures taken to prevent such an event.If the decision to close a bank is left to a regulator, then their own interests will dominate the decision and they will not necessarily make the decision that is socially optimal;

banks might be closed too readily or be kept operating for too long. Such distortions in the decision-making are the result of the interests of regulators seeking to minimize their own costs, but also the result of banks reacting strategically to the threat of closure by changing their decisions and affecting thereby the costs of regulators and hence their decisions. Any such distortions can be minimised by selecting the type regulator such that these are minimised.

Chapter 37

Regulatory coordination and competition

Banks are commonly not only regulated by a single set of rules, but are subjected to multiple regulators. This situation is common for banks operating in multiple jurisdictions where each jurisdiction has their own set of regulatory constraints. However, banks are not only subjected to different regulations, they can actively exploit any differences between regulations by moving their head office to another country or deciding to seek deposits in one country but provide loans in another country. This will induce competition between jurisdictions and we will use this chapter to evaluate the consequences of such competition.

We will investigate in chapter 37.1 how banks can exploit different regulatory restrictions and how this competition between regulations will affect the regulation itself. Banks being subjected to two regulators, one being a regulator in one jurisdiction only and the other covering multiple jurisdiction is the point of interest in chapter 37.2. Finally, in chapter 37.3 we consider a situation in which a single regulator covers multiple countries in a so-called banking union.

37.1 Regulatory arbitrage

If countries have a free-trade agreement, then banks will be able to be located in one country and operate in another country. Thus they can choose which country's regulatory regime they prefer. Exploiting any differences in regulation by choosing the location that is most beneficial to them is known as regulatory arbitrage. We will explore how competition between countries to attract banks will affect any regulation they are seeking to enact.

We assume there are two countries and in each of these countries two types of banks operate. A fraction p of banks provide loans L that are repaid with probability π_H, including interest r_L, and a fraction $1-p$ of banks whose loans are repaid with probability $\pi_L < \pi_H$. We further assume that $\pi_H(1+r_L) > \pi_L(1+r_L) > 1$, implying that both bank types are profitable. In addition, there is a probability γ that all banks are experiencing a macroeconomic shock. In this case, the loans of the banks lose value, for example due to an increase in the default rate, and are worth only a fraction $\lambda < 1$; if this happens we assume that the bank is liquidated. With a macroeconomic shock we find that $\lambda\pi_H(1 + r_L) > 1 > \lambda\pi_L(1 + r_L)$ and bank having taken the higher risks, π_L, would no longer be profitable.

A regulator imposes liquidity requirements R_i onto the banks, where $R_i = \rho_i D$, as well as a tax $T_i = \tau_i D$. The revenue of this tax is raised from banks and given to depositors. We can interpret it either as a bank tax whose revenue is returned through government spending or it represents wages of employees that are also depositors. liquidity requirements and taxes might be charged according to the risks banks are taking, thus they will depend son whether the repayment rates for loans are high or low. Bank profits are then given by

$$\begin{aligned} \Pi_B^i &= (1 - \gamma)(\pi_i(1 + r_L)L_i + R_i - (1 + r_D)D - T_i) \\ &= (1 - \gamma)(\pi_i(1 + r_L)(1 - \rho_i) + \rho_i - (1 + r_D) - \tau_i)D, \end{aligned} \tag{37.1}$$

with the loans and cash reserves jointly being financed by deposits such that $L_i = D - R_i = (1 - \rho_i)D$. The first term shows that profits are only realized if no macroeconomic shock is realized. The profits then consist of the return to the loans, the retained cash reserves, less the repayment of deposits and the tax imposed on the bank. As $\pi_i(1 + r_L) > 1$ it is obvious that banks would choose $R_i = 0$ and $T_i = 0$ to maximize their profits, hence any regulation that would be imposed is binding.

Bank disclosures We assume that the regulator cannot distinguish between the two types of banks, but rather has to rely on the declaration of banks of their type. In order for banks to disclose their type correctly, it must be less profitable to declare the other type and obtain liquidity requirements and taxation for that type. Hence

$$\begin{aligned} \Pi_B^{HH} &= (1 - \gamma)(\pi_H(1 + r_L)(1 - \rho_H) + \rho_H - (1 + r_D) - \tau_H)D \\ &\geq (1 - \gamma)(\pi_H(1 + r_L)(1 - \rho_L) + \rho_L - (1 + r_D) - \tau_L)D \\ &= \Pi_B^{HL}, \\ \Pi_B^{LL} &= (1 - \gamma)(\pi_L(1 + r_L)(1 - \rho_L) + \rho_L - (1 + r_D) - \tau_L)D \\ &\geq (1 - \gamma)(\pi_L(1 + r_L)(1 - \rho_H) + \rho_H - (1 + r_D) - \tau_H)D \\ &= \Pi_B^{LH}. \end{aligned} \tag{37.2}$$

The first (second) inequality states that it is preferable for a bank that has given loans with high (low) repayment rates to declare this fact rather than wrongly declare they have chosen loans with low (high) repayment rates. If the bank correctly declares the repayment rate, the profits are obtained from equation (37.1) and if the bank declares the wrong repayment rate, we use the true repayment rate, but the wrong liquidity requirements and tax rates.

These inequalities solve for

$$\begin{aligned} (\pi_H\,(1+r_L)-1)\,(\rho_L-\rho_H)-(\tau_H-\tau_L) &\geq 0, \\ (\pi_L\,(1+r_L)-1)\,(\rho_L-\rho_H)-(\tau_H-\tau_L) &\leq 0. \end{aligned} \tag{37.3}$$

We can rewrite equation (37.2) also as

$$\Pi_B^{HL} = \Pi_B^{LL} + (1-\gamma)\,(\pi_H-\pi_L)\,(1-\rho_L)\,D \leq \Pi_B^{HH}, \tag{37.4}$$

From this relationship we get that $\Pi_B^{HH} \geq \Pi_B^{LL} \geq 0$, the latter inequality we impose to ensure the economy consists of both banks. The regulator can extract all surplus from the bank that has provided loans with a low repayment rate such that $\Pi_B^{LL} = 0$. With this assumption we obtain

$$\Pi_B^{HH} = (1-\gamma)\,(\pi_H-\pi_L)\,(1-\rho_L)\,(1+r_L)\,D, \tag{37.5}$$

if the regulator extracts all additional surplus from bank the bank that provides low-risk loans but still allowing the constraint in equation (37.2) to be fulfilled, implying from equation (37.4) that $\Pi_B^{HH} \geq \Pi_B^{HL}$. From equation (37.2) we then obtain

$$(1-\gamma)\,(\pi_H-\pi_L)\,(\rho_H-\rho_L)\,(1+r_L)\,D \leq 0, \tag{37.6}$$

which requires

$$\rho_H \leq \rho_L. \tag{37.7}$$

The liquidity requirements on banks that provide loans with high repayment rates have to be less than those of banks providing loans with low repayment rates.

Using that $\Pi_B^{LL} = 0$ and equation (37.5) in the profits of the bank in equation (37.1), the tax rates are given by

$$\begin{aligned} \tau_L &= \pi_L\,(1+r_L)\,(1-\rho_L)+\rho_L-(1+r_D)\,, \\ \tau_H &= \pi_H\,(1+r_L)\,(\rho_L-\rho_H)+(\rho_H-\rho_L)+\tau_L \\ &= (\pi_H\,(1+r_L)-1)\,(1\rho_L-\rho_H) \\ &\quad +\pi_L\,(1+r_L)\,(1-\rho_L)+\rho_L-(1+r_D)\,. \end{aligned} \tag{37.8}$$

The differences in the tax rate are easily determined as

$$\tau_H-\tau_L = (\pi_H\,(1+r_L)-1)\,(\rho_L-\rho_H) \tag{37.9}$$

and as from equation (37.7) we know that $\rho_H \leq \rho_L$, we find that

$$\tau_H \geq \tau_L. \tag{37.10}$$

Thus the regulator would charge a higher tax rate to banks providing low-risk loans than those providing high-risk loans. This will at least partially offset the benefits that low-risk banks have from lower liquidity requirements.

Optimal regulation We assume that a regulator would seek to maximize the welfare of depositors by choosing the liquidity requirements and tax rates optimally. Depending on the type of bank, depositors obtain

$$\Pi_D^i = (1 + r_D + \tau_i)\, D \tag{37.11}$$

if the bank continues as they obtain their deposits and interest r_D on them, together with the tax revenue. If the bank is liquidated after a macroeconomic shock, the depositors obtain the proceeds of the liquidated loan and the cash reserves the bank has held, but there are no taxes to be collected. Hence depositors obtain

$$\hat{\Pi}_D^i = (\lambda\pi_i\,(1 + r_L)\,(1 - \rho_i) + \rho_i)\, D. \tag{37.12}$$

We easily obtain the following partial derivatives of these depositor outcomes:

$$\begin{aligned}
\frac{\partial \Pi_D^{HH}}{\partial \rho_H} &= (1 - \pi_H\,(1 + r_L))\, D, \\
\frac{\partial \Pi_D^{HH}}{\partial \rho_L} &= (\pi_H - \pi_L)\,(1 + r_L)\, D, \\
\frac{\partial \Pi_D^{LL}}{\partial \rho_H} &= 0, \\
\frac{\partial \Pi_D^{LL}}{\partial \rho_L} &= (1 - \pi_L\,(1 + r_L))\, D, \\
\frac{\partial \hat{\Pi}_D^{HH}}{\partial \rho_H} &= (1 - \lambda\pi_H\,(1 + r_L))\, D, \\
\frac{\partial \hat{\Pi}_D^{HH}}{\partial \rho_L} &= 0, \\
\frac{\partial \hat{\Pi}_D^{LL}}{\partial \rho_H} &= 0, \\
\frac{\partial \hat{\Pi}_D^{LL}}{\partial \rho_L} &= (1 - \lambda\pi_L\,(1 + r_L))\, D.
\end{aligned} \tag{37.13}$$

With $U(\cdot)$ denoting the utility of these payoffs, where $U'(\Pi_D) = \frac{\partial U(\Pi_D)}{\partial \Pi_D} > 0$ denotes the marginal utility and $\frac{\partial^2 U(\pi_D)}{\partial \Pi_D} < 0$, the expected utility of depositors is given by

$$U_D = p\left(\gamma U\left(\Pi_D^{HH}\right) + (1 - \gamma)\, Y\left(\hat{\Pi}_D^{HH}\right)\right) + (1 - p)\left(\gamma U\left(\Pi_D^{LL}\right) + (1 - \gamma)\, U\left(\hat{\Pi}_D^{LL}\right)\right), \tag{37.14}$$

where we look at cases for both types of banks, p, and the existence of macroeconomic shocks, γ. After inserting from equation (37.8) into equation (37.11), using the expressions from equation (37.13), we get the first order conditions of maximizing this expression for the optimal liquidity requirements for depositors at low-risk and high-risk banks, respectively, as

$$\frac{\partial U_D}{\partial \rho_H} = p\left(\gamma\left(1-\pi_H\left(1+r_L\right)\right)U'\left(\Pi_D^{HH}\right)\right. \tag{37.15}$$
$$\left.+\left(1-\gamma\right)\left(1-\lambda\pi_H\left(1+r_L\right)\right)U'\left(\hat{\Pi}_D^{HH}\right)\right)D=0$$
$$\frac{\partial U_D}{\partial \rho_L} = \left(p\gamma\left(\pi_H-\pi_L\right)\left(1+r_L\right)U'\left(\Pi_D^{HH}\right)\right.$$
$$+\left(1-p\right)\left(\gamma\left(1-\pi_L\left(1+r_L\right)\right)U'\left(\Pi_D^{LL}\right)\right.$$
$$\left.\left.+\left(1-\gamma\right)\left(1-\lambda\pi_L\left(1+r_L\right)\right)U'\left(\hat{\Pi}_D^{LL}\right)\right)\right)D.$$

These conditions can be rewritten as

$$\left(1-\gamma\right)\left(1-\lambda\pi_L\left(1+r_L\right)\right)U'\left(\hat{\Pi}_D^{LL}\right) \tag{37.16}$$
$$=\gamma\left(\pi_L\left(1=r_L\right)-1\right)U'\left(\Pi_D^{LL}\right)$$
$$-\frac{p}{1-p}\gamma\left(\pi_H-\pi_L\right)\left(1+r_L\right)U'\left(\Pi_D^{HH}\right)$$
$$=\gamma\pi_L\left(1+r_L\right)\left(U'\left(\Pi_D^{LL}\right)-U'\left(\Pi_D^{HH}\right)\right)$$
$$+\gamma\pi_H\left(1+r_L\right)U'\left(\Pi_D^{HH}\right)$$
$$-\gamma U'\left(\Pi_D^{LL}\right)-\frac{1}{1-p}\gamma\left(\pi_H-\pi_L\right)\left(1+r_L\right)U'\left(\Pi_D^{HH}\right)$$
$$=\gamma\left(\pi_L\left(1+r_L\right)-1\right)\left(U'\left(\Pi_D^{LL}\right)-\left(\Pi_D^{HH}\right)\right)$$
$$+\left(1-\gamma\right)\left(1-\lambda\pi_H\left(1+r_L\right)\right)U'\left(\hat{\Pi}_D^{HH}\right)$$
$$-\frac{1}{1-p}\gamma\left(\pi_H-\pi_L\right)\left(1+r_L\right)U'\left(\Pi_D^{HH}\right).$$

To obtain these results, we have inserted for the expression $\gamma\left(1-\pi_H\left(1+r_L\right)\right)U'\left(\Pi_D^{HH}\right)$ from the first line in equation (37.15). We see that the second term on the right hand side is negative as we assumed $\lambda\pi_H\left(1+r_L\right)>1$ and the final term is negative as well. As $\lambda\pi_L\left(1+r_L\right)<1$ by assumption, the left hand side is positive, implying that the first term on the right hand side must be positive, thus $U'\left(\Pi_D^{LL}\right)>U'\left(\Pi_D^{HH}\right)$ or $\Pi_D^{LL}<\Pi_D^{HH}$. Inserting from equation (37.11) and using equation (37.8), this becomes $\left(\pi_H\left(1+r_L\right)-1\right)\left(\rho_L-\rho_H\right)>0$ or $\rho_L>\rho_H$.

We therefore see that the optimal liquidity requirements are consistent with the required incentives for banks to disclose the risks of the loans they have provided, which requires that less cash reserves are held by the less risky bank; this bank is then taxed more than the more risky bank. The higher liquidity requirements in this situation allows for a higher payment to depositors in case of a macroeconomic shock.

We now have established properties of the liquidity requirements and the taxation in a single country. We will now continue to assess the impact of two countries competing for these banks.

Competition between countries Let us continue by assuming that there are two countries with an identical composition of banks; these banks are free to relocate between these countries. The tax is always payable in the country the bank is resident in and whose regulations have to be followed. Thus a country with two banks has tax revenue $2\tau_i$ and a country with no bank has tax revenue $\tau_i = 0$.

Suppose $\tau_L > 0$ in one of the two countries. The other country could lower their tax rate to $\hat{\tau}_L = \tau_L - \varepsilon$ and, using the same liquidity requirements ρ_L, attract both banks. Hence we get revenue for depositors in this country of

$$\Pi_D^{LL} = (1 + r_D + 2\hat{\tau}_L)\, D, \tag{37.17}$$

Hence from equation (37.14) the expected utility increases as long as $2\hat{\tau}_L > \tau_L$, or $\tau_L > 2\varepsilon$. Therefore as long as $\tau_L > 0$ the other country will be able to generate a higher utility to its depositors by undercutting the tax rate, leading to the only solution that $\pi_L = 0$. The same argument applies for $\tau_H > 0$. Therefore, tax rates cannot be positive. Subsidies, such as setting $\tau_i < 0$ is also not possible as setting $\hat{\tau}_i = \tau_i + \varepsilon$ in this case has the same impact. It therefore follows that with two countries competing for the location of the bank we obtain

$$\tau_L = \tau_H = 0. \tag{37.18}$$

Inserting this result into equation (37.2) we find that in order to ensure banks disclose the risks of the loans they provide correctly, we require

$$(\pi_H (1 + r_L) - 1)(\rho_L - \rho_H) \leq 0 \leq (\pi_L (1 + r_L) - 1)(\rho_L - \rho_H)\,. \tag{37.19}$$

Using our assumption that $\pi_H (1 + r_L) > 1 > \pi_L (1 + r_l)$ this simplifies to $\rho_L - \rho_H \geq 0 \geq \rho_L - \rho_H$, which can only be fulfilled if

$$\rho_L = \rho_H. \tag{37.20}$$

We find that competition between countries reduces the tax burden as the countries compete in attracting banks until the tax is completely eliminated. As depositors benefit from the tax raised directly, they will undercut each other to attract both banks until taxes are eliminated. A subsidy is not sustainable either as the countries would make a loss from attracting banks and thus reduce them to entice banks away from their territory. Consequently, tax rates are zero.

Having the same zero tax rate for banks of both types eliminates the possibility of a trade-off between a higher tax rate for low-risk banks in exchange for lower liquidity requirements. Hence banks cannot be provided with any incentives to disclose their type truthfully, necessitating to have uniform liquidity requirements across banks. Therefore, competition does not allow to differentiate between banks with different risks. Of course, the liquidity requirements in both countries need to be identical as otherwise banks would be able to move to the country that offers the lower liquidity rate.

From equation (37.1) it is easy to see that $\Pi_B^{HH} > \Pi_B^{LL} \geq 0$. With $\tau_i = 0$ we get using equations (37.11) and (37.12) that depositors obtain

$$\begin{aligned} \Pi_D^i &= (1 + r_D)\, D, \\ \hat{\Pi}_D^i &= (\lambda \pi_i (1 + r_L)(1 - \rho) + \rho)\, D, \end{aligned} \tag{37.21}$$

for the case the loans are repaid and the loans default, respectively. This implies from equation (37.14) that the objective function of depositors is given by

$$U_D^* = \gamma U \left(\Pi_D^i\right)(1 - \gamma)\left(pU\left(\hat{\Pi}_D^{HH}\right) + (1 - p)\, U\left(\hat{\Pi}_D^{LL}\right)\right), \tag{37.22}$$

from which we get the first order conditions as

$$\begin{aligned} \frac{\partial U_D^*}{\partial \rho} = (1 - \gamma) \Big(p\, (1 - \lambda \pi_H (1 + r_L))\, U'\left(\hat{\Pi}_D^{HH}\right) \\ + (1 - p)(1 - \lambda \pi_L (1 + r_L))\, U'\left(\hat{\Pi}_D^{LL}\right)\Big)\, D = 0 \end{aligned} \tag{37.23}$$

which can be solved for the optimal liquidity requirements.

An upper bound of liquidity requirements are given by the participation constraint $\Pi_B^{LL} \geq 0$. This requires

$$\rho \leq \rho^* = \frac{(1 + r_D) - \pi_L (1 + r_L)}{1 - \pi_L (1 + r_L)}, \tag{37.24}$$

noting that $\pi_L (1 + r_L) > 1 + r_D > 1$.

Summary We have thus seen that banks will seek to exploit differences in regulatory regimes between countries, leading to countries competing for banks, for example to raise additional taxes. This leads to an erosion of taxes in both countries. A consequence of the competition between countries for banks is that the distinction between banks with different risks is no longer possible as banks have no incentives to disclose these risks truthfully. This leads to common liquidity requirements for all banks, regardless of the risks they are taking. If there was no competition between banks, then regulators could use a combination of different tax rates and liquidity requirements to entice banks to disclose their risks truthfully. Competition between regulations and banks exploiting this competition in what is known as regulatory arbitrage, will not allow regulators to design incentives to this effect.

Reading Boyer & Kempf (2020)

37.2 Local and global regulators

Banks are subjected to multiple regulators, often a local regulator that sets restrictions for banks for the benefit of the local economy and then another regulator that takes a more global view by taking into account the effect of banks on the wider economy as a whole. We will here investigate which arrangement is preferable.

Depositors at a bank can be domestic depositors, D_H, or foreign depositors D_F, where the total deposits are given by $D = D_H + D_F$; all loans the bank provides are domestic and fully financed by deposits such that $L = D$. Using these loans, companies invest into projects that yield a return of R with probability π and zero otherwise. Hence the total social surplus of successful investment is

$$S_G = (1 + R) L - D = RD. \tag{37.25}$$

The global surplus consists of the return on the investment, less the financing costs of these investments, which are the deposits. We here assume that deposits do not attract interest. The surplus accruing domestically, disregards the effect on foreign depositors such that

$$S_H = (1 + R) - D_H = RD_H + (1 + R) D_F. \tag{37.26}$$

If a regulator decides to close down a bank, the loans this bank has provided are liquidated raising only a fraction $\lambda < 1$ to distribute equally between all depositors. The social surplus is only obtained if the loans are repaid, π, and this expected surplus needs to be compared to the payments the depositors obtain, λD globally and λD_H for a regulator that is only concerned about their domestic depositors. Hence a regulator concerned about the global and domestic welfare, respetively, will close the bank if

$$\begin{aligned} \pi S_G &\leq \lambda L, \\ \pi S_H &\leq \lambda D_H. \end{aligned} \tag{37.27}$$

This condition can be solved for

$$\begin{aligned} \pi \leq \pi_G^* &= \frac{\lambda}{R}, \\ \pi \leq \pi_H^* &= \frac{\lambda D_H}{RD_H + (1 + R) D_F}. \end{aligned} \tag{37.28}$$

We easily see that $\pi_H^* < \pi_G^*$ and hence the local regulator will be less strict in its closure decision. This result arises from the externality imposed on foreign depositors whose welfare is ignored.

As the bank will be liquidated if the repayment rate of loans, π, is below its critical level π_i^* as determined in equation (37.28), and assuming that the default rates of loans are not known ex-ante, but only their distribution $F(\cdot)$, we have the bank profits given by

$$\hat{\Pi}_B\left(\pi_i^*\right) = \int_{\pi_i^*}^{1} \pi\left(\left(1+r_L\right)L - \left(1+r_D\right)D\right)dF\left(\pi\right) \tag{37.29}$$
$$= \left(r_L - r_D\right)L\int_{\pi_i^*}^{1}\pi dF\left(\pi\right).$$

Let us now assume that the bank can make it more difficult for the regulator to assess the risks they have taken, thus to know the repayment rate π. Suppose that with probability p the regulator cannot assess π and consequently cannot close the bank. The costs of banks to increase the difficulty to regulators is increasing in the amount of loans and the probability p such that $C = \frac{1}{2}cp^2L$. The profits of the bank are then given by

$$\Pi_B = p\hat{\Pi}\left(0\right) + \gamma\left(1-p\right)\hat{\Pi}_B\left(\pi_G^*\right) + \left(1-\gamma\right)\left(1-p\right)\hat{\Pi}_B\left(\pi_H^*\right) - C. \tag{37.30}$$

The first term denotes the case where the regulator cannot identify the risk π, and the threshold for closing the bank becomes $\pi_i^* = 0$ and the bank is never liquidated. The second and third term, respectively, denote the profits made when assessed by a global regulator, with probability γ, and a local regulator with probability $1-\gamma$, and the regulator obtaining information about the risk π. We thus assume here that the bank is subject to two possible regulators, either a local regulator or a global regulator.

The optimal level of supervisory difficulty, p, is then obtained by maximizing the profits of the banks, giving us the first order condition

$$\frac{\partial\Pi_B}{\partial p} = \hat{\Pi}\left(0\right) - \gamma\hat{\Pi}_B\left(\pi_G^*\right) - \left(1-\gamma\right)\hat{\Pi}_B\left.\right)\pi_H^*\right) - cLp = 0,$$

which solves for the optimal level of difficulty

$$p^* = \frac{\hat{\Pi}_B\left(0\right) - \gamma\hat{\Pi}\left(\pi_G^*\right) - \left(1-\gamma\right)\hat{\Pi}_B\left(\pi_H^*\right)}{cL}. \tag{37.31}$$

We can now determine whether a global or a local regulator should supervise banks. We assume that this is done such that global welfare is maximized. The global welfare is given by

$$\begin{aligned}\Pi_W = {} & p^* \int_0^1 \pi RLdF(\pi) + \\ & (1-p^*)\gamma \left(\int_0^{\pi_G^*} \lambda LdF(\pi) + \int_{\pi_G^*}^1 \pi RLdF(\pi) \right) \\ & + (1-p^*)(1-\gamma) \left(\int_0^{\pi_H^*} \lambda LdF(\pi) + \int_{\pi_H^*}^1 \pi RLdF(\pi) \right) \\ = {} & \left(\int_0^1 \pi dF(\pi) + (1-p^*) \left(\gamma \int_0^{\pi_G^*} \left(\pi_G^* - \pi\right) dF(\pi) \right.\right. \\ & \left.\left. + (1-\gamma) \int_0^{\pi_H^*} \left(\pi_G^* - \pi\right) dF(\pi) \right) \right) RL,\end{aligned} \tag{37.32}$$

using from equation (37.28) that $\lambda = \Pi_G^* R$ and merging integrals. The first term represents the case in which the regulators have no information on the bank and then the subsequent terms are for global and local regulators, using our knowledge that banks with $\pi < \pi_i^*$, are liquidated. The optimal fraction of global regulation is then obtained by maximizing this expressions, giving us the first order condition

$$\begin{aligned}\frac{\partial \Pi_R}{\partial \gamma} = {} & \left((1-p^*) \int_{\pi_H^*}^{\pi_G^*} \left(\pi_G^* - \pi\right) dF(\pi) - \frac{\partial p^*}{\partial \gamma} \gamma \int_{\pi_H^*}^{\pi_G^*} \left(\pi_G^* - \pi\right) dF(\pi) \right. \\ & \left. - \frac{\partial p^*}{\partial \gamma} \int_0^{\pi_H^*} \left(\pi_G^* - \pi\right) dF(\pi) \right) RL = 0,\end{aligned}$$

where we have merged integrals for convenient. Using that from equation (37.31) we can obtain

$$\begin{aligned}\frac{\partial p^*}{\partial \gamma} &= \frac{\hat{\Pi}\left(\pi_H^*\right) - \hat{\Pi}\left(\pi_G^*\right)}{cL} = \frac{r_L - r_D}{c} \int_{\pi_H^*}^{\pi_G^*} \pi dF(\pi) > 0, \\ p^* &= \frac{\hat{\Pi}(0) - \hat{\Pi}\left(\pi_H^*\right)}{cL} + \gamma \frac{\partial p^*}{\partial \gamma},\end{aligned} \tag{37.33}$$

we get the optimal fraction of global regulation to be

$$\gamma = \frac{\int_{\pi_H^*}^{\pi_G^*} \left(\pi_G^* - \pi\right) dF(\pi) \left(1 - \frac{\partial p^*}{\partial \gamma}\right) - \frac{\partial p^*}{\partial \gamma} \int_0^{\pi_H^*} \left(\pi_G^* - \pi\right) dF(\pi)}{2 \frac{\partial p^*}{\partial \gamma} \int_{\pi_H^*}^{\pi_G^*} \left(\pi_G^* - \pi\right) dF(\pi)}. \tag{37.34}$$

Hence it is optimal to have a combination of local and global regulators as in general we will have $0 \leq \gamma \leq 1$.

We find that for the overall global welfare it is best to have a combination of local and global regulators. Having only a local regulator would lead to a situation in which banks are not closed down fast enough as regulators are not concerned about the welfare impact a failing bank has on foreign depositors. On the other hand, having only a global regulator, the welfare of domestic depositors would not be taken into account fully as the concerns of foreign depositors are also considered. It is thus a combination of these two types of regulators that is optima. The arrangement of having regulations that take into account local concerns on the one hand, but also global concerns on the other hand are optimal.

Reading Colliard (2020)

37.3 Banking unions

In a banking union two or more countries decide on a joint regulation of all banks in their respective countries and this regulation will take into account the interests of all countries to the same degree. We will therefore determine here under which condition a banking union is preferable to each country conducting their own regulation.

We consider banks in two countries. If banks fail in country i, a fraction γ_i of banks are bailed out by a regulator. The costs of bailing out these banks are increasing in the fraction of banks bailed out and the size of the bank such that we have $c\gamma_i^2 D$. The amount bailed out is $\gamma_i L$ as each bank gives loans to the amount of L. The deposits a country provides, D, are held in this country as well as the other country in fractions α and $1-\alpha$, respectively.

Individual bailouts Let us initially consider a regulator conducting a bailout who is only concerned about the costs the bailout imposes on this country. The costs of bailing out the banks in a country consist of the costs of the bailout itself, $c\gamma_i^2 D$, and the losses incurred by the domestic depositors in the not bailed out banks of that country, $\alpha(1-\gamma_i)D$. Thus total costs of a bailout in a country is given by

$$C_i = c\gamma_i^2 D + \alpha(1-\gamma_i)D. \tag{37.35}$$

The optimal level of a bailout is then given by minimising these costs, giving us the first order condition

$$\frac{\partial C_i}{\partial \gamma_i} = 2c\gamma_i L - \alpha D = 0, \tag{37.36}$$

from which we obtain the optimal fraction of banks to be bailed out as

$$\gamma_i = \frac{\alpha}{2c}. \tag{37.37}$$

Inserting this result back into equation (37.35), the costs of a bailout are given by

$$C_i = \alpha\left(1-\frac{\alpha}{4c}\right)D. \tag{37.38}$$

By not bailing out all banks, the country imposes costs on the other country as not all their depositors are benefitting from the bailout and will hence make losses. These losses imposed on depositors of other countries are given by

$$C_{ij} = (1-\alpha)(1-\gamma_i) D = (1-\alpha)\left(1-\frac{\alpha}{2c}\right) D. \tag{37.39}$$

The costs are imposed on non-domestic deposits, $1-\alpha$, that are not bailed out, $1-\gamma$. The second equality arises if we insert for the optimal bailout from equation (37.37).

If we assume that both countries are equal in that they face the same bailout costs and deposits are of the same size, the total costs of this bailout strategy for each country would be

$$C = C_i + C_{ji} = \left(1 - \frac{\alpha(2-\alpha)}{4c}\right) D. \tag{37.40}$$

These costs of a country optimizing their own bailout can now be compared to a situation of a banking union where the overall costs of a bailout are considered.

Bailouts in a banking union In a banking union, in contrast, the total costs would be minimized. With deposits on the two countries together being $2D$, we have these costs given as

$$\hat{C} = 2c\hat{\gamma}^2 D + 2(1-\hat{\gamma}) D, \tag{37.41}$$

in analogy to the costs in equation (37.35). Minimizing these costs yields the first order condition

$$\frac{\partial \hat{C}}{\partial \hat{\gamma}} = 4\hat{\gamma} c D - 2D = 0, \tag{37.42}$$

from which we obtain

$$\hat{\gamma} = \frac{1}{2c}. \tag{37.43}$$

Inserting this result into equation (37.41), we obtain the total costs as

$$\hat{C} = 2\left(1 - \frac{1}{4c}\right) D. \tag{37.44}$$

As the two countries are identical, they would share these costs equally such that the costs for each country are

$$\hat{C}_i = \frac{1}{2}\hat{C} = \left(1 - \frac{1}{4c}\right) L. \tag{37.45}$$

Risk-taking by banks The potential bailouts will affect the risk-taking of banks. Assume that banks can adjust their risk-level from a base repayment rate of π, which represents the probability of the loan being repaid, to

$$\hat{\pi}_i = (1-\theta_i)\pi. \tag{37.46}$$

This increased risk causes the loan rate, r_L, to increase such that

$$1 + \hat{r}_L^i = (1+\theta_i)(1+r_L). \tag{37.47}$$

If banks are bailed out after they fail, we assume that the bailout is such that their loans are repaid without interest. This gives us bank profits of

$$\begin{aligned}\Pi_B &= \hat{\pi}_i \left(\left(1+\hat{r}_L^i\right) L - \left(1+r_D\right) D\right) + \left(1-\hat{\pi}_i\right)\gamma_i L \\ &= \left(\left(1-\theta_i^2\right)\pi\left(1+r_L\right) - 1 + \gamma_i\left(1-\pi\left(1-\theta_i\right)\right)\right) L,\end{aligned} \tag{37.48}$$

where for the second equality we used that loans are fully financed by deposits, $L = D$ and inserted from equations (37.46) and (37.47). We also assumed that $\hat{\pi}_i\left(1+r_D\right) = 1$ to make deposit markets competitive. The first term denotes the profits of the bank if the loan is repaid and the second term the payment of the defaulting loan, provided the bank is bailed out. Note that the expected return on loans, $\left(1-\theta_i^2\right)\pi\left(1+r_L\right)$ is decreasing as the risk is increasing, that is with θ_i increasing.

Maximizing these bank profits over the risks the banks seek to take, we get the first order condition as

$$\frac{\partial \Pi_B}{\partial \theta_i} = \left(-2\theta_i \pi\left(1+r_L\right) + \pi\gamma_i\right) L = 0, \tag{37.49}$$

which solves for the optimal risk level being

$$\theta_i = \frac{\gamma_i}{2\left(1+r_L\right)}. \tag{37.50}$$

The optimal risk will depend on the bailout policy. The expected bailout costs for individual countries setting the bailout policy are now given by $\left(1-\hat{\pi}_i\right) C$, where $\hat{\pi}_i = \pi\left(1-\frac{\gamma_i}{2(1+r_L)}\right) = \pi\left(1-\frac{\alpha}{4c(1+r_L)}\right)$ when inserting all the relevant expressions. The total costs C are only incurred if the bank fails, which happens with probability $\hat{\pi}_i$. Similarly, in the banking union these costs are then $\left(1-\hat{\pi}\right)\hat{C}_i$, where $\hat{\pi} = \pi\left(1-\frac{\hat{\gamma}}{2(1+r_L)}\right) = \pi\left(1-\frac{1}{4c(1+r_L)}\right)$. The banking union is now preferred if the costs are smaller than the costs of each country making their own bailout decisions. Thus we require

$$\begin{aligned}&\left(1-\pi\left(1-\frac{1}{4c\left(1+r_L\right)}\right)\right)\left(1-\frac{1}{4c}\right) \\ &\quad\leq \left(1-\pi\left(1-\frac{\alpha}{4c\left(1+r_L\right)}\right)\right)\left(1-\frac{\alpha\left(2-\alpha\right)}{4c}\right)\end{aligned} \tag{37.51}$$

This expression can be solved for

$$\begin{aligned}&\alpha^2\frac{\pi}{4c\left(1+r_L\right)} + \alpha\left(1-\pi\left(1+\frac{1}{4c\left(1+r_L\right)}\right)\right) \\ &\quad-\left(\left(1-\pi\right) - \frac{\pi}{1+r_L}\left(1-\frac{1}{4c}\right)\right) \leq 0.\end{aligned} \tag{37.52}$$

The critical value for the size of the country α can be obtained by setting this relationship equal. Denoting the left-hand side by Ψ, we have

$$\begin{aligned} \frac{\partial \Psi}{\partial \alpha} &= 2\alpha \frac{\pi}{4c\left(1+r_L\right)} + \left(1 - \pi\left(1 + \frac{1}{4c\left(1+r_L\right)}\right)\right) \geqslant 0, \qquad (37.53)\\ \frac{\partial \Psi}{\partial\left(1+r_L\right)} &= -\frac{\alpha^2 \pi}{4c\left(1+r_L\right)^2} + \frac{\alpha\pi}{4c\left(1+r_L\right)^2} - \frac{\pi}{\left(1+r_L\right)^2}\left(1 - \frac{1}{4c}\right) < 0,\\ \frac{\partial \Psi}{\partial c} &= \left(1 + \alpha - \alpha^2\right)\frac{\pi}{4c^2\left(1+r_L\right)} > 0. \end{aligned}$$

The second equation is negative as $c > \frac{1}{3}$ given we assumed that c are the costs of a bail out, for example their financing costs, such that we can suppose that $c > 1$. The first equation is positive if

$$\alpha > \alpha^* = \frac{1}{2} - 2c\left(1+r_L\right)\frac{1-\pi}{\pi}. \qquad (37.54)$$

Using the implicit function theorem we then have

$$\begin{aligned} \frac{\partial \alpha}{\partial\left(1+r_l\right)} &= \begin{cases} -\frac{\frac{\partial \Psi}{\partial(1+r_L)}}{\frac{\partial \Psi}{\partial \alpha}} > 0 \text{ if } \alpha > \alpha^* \\ -\frac{\frac{\partial \Psi}{\partial(1+r_L)}}{\frac{\partial \Psi}{\partial \alpha}} \leq 0 \text{ if } \alpha \leq \alpha^* \end{cases}, \qquad (37.55)\\ \frac{\partial \alpha}{\partial c} &= -\frac{\frac{\partial \Psi}{\partial c}}{\frac{\partial \Psi}{\partial \alpha}} > 0. \end{aligned}$$

Figure 37.1 illustrates these results. We see that a banking union is preferred for banks with high loan rates, and thus high profitability. As a banking union leads to a higher bailouts as we can see when comparing equations (37.37) and (37.43), this will result in a higher risk shifting due to equation (37.50). The higher loan rates limit this risk shifting, making the benefits of the banking union in terms of higher bailouts and less losses to depositors more important. These benefits arise from the absence of the externalities imposed on foreign depositors that are bailed out less frequently than is optimal. the more deposits are held domestically, a higher α, the less important this externality becomes and the banking union will be less attractive. For cases that most deposits are held overseas, a low α, a banking union might not be attractive as the depositors are mostly invested overseas and thus similar to domestic depositors there will mostly be bailed out by the other country; some externality remains, though, causing the asymmetry we observe in figure 37.1.

Summary We have thus seen that a banking union can address the externality of a national regulator not considering the welfare of depositors from other countries in their decision-making and hence bailing banks out not frequently enough. In a banking union all depositors are considered equally, but the more frequent bailout also increase the moral hazard of banks increasing the risks they are taking. This

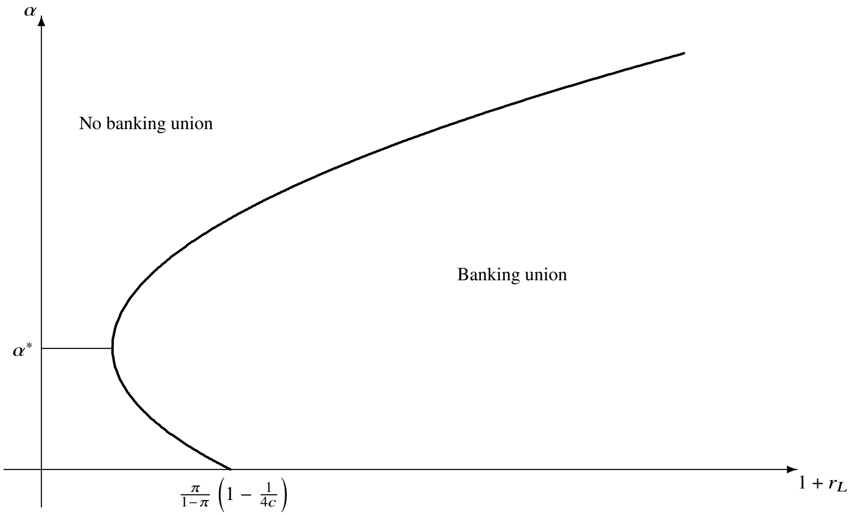


Fig. 37.1: Optimal banking unions

offsets some of the benefits a banking union has and thus it will not be optimal in all circumstances. If foreign deposits are rare, the externality of lower bailouts are negligible and banking unions offer very few benefits. The same is the case if foreign deposits are very high, in this case the foreign deposits are dominating the market and in a reversal, of the previous argument, their interest will be taken into account indirectly through the externality from the other country. It is thus in particular that if cross-border deposits are substantial but not dominating that banking unions are most beneficial.

Reading Haufler (2021)

Conclusions

We have seen that regulation across countries should be coordinated in many cases. We have seen that internationally active banks can exploit differences in regulations and through regulatory arbitrage induce countries to compete with each other through ever more lenient regulations. This might make existing regulations in the end not only less strict than they should be, but can also make them less effective in inducing specific behaviour in banks. Such coordination of regulation can be a two-tier process where local, or national, regulators set rules, but these are complemented by global rules. This would allow a balance between local interests and global interest, minimising any externalities. It can even be optimal to have only a single global regulator, a banking union, to eliminate any externalities arising from the regulation in one country onto banks and individuals in another country.

Review

The regulation of banks is not very straightforward as not only are there many different aspects to consider, but regulatory constraints will also affect the decision-making by banks, making regulation less effective or even having the opposite of the desired effect. In addition, the different regulations banks are subjected to need to be coordinated to ensure the are not offsetting each other. The most common regulation is to impose minimum capital requirements on banks, which has the aim to reduce the risks of banks by providing incentives to reduce risks, in addition to providing a larger cushion against losses. However, if capital requirements are not implemented optimally, that is, based on the risks banks are taking, they might distort incentives and risks taken by banks might actually increase. The effect of capital requirements is also undermined by well-meaning bailouts of banks to ensure that depositors are repaid or the introduction of deposit insurance, both can increase the propensity of banks to take risks, making higher capital requirements necessary.

Similarly, liquidity reserves are supposed to ensure banks have sufficient cash reserves to meet the withdrawals of depositors and interbank loans; this should ensure the banking system is stable and bank runs unlikely to occur. However, we have seen that imposing liquidity requirements can increase the risks banks are taking, ensuring the liquidity of the bank, but negatively affecting their solvency. This would imply higher capital requirements are needed to limit this effect. This simple example highlights the complexity of bank regulations.

Banking regulation is not limited to the provision of loans and the holding of cash reserves, it also includes the procedures for closing a bank. Closing a bank can here also include that the bank is bought by another bank at a favourable price or it may be liquidated. Both will impose losses on the bank they would rather seek to avoid. Such closure decisions are not merely reactive to banks in distress, but might be used pro-actively before a bank is actually in distress. A regulator will consider the risks of a bank continuing to operate against the costs of closing it down instantly. Here the costs to the regulator may not be reflective of the social costs and hence decisions might be distorted by the incentives of the regulator making the closure decision. However, to make the consideration even more complicated, the threat of a bank closure will also affect the behaviour of banks, they might take decisions that will make it less likely that a bank is closed as it continuing to operate will be less risky or the costs of its immediate closure will be higher.

While assessing these regulations and their implications is already challenging, many banks are operating internationally and are subject to different regulatory regime sin each country. They can take advantage of different regulatory regimes across countries, leading to competition between countries that may reduce regulatory standards; it may even be that remaining regulations becoming less effective. These concerns suggest that a degree of regulatory cooperation across countries is desirable to ensure that such externalities are reduced.

While most non-financial companies are subject to a wide range of regulations, from consumer-protection to safety requirements for products and working practices and employment rights, they are rarely directly restricted on how to use the resources they have obtained. Banks are subjected to restrictions on how much loans they can provide for a given amount of capital and how much cash reserves they need to hold. This adds more distortions to the incentives of banks than for other companies, which need to be considered to assess the impact of any such regulation.

Part VIII
Macroeconomic implications

Banks are an important part of the economy. By providing loans they facilitate investment and consumption and thus support economic growth; with deposits they also enable excess funds to be invested productively with the ability to withdraw these at any time if the need arises. Furthermore, banks are used by central banks to implement their monetary policy decisions; it is the lending of funds to banks, whether it is the interest rate they are charged or the amount of funds provided, that will affect the interest rates that banks charge on loans and provide depositors with. These loan and deposit rates will then in turn affect decision-making by companies and depositors, affecting macroeconomic outcomes. While macroeconomic models have often acknowledged the importance of credit in an economy, models that explicitly take into account behaviour typical of banks are not as common as this importance suggests. In many cases banks are introduced into macroeconomic models, but they are mainly used to introduce frictions into the provision of loans. Such friction include interest rate differences between deposits and loans, moral hazard between banks and companies, or asymmetric information. However, these models often do not acknowledge that banks can generate loans without having obtained deposits, referred to as savings in macroeconomic models, beforehand.

We will here not seek to discuss comprehensive macroeconomic models that include banks or entities that are suppose to represent banks, but instead focus on the influence that banks have on macroeconomic outcomes and how banks are affected by economic policy decisions, in particular monetary policy. To this effect we will look at the implications that monetary policy decisions by central banks has on the decisions banks take in chapter 39, looking at interest rate changes the central bank announces as well as the provision of liquidity to the banking system. We will also see how other regulations may supplement monetary policy. In chapter 40 we will then investigate how the behaviour of banks can lead to fluctuations in the economy, which may give rise to business cycles or exacerbate such business cycles. Finally we will re-visit the topic of credit rationing in chapter 41; although credit rationing was discussed in chapter , we will take a specific macroeconomic look at the provision of loans and see that in times of high demand loans may not able available to all borrowers that seek them. However, we commence with chapter 38 by conducting an analysis of banks to create deposits themselves, and thus to create money.

This part will show that the behaviour of banks cannot be analysed in isolation of the macroeconomic conditions, or be taken as an external factor in the bank's decision-making, but instead have to form part of a more holistic analysis of banks. When regulating banks or using them as tools for the implementation of monetary policy, they might not always react as intended or their reaction has consequences that were not intended. For a full assessment of the consequences of any policy decision, economic policy or banking regulation, it is important to thoroughly understand the complex interactions between banks and their economic environment.

Chapter 38

Money creation

A common approach when modelling banks is to assume that banks collect deposits from individuals that seek to not consume this part of their wealth. Having deposited their funds, the bank then proceeds to use these deposits to provide loans. In doing so, the bank might retain some fraction of the deposits in cash or invest into other assets, it might also have some other sources of funds available to them. After a number of time periods, this loan is repaid and the proceeds thus obtained, used to repay the depositors. The proceedings of the loan act as an implicit collateral for the banks' obligations to repay deposits.

Banks, however, do not require deposits, or any other sources of funding, to provide loans. Suppose a bank, having no funding at all, decides to provide a loan to a company. It will do so by crediting the loan amount to an account the company has with the bank. This credit in the account of the company is a deposit. Therefore, the loan has created its own deposit and to provide a loan, the bank does not even require a deposit. That the company may subsequently use its deposit to transfer it to another bank or take out as cash, does not affect the creation of the deposit in the first place. It is not only that a deposit is not required to provide a loan, it is even that a deposit cannot be directly used to provide a loan. Even if a bank has deposits, the loan will nevertheless create its own deposit as before. The deposit provided to the bank externally can merely be used to act as a cash reserve, enabling the bank to allow the transfer of the deposits it has generated through the provision of loans or them being taken out as cash. The vast majority of deposits banks hold are generated from the provision of loans. It is ultimately only deposits that have been created by customers using cash or resulting from central bank loans that are of a different origin.

Whether deposits are generated from funds given to the bank externally or arising from the provision of loans, are for most aspects of banking of very little consequence. If considering the costs of loans, we still have to take into account that deposits attract interest and need that the bank needs to allow the depositor to withdraw it. Thus, for the lending decision it is of little consequence what the origin of a deposit is. Similarly, it is of no consequence for the depositor how the deposit it holds has been generated. In all cases, they expect their deposits to be repaid if they demand it. Where it is of some relevance, though, is when considering the level of loans a bank can provide to the wider economy. With the aforementioned creation of deposits from loans, there is no limit to how much loans a bank could provide, apart from regulatory constraints. If we assume that loans are used for investment purposes, there is no limit to the amount of investment that can be conducted; this is very much in contrast to the widespread view that investment is limited by the amount of savings (deposits) in an economy.

Of course, the amount of loans that can be given in an economy will be limited by the existence of profitable investment opportunities. Such limits are the result of the funds available for consumers to purchase the products that are produced. We will see in chapter 38.2 how banks allow a higher level of investment than would be possible in their absence, but also how the level of investment will be limited. Before that in chapter 38.1, however, we will explore the impact additional deposits from central banks or cash deposits have on a bank's ability to lend.

38.1 The money multiplier

Central banks conduct monetary policy by changing interest rates which they apply to loans they provide to banks or interest they pay on deposits banks keep at the central bank. Another way, central banks conduct monetary policy is by increasing or decreasing the amount they lend to banks. These two forms of monetary policy are, of course related, as an increase in the loans the central bank provides to banks should lead to a reduction in the interest rate they can charge. Regardless of the mechanism used in conducting their monetary policy, we will consider a central bank changing the amount it lends to bank by an amount of M and how this affects the amount of lending the banks are able to do. We will consider firstly a situation where the lending is not sensitive to interest rates and then consider how changed lending might affect interest rates, which will again affect the amount of lending that is feasible.

Lending expansion Let us assume the central bank provides the banks with additional funds of M that are credited as deposits in the banks; we might also interpret them as cash deposits by any customer. If we assume that banks need to hold a fraction $\rho < 1$ of their deposits as cash reserves, they can lend out an amount of $(1-\rho)M$. This amount of $(1-\rho)M$ is again generating deposits of $(1-\rho)M$ with the money being paid into customer accounts and transferred within the banking system. Of these new deposits banks receive, again a fraction $1-\rho$ can be lent out, thus $(1-\rho)^2 M$, generating new deposits again, such that this process continues ad infinitum. The total deposits generated are thus

$$D = \sum_{i=0}^{+\infty} (1-\rho)^i M = \frac{M}{\rho}. \tag{38.1}$$

The marginal impact of increasing the central bank money therefore is given by

$$\frac{\partial D}{\partial M}=\frac{1}{\rho}>1. \tag{38.2}$$

The amount of loans would be equivalent to all deposits, except the cash holdings, ρD, giving us $L=(1-\rho)D$. The increase in loans is then

$$\frac{\partial L}{\partial M}=\frac{1-\rho}{\rho}, \tag{38.3}$$

which is larger than 1 as long as $\rho<\frac{1}{2}$. Hence, lending is increasing more than the additional central bank funds, making the injection of central banks more expansionary.

Interest sensitive lending Assume the bank lends L at a rate r_L, where $\frac{\partial L}{\partial r_L}<0$, i.e. the higher the loan rate the less the demand for loans, and has deposits D on which it pays interest r_D with $\frac{\partial D}{\partial r_D}>0$, i. e. higher interest rates induce more deposits. In a competitive market, the interest rate is not affected by the amounts lent, L, and deposits D. Let us finally assume that any available deposits not held as cash, $(1-\rho)D$, and not lent out, $G=(1-\rho)D-L\leq 0$, can be invested in government securities at the risk-free rate r, while cash held does not generate any profits. Thus the profits of banks are given as

$$\begin{aligned}\Pi_B &= \pi\left(1+r_L\right)L+\left(1+r\right)G-\left(1+r_D\right)D \\ &= \left(\pi\left(1+r_L\right)-\left(1+r\right)\right)L \\ &\quad +\left(\left(1-\rho\right)\left(1+r\right)-\left(1+r_D\right)\right)D,\end{aligned} \tag{38.4}$$

where π is the probability of a loan being repaid.

The optimal amounts of borrowing and lending are given such that the following first order conditions are fulfilled:

$$\begin{aligned}\frac{\partial \Pi_B}{\partial L} &= \pi\left(1+r_L\right)-\left(1+r\right)=0, \\ \frac{\partial \Pi_B}{\partial D} &= \left(1-\rho\right)\left(1+r\right)-\left(1+r_D\right)=0,\end{aligned} \tag{38.5}$$

which solve for

$$\begin{aligned}1+r_L &= \frac{1+r}{\pi}, \\ 1+r_D &= \left(1-\rho\right)\left(1+r\right).\end{aligned} \tag{38.6}$$

If the central bank increases deposits in the bank by M through liquidity injection, the total amount of deposits is then $\hat{D}=D+M$. Using that $L=(1-\rho)D$ we can write this as $\hat{D}-\frac{L}{1-\rho}-M=0$. Totally differentiating this expression, we get

$$\left((1-\rho)\frac{\partial\hat{D}}{\partial r_D}-\frac{1}{\pi(1-\rho)}\frac{\partial L}{\partial r_L}\right)dr-dM=0, \tag{38.7}$$

using $\frac{\partial\hat{D}}{\partial r}=\frac{\partial\hat{D}}{\partial r_D}\frac{\partial r_D}{\partial r}=\frac{\partial\hat{D}}{\partial r_D}(1-\rho)$ and $\frac{\partial L}{\partial r}=\frac{\partial L}{\partial r_L}\frac{\partial r_L}{\partial r}=\frac{\partial L}{\partial r_L}\frac{1}{\pi}$ when inserting from (38.6). Using that $\frac{\partial\hat{D}}{\partial r_D}=\frac{\partial D}{\partial r_D}$, this can be solved for

$$\frac{\partial r}{\partial M}=\frac{1}{(1-\rho)\frac{\partial D}{\partial r_D}-\frac{1}{\pi(1-\rho)}\frac{\partial L}{\partial r_L}}>0. \tag{38.8}$$

Thus, if the deposits are increased by providing liquidity from the centra bank, the interest rate on government securities, r, rises, and thus through equation(38.6) do the deposit and loan rates. The reason is that the increased deposits the bak had need to be held, and this is only possible if the deposit rate increases; these increased costs are then also reflected in higher loan rates.

From $D=\hat{D}-M$, we get using (38.8) that

$$\begin{aligned}\frac{\partial D}{\partial M}&=\frac{\partial\hat{D}}{\partial M}-1=\frac{\partial\hat{D}}{\partial r}\frac{\partial r}{\partial M}-1=\frac{\partial\hat{D}}{\partial r_D}\frac{\partial r_D}{\partial r}\frac{\partial r}{\partial M}-1\\&=\frac{\partial\hat{D}}{\partial r_D}(1-\rho)\frac{\partial r}{\partial M}-1=\frac{\frac{\partial L}{\partial r_L}}{\pi(1-\rho)^2\frac{\partial D}{\partial r_D}-\frac{\partial L}{\partial r_L}}<0.\end{aligned} \tag{38.9}$$

Due to $\frac{\partial D}{\partial r_D}>0$ and $\frac{\partial L}{\partial r_L}<0$, we see also that $\left|\frac{\partial D}{\partial M}\right|<1$. Hence, the effect of interest rates rising when liquidity is injected into the banking system, causes the deposits to reduce due to the lower demand for loans.

Combined effect The increase in interest rates will thus partially off-set the increase in deposits as obtained in equation (38.2). The process of repeatedly creating new deposits will be limited by the effect of the interest rising.

The total effects on deposits are thus from (38.2) and (38.9)

$$\begin{aligned}\frac{\partial D}{\partial M}&=\left(1+\frac{\frac{\partial D}{\partial r_D}}{\pi(1-\rho)^2\frac{\partial D}{\partial r_D}-\frac{\partial L}{\partial r_L}}\right)\frac{1}{\rho}\\&=\frac{1}{\rho}\frac{\pi(1-\rho)^2\frac{\partial D}{\partial r_D}}{\pi(1-\rho)^2\frac{\partial D}{\partial r_D}-\frac{\partial L}{\partial r_L}}.\end{aligned} \tag{38.10}$$

With $L=(1-\rho)D$, we easily get

$$\frac{\partial L}{\partial M}=(1-\rho)\frac{\partial D}{\partial M}=\frac{1}{\rho}\frac{\pi(1-\rho)^3\frac{\partial D}{\partial r_D}}{\pi(1-\rho)^2\frac{\partial D}{\partial r_D}-\frac{\partial L}{\partial r_L}}>0. \tag{38.11}$$

If the interest rate sensitivity of loans is not too high, i. e. $-\frac{\partial L}{\partial r_L}<\pi(1-\rho)^2(1-2\rho)\frac{\partial D}{\partial r_D}$, we find that $\frac{\partial L}{\partial M}>1$ and loan expand more than the liquidity injection by the central bank. If $\rho > \frac{1}{2}$ this expression can never be fulfilled, but for the realistic case of small values for the fraction of cash reserves ρ, this expression can be fulfilled if the demand for loans is not too sensitive to loan rates.

Summary Injecting liquidity into the banking sector will increase deposits and loans cumulatively as the new funds are lent out, leading to more deposits being created by banks. This process is limited by the need to retain cash reserves and thus not all deposits can be lent out, and the sensitivity of loans to increasing loan rates that arise from the increased deposit rates, which are required to ensure all deposits created are actually held. This increases costs to the bank, necessitating higher loan rates. Thus injecting liquidity causes interest rates to rise and thereby partially offsets the increase in lending from this liquidity injection.

Reading Keiding (2016a, Chapter 11.2)

38.2 Limits to loan provision

In an aggregate view of the economy, companies pay the wages of their workers, who then use these wages to buy the products that have been produced. However, it is only this sale of products that generates the funds to pay the wages in the first place. This is often referred to as the 'cash-in-advance constraint' in an economy. Banks can help to overcome this problem by providing loans that allow the payment of wages before the products are sold. Banks can provide loans to companies such that they can conduct their investments and pay the wages of workers, who subsequently buy their products, without having to rely on deposits (wages) in the first place.

Let us consider an economy with three time periods, where in the first time period the investment is made by the company, the second time period sees the outcome of the investment (production), and finally in the third time period, the goods are sold and consumption takes place. We will analyse the constraints and incentives of each market participant in turn, companies, workers (consumers), and banks, before obtaining the overall equilibrium.

Companies Companies seek to maximize their final outcome. Due to the absence of production in the final period, this consists of their deposits in banks, denoted D_3^C, on which they receive interest r_D^3, in addition to the cash they are holding, C_3^C. We assume that the value of cash deteriorates at a rate $\rho \leq 1$ in each time period. This may be the result of inflation if we interpret all interest rates as real variables, or we might want to take into account that banks often charge account fees. Hence companies seek to maximize

$$\Pi_C = (1 + r_D^3)D_3^C + \rho C_3^C. \tag{38.12}$$

In this final time period, the funds available arise from the deposits in banks made initially in period 1, D_1^C, and the interest paid on this, r_D^1, the initial amount invested into cash, C_1^C, and the proceeds of production, V from time period 2, which from equation (38.16) we will see is assumed to be risk-free. These funds are used to invest into deposits, D_3^C, cash, C_3^C, and the repayment of any loan they might have taken out, L, on which an interest rate of r_L is applied. Thus we find that

$$(1 + r_D^1)D_1^C + \rho C_1^C + V = D_3^C + C_3^C + (1 + r_L)L. \tag{38.13}$$

The right-hand side denotes the funds available at the start of time period 3, deposits with interest from time period 1, cash from time period 1, and the proceeds from production in time period 2. The left-hand side shows the use of funds in time period 3, investment into deposits and cash, and the repayment of the loan.

The initial time period sees the funds of the company, their own capital K and the loan L, invested into deposits, D_1^C, cash C_1^C, the capital investment I, and the requisite labour input of N units, with a wage of w per unit. Hence

$$K + L = D_1^C + C_1^C + I + Nw. \tag{38.14}$$

We assume there is no mechanism to force companies to repay the loan, thus incentives need to be provided such that this repayment is ensured and thereby avoid strategic defaults by companies. The total gross proceeds after production is completed is $(1 + r_1^D)D_1^C + \rho C_1^C + V$ and we assume that this could be held in cash without repaying the loan. Alternatively, companies could repay the loan and invest the remaining proceeds into deposits, earning interest on them. Hence in order for companies to repay loans, we require that investment of all proceeds, less the loan repayment, into deposits yields a higher outcome than investing all proceeds into cash without repaying the loan. This gives us thr condition that

$$\begin{aligned}(1 + r_D^3)\left((1 + r_D^1)D_1^C + \rho C_1^C + V - (1 + r_L)L\right)& \\ \geq \rho\left((1 + r_D^1)D_1^C + \rho C_1^C + V\right).&\end{aligned} \tag{38.15}$$

We assume that when not repaying the loan, the proceeds can only be held in cash as depositing monies with any other bank would become known to the original bank and they then can enforce the repayment of the loan.

Finally, production is assumed to use a Leontief production function, the inputs consisting of capital investment, I, and labour, N. We set

$$V = (1 + R) \min\{I, \alpha N\}, \tag{38.16}$$

where R denotes a return on investment and α the relative importance of labour and capital investment, or the relative productivity.

Workers The wealth of workers in the final period consists of their final deposits, D_3^W, their final cash holdings, C_3^W, less the units of labour supplied, N, such that we have

$$\Pi_W = (1 + r_D^3)D_3^W + \rho C_3^W - N. \tag{38.17}$$

With no work being conducted in the final period, cash and deposits are rolled over from what the workers invested in the first time period, such that

$$D_3^W + C_3^W = (1 + r_D^1)D_1^W + \rho C_1^W. \tag{38.18}$$

The initial deposits and cash are obtained from the wages paid

$$D_1^W + C_1^W = wN \tag{38.19}$$

as there is no other source of income and workers have no initial wealth. We thus assume that workers are paid before production commences, as this allows them to subsequently purchase the produced goods from the company.

Banks The banks' final position will be consisting of the cash held, less the deposits to be repaid. There is no lending in time period 2 as all loans are being repaid. Therefore we have

$$\Pi_B = C_3^B - (1 + r_D^3)D_3, \tag{38.20}$$

where $D_3 = D_3^C + D_3^W$ denotes the deposits of companies and workers. This cash is accumulated from the repaid loans, new deposits made, and the previously owned amount of cash. This cash is not reduced by a factor ρ as we assume that banks have a competitive advantage in holding cash, e.g. in the form of depositing it with the central bank who pays them interest at a rate equal to inflation. They also repay existing deposits, such that their cash position is given by

$$C_3^B = (1 + r_L)L + D_3 + C_1^B - (1 + r_D^1)D_1, \tag{38.21}$$

with $D_1 = D_1^C + D_1^W$ the deposits of companies and workers that have been repaid with interest after period 1. The initial cash position consists of the deposits not lent out:

$$C_1^B = D_1 - L. \tag{38.22}$$

Inserting equations (38.21) and (38.22) into equation (38.20) we get for the bank profits that

$$\Pi_B = r_L L - r_D^1 D_1 - r_D^3 D_3, \tag{38.23}$$

representing the interest received on loans less the interest paid on deposits in either time period.

In a competitive banking environment, we have $\Pi_B = 0$ and hence

$$r_L = r_D^1 \frac{D_1}{L} + r_D^3 \frac{D_3}{L}. \tag{38.24}$$

Having characterised the behaviour of all market participants, we can now proceed to derive the equilibrium. This will include the determination of interest rates, deposits, cash holdings and bank lending.

Equilibrium Inspecting equation (38.23), we note that if $r_D^t < 0$, the bank would demand an infinite amount of deposits D_t as this increases its profits. Similarly, for $r_D^t > 0$ the demand for deposits would be as small as possible, both cannot be equilibria as the supply of deposits is finite and we thus require $r_D^t = 0$. In this case the bank would be indifferent to the amount of deposits it holds. Hence in equilibrium

$$r_D^t = r_L = 0, \tag{38.25}$$

where the result on r_L is derived from equation (38.24). The result of zero interest rates arises from our assumption that production is risk-free and no alternative investments exist, only loans for banks and deposits or cash for companies and workers.

From equations (38.12) and (38.17), it is obvious that holding cash in the final time period is not optimal for companies or workers as they would lose value given $\rho < 1$ and $r_D^3 = 0$. The same argument can be made from equations (38.13) and (38.18) for the cash holding in period 1, as the reduction in value reduces the value of future deposits, which in turn reduces future profits. Therefore we have

$$C_1^C = C_3^C = C_1^W = C_3^W = 0, \tag{38.26}$$

i. e. cash is not held by either companies or workers. For workers, equation (38.18) then implies $D_1^W = D_3^W$ as $C_t^W = 0$ and $r_D^1 = 0$. Equation (38.19) similarly gives us that $D_3^W = wN$ using $C_1^W = 0$, hence

$$D_1^W = D_0^W = wN \tag{38.27}$$

and equation (38.17) simplifies to

$$\Pi_W = (w - 1)N. \tag{38.28}$$

If $w > 1$ workers would supply an infinite amount of labour and for $w < 1$ no labour at all as this would maximize the profits of workers, hence we need

$$w = 1 \tag{38.29}$$

in equilibrium and workers are indifferent to the amount of work offered.

Optimal production requires from equation (38.16) that $I = \alpha N$ to avoid wasting any resources and hence

$$V = (1 + R)\alpha N. \tag{38.30}$$

Inserting equations (38.25), (38.26), (38.29), and (38.30) into equation (38.14), we obtain

$$L = D_1^C + (1 + \alpha)N - K \tag{38.31}$$

and from equation (38.13) that

$$D_3^C = (1 + R)\,\alpha N - (1 - \alpha)\,N + K = (\alpha R - 1)\,N + K. \tag{38.32}$$

From equation (38.12) we now get that $\Pi_C = D_3^C$, noting that $r_D^3 = 0$ and $C_3^C = 0$ from our above results. Thus companies maximizing their profits, will maximize the deposits in the final time period. From equation (38.32) we see that if $\alpha R > 1$ this increases in N, which we assume here to be fulfilled. This assumption means that we only consider companies that have investment projects that are sufficiently profitable (high R) and are not too capital intensive (high α) as we require a sufficient amount of labour income that can buy the products the company produces. Therefore, we seek the maximum number of working hours that are feasible to maximize company profits. Workers were shown to be indifferent to the number of hours worked and banks indifferent to the deposits this generates, it would thus be an equilibrium to find the maximum amount of work. Work amount N cannot become arbitrarily large as we need to ensure there are sufficient incentives to repay the loan, which requires equation (38.15) to be fulfilled. Inserting our previous results that $r_D^t = r_L = 0$, $C_1^C = 0$, and $V = (1 + R)\,\alpha N$, we obtain $D_1^C + (1 + R)\,\alpha N - L \geq \rho\left(D_1^C + (1 + R\alpha N)\right)$. Replacing L from equation (38.31), this becomes after re-arranging

$$K - \rho D_1^C \geq -N\left(\left(R - \rho(1 + R)\right)\alpha - 1\right). \tag{38.33}$$

If we assume that $\left(R - \rho\,(1 + R)\right)\alpha < 1$ the least constraining choice is

$$D_1^C = 0 \tag{38.34}$$

and even then this constraint will become binding as N increases. Hence inserting this result into equation (38.33) as an equality, we obtain the maximum working hours to be

$$\begin{aligned} N &= \frac{K}{1 - \alpha\left(R - \rho\,(1 + R)\right)}, \\ I &= \frac{\alpha K}{1 - \alpha\left(R - \rho\,(1 + R)\right)}, \\ L &= \frac{\alpha\,(1 - \rho)\,(1 + R)}{1 - \alpha\left(R - \rho\,(1 + R)\right)} K, \end{aligned} \tag{38.35}$$

where the second equation was obtained from our earlier result that $I = \alpha N$ and the final equation from (38.31), noting that $D_1^C = 0$ from equation (38.34).

Without banks, we would at most invest K into the capital expenditure I and labour wN. In the presence of banks, though, we have the ratio of the investment and the initial capital of the company given as

$$\Lambda = \frac{I + wN}{K} = \frac{1 + \alpha}{1 - \alpha\left(R - \rho(1 + R)\right)} \geq 1. \tag{38.36}$$

Thus, with bank loans, we observe larger investments than the company could achieve independently using its own capital. We can easily show that

$$\begin{aligned}
\frac{\partial \Lambda}{\partial \alpha} &= \frac{(1+R)(1-\rho)}{(1-\alpha\,(R-\rho(1+R)))^2} > 0, \\
\frac{\partial \Lambda}{\partial R} &= \frac{(1+\alpha)(1-\rho)}{(1-\alpha\,(R-\rho(1+R)))^2} > 0, \\
\frac{\partial \Lambda}{\partial \rho} &= -\frac{\alpha(1+\alpha)(1-\rho)}{(1-\alpha\,(R-\rho(1+R)))^2} < 0.
\end{aligned} \tag{38.37}$$

Hence, the more productive the technology is (R) and the more it relies on labour input (α), the more beneficial are banks as seen in the first two equations. Furthermore, if banks have more advantages, i. e. ρ is lower, the more beneficial are banks. Note that for $\rho = 1$ we have $\Lambda = 1$ and hence it is essential that banks have a comparative advantage in holding cash.

Even through the bank does not receive interest on cash, it will not lend out all deposits as from equations (38.22), (38.29), (38.34), and (38.35) we obtain

$$\begin{aligned}
C_1^B &= D_1 - L = D_1^W - L = N - L = K - \alpha N \\
&= \frac{1-\alpha(1+R)(1-\rho)}{1-\alpha\,(R-\rho(1+R))} > 0.
\end{aligned} \tag{38.38}$$

The reason for this result is, that banks cannot lend the entire deposits they hold, because this would violate the incentive constraint of the company, equation (38.15), and thus loans would not be repaid. Therefore, cash balances are held not for purposes of liquidity management or to account for the withdrawal of deposits, but to ensure not too much lending is conducted, which would result in loans not being repaid.

Banks creating their own deposits While the above results are an equilibrium, it should not be achievable in praxis. The reason is that deposits by workers, which are then lent out, are only received once investment takes place. These investments, however, require loans, which in the absence of deposits cannot be given. It is therefore that banks would need deposits to provide loans, but these deposits are only created once loans are given. It should be impossible initiate the first loan and hence the derived equilibrium would not be obtainable.

However, banks are able to achieve this equilibrium by creating their own deposits. They provide a loan to a company, but this loan is not financed by deposits the bank has collected, but instead the bank credits the loan into an account held by the company, it thus creates a deposit of equal size to the loan. The company then uses these newly created deposits to pay for their investments and the wages of the workers. These deposits are backed by the future revenue from the company generated from production, which it uses to repay the loan to the bank. These proceedings from the loan are then used to return the deposits to workers and the company, for which it constitutes its profits.

We can thus interpret banks as institutions that can finance a loan to companies through obligations of making payments in the future (repay deposits), backed by the repayments of these loans. Banks can create their own deposits that are backed by future income streams from loan repayments.

Summary If we assume that banks have a competitive advantage in managing cash, they can, through the provision of loans, expand the scale of investment in an economy and thereby increase economic welfare. They are able to do so even in the absence of deposits, i. e. savings in the economy, that they could use to provide these loans; instead banks are able to create their own deposits. These deposits are not backed by actual savings such as cash that depositors have provided to the bank, but by the promise to pay these deposits in the future based on returns generated from the loans they have given. Once these loans are repaid, banks are in a position to fulfill their obligation arising from the deposits they have created. The amount of lending banks can do is, however, limited by the need to ensure that loans are repaid. This does not allow for an unlimited provision of loans as the high amounts lent out would provide incentives for companies to not repay their loan. The more labour intensive the production technology is, the more loans can be created as the amount of wages paid increases and thereby the money by workers to buy the products of the company. Similarly a more profitable technology allows for more deposits to be created as the high profits of the company provide sufficient assurance that the deposits can be repaid.

Reading Donaldson, Piacentino, & Thakor (2018)

Conclusions

Banks do not rely on deposits to provide loans. Instead, each loan they give creates its own deposit as the loan amount is credited as a deposit to the borrower, who then can use this deposit to make payment for the purpose of the loan. In principle, this process can give rise to an infinite amount of loans and deposits in an economy, but in reality this will be limited by the amount of profitable investment opportunities that generate enough income so that workers can purchase the products. Banks hereby allow to circumvent the cash-in-advance fallacy in that the loan provided by the bank allows payment of the workers and the investment, prior to completing the production and sale of products. Providing too many loans, and subsequently a too large amount of products in the market, would not create sufficient demand as the wages paid to workers are not sufficient to pay for a large production.

Another limit emerges from the need to retain cash reserves. While deposits and loans can be created without bounds, as long as sufficient profitable investment opportunities are available for companies, the short-term nature of deposits requires banks to retain a certain amount of deposits as cash reserves, as outlined in chapter 4.1. Following a liquidity injection by the central bank, the obtained funds are treated as deposits that can be partially lend out, creating new deposits, giving rise to an expansion in bank lending. The higher the cash reserves are, the less loans are created from this liquidity injection. The expansion of loans is further more limited by the need to raise loan and deposit rates, mainly because deposits need to be held and more deposits in an economy require higher deposit rates, which in tern increase the costs of loans to banks.

Therefore, overall, banks can grant loans without relying on deposits to fund these, but instead create their own deposits. The amount of loans banks grant, and hence the amount of deposits created are only bound by the limits to profitable investment opportunities of companies that require loan financing and the amount of cash reserves that are required.

Chapter 39

Monetary policy

Banks are one of the key players in the transmission of monetary policy. It is the borrowing and lending rates of banks with the central bank that serve as a benchmark for deposit rates and many loan rates. Banks are also provided with liquidity by the central bank or have liquidity withdrawn, allowing them to expand lending or having to reduce lending. It is beyond the scope of this book to assess the impact such monetary policy decision has on the wider economy, instead we will focus on how monetary policy affects banks by looking at the impact monetary decisions, such as interest rates and liquidity provision, have on the decisions of banks.

Monetary policy has direct effects on investments by companies, that are strengthened by any lending constraints that banks might have, as chapter 39.1 will show, but it also affects the incentives of banks to provide loans as it changes the deposit rate as well as the loan rate, affecting the risks banks are willing to take as chapter 39.2 will show. The reduction of interest rates during the financial crisis 2008/9 and its aftermath has not only been justified with lowering loan cost to borrowers, but was also seen as essential to support banks during that time period. Chapter 39.3 will evaluate how lowering interest rates can incentivise banks to continue lending, even when companies are faced with adverse conditions.

Monetary policy can be complemented or replaced with other forms of liquidity management; chapter 39.4 will show how minimum liquidity requirements can affect the interest rates in the interbank market and thus the funding costs of banks. While liquidity injections should reduce the interest rates banks face due to mare liquidity being available in the market, we show in chapter 39.5 that an anticipation of its future withdrawal can actually increase the interest rates banks face in interbank markets at the time the liquidity provision is increased. We will finally consider in chapter 39.6 how internationally active banks react to changes in interest rates changes of one country by moving the provision of loans from one country to another.

39.1 The effect of lending constraints

Let us assume the central bank conducts monetary policy such that the loan rates banks charge, change. The might do so by changing the interest rate that banks can borrow from the central bank. The production of a company getting a loan L and holding capital K is given by

$$V = (1 + R)\, I, \tag{39.1}$$

where R denotes the return investments $I = L + K$. We assume that this return decreases as we increase the total investment I, thus $\frac{\partial(1+R)}{\partial I} < 0$, and $\frac{\partial^2(1+R)}{\partial I^2} > 0$. Nevertheless, we assume that the total output is increasing in its investment but at a diminishing rate, thus

$$\begin{aligned} \frac{\partial V}{\partial I} &= \frac{\partial\,(1+R)}{\partial I} I + (1+R) \geq 0, \\ \frac{\partial^2 V}{\partial I^2} &= \frac{\partial^2\,(1+R)}{\partial I^2} I + 2\frac{\partial V}{\partial I} \leq 0. \end{aligned} \tag{39.2}$$

Investment with lending constraints The company will now maximize their profits

$$\Pi_C = V - (1 + r_L)\, L, \tag{39.3}$$

where r_L denotes the loan rate. Companies are subject to a lending constraint in that the total amount to be repaid, $(1 + r_L)\, L$, must not exceed κK, i.e. there is a maximum leverage κ that limits the size of the loan. Thus we need to impose the constraint that

$$(1 + r_L)\, L \leq \kappa K. \tag{39.4}$$

Such a constraint might be imposed by regulation or be a the result of the lending policies of banks to ensure the risks companies take are sufficiently small, as discussed in chapter 3.2.3.

Maximizing equation (39.3) subject to equation (39.4), we get with λ denoting the Lagrange multiplier, that

$$\frac{\partial V}{\partial L} - (1 + r_L) - \lambda = 0. \tag{39.5}$$

We note that $\frac{\partial V}{\partial I} = \frac{\partial V}{\partial L}\frac{\partial L}{\partial I} = \frac{\partial V}{\partial L}$ as the loan and investment size, for given capital K are increasing at the same rate. If the constraint in equation (39.4) is binding, we have $\lambda > 0$ and equation (39.5) can be rewritten as

$$\frac{\partial V}{\partial I} = 1 + r_L + \lambda. \tag{39.6}$$

Firstly, we observe that a higher loan rate, as induced by monetary policy of the central bank, necessitates a higher marginal product of production. Using from equation (39.2) that the marginal product is decreasing in I, a higher loan rate implies a higher marginal rate, which in turn necessitates a lower level of investment

I and hence for a given capital K a lower loan L. Furthermore, an increase in the loan rate will cause the constraint (39.4) to become more binding. This in turn will increase the Lagrange multiplier λ. Therefore, a further reduction of investment would be optimal.

We thus find that with a binding lending constraint, investments and hence loans are reduced more than without a binding lending constraint, i. e. $\lambda = 0$.

Summary Higher loan rates can make lending constraints more binding as the amount to be repaid on the maturity of the loan is higher. The consequence that companies scale back their investments and thus lending more than when lending constraints are not binding. Therefore, monetary policy that affects loan rates has a more pronounced effect in companies that are subject to such lending constraints. Apart from regulatory limits on leverage, such constraints can also emerge from the availability of collateral or the need for banks to ensure that companies do not seek more risky investments with the loan proceeds.

Reading Freixas & Rochet (2008b, Ch. 6.2.2)

39.2 Risk taking

We assume that a central bank can control the risk-free rate r and uses this tool for purposes of monetary policy. Banks provide loans L at a rate r_L, financed by deposits D at interest r_D and equity K for which a return r_K is generated. Loans are repaid with probability π, which through monitoring can be influenced by banks at costs C that increase the less risky the loan becomes; specifically we assume these costs to be $\frac{1}{2}C\pi^2L$. Define the capital ratio κ as $K = \kappa L$ such that $D = L - K = (1-\kappa)L$. Banks have limited liability and can repay deposits only if the loan they granted is repaid. With depositors having the alternative to invest into risk-free securities that yield a return of r, we thus require $\pi\,(1+r_D)\,D = (1+r)\,D$, which solves for

$$1 + r_D = \frac{1+r}{\pi}. \tag{39.7}$$

If the loan is repaid, shareholders of the bank obtain a return of r_K, such that their expected outcome is $\pi\,(1+r_K)$. We decompose the expected return on equity into the risk free rate r and a risk premium δ, such that the total return equals $1+r+\delta$, and we have

$$1 + r_K = \frac{1+r+\delta}{\pi}. \tag{39.8}$$

Bank profits are now given by the repaid loans, less the repaid deposits, the required payments to shareholders, and the costs of monitoring loans. This gives us

$$\Pi_B = \pi\left(\left(1+r_L\right)L - \left(1+r_D\right)D - \left(1+r_K\right)K\right) - \frac{1}{2}C\pi^2L \tag{39.9}$$

Banks will now seek to maximize their profits by choosing the risk level π, the loans rate r_L, and the capital ratio κ.

Optimal risk level Assuming for now the loan rate r_L and capital ratio κ as given, we can determine the optimal risk π. Inserting for $D = (1-\kappa)L$, $K = \kappa L$, and equation (39.8) into the bank profits of equation (39.9), we get the first order condition that

$$\frac{\partial \Pi_B}{\partial \pi} = \left((1+r_L) - (1+r_D)(1-\kappa) - C\pi\right) L = 0, \tag{39.10}$$

which solves for

$$\pi = \frac{(1+r_L) - (1+r_D)(1-\kappa)}{C} \tag{39.11}$$

Using equation (39.7), to replace r_D, obtain after solving again for π that

$$\pi = \frac{(1+r_L) + \sqrt{(1+r_L)^2 - 4C(1+r)(1-\kappa)}}{2C} \tag{39.12}$$

Thus we easily see that with a given capital ratio κ

$$\frac{\partial \pi}{\partial (1+r)} = -\frac{1-\kappa}{\sqrt{(1+r_L)^2 - 4C(1+r)(1-\kappa)}} < 0 \tag{39.13}$$

and a higher risk-free rate set by the central bank will increase the risk banks are taking. The reason is that as the risk free rate increases, equation (39.7) suggests that the deposit rate increases, too. Thus, the profits of banks would increase and banks can reduce costs be reducing the level of monitoring, causing the risk to increase.

Optimal loan rate If we insert equations (39.11) and (39.8), while replacing $D = (1-\kappa)L$ as well as $K = \kappa L$, into the profit function of equation (39.9), we get

$$\Pi_B = \left(\frac{\left((1+r_L) - (1+r_D)(1-\kappa)\right)^2}{2C} - (1+r+\delta)\kappa\right) L. \tag{39.14}$$

The demand for loans, L, is not exogenously given, but will be affected by the loan rate r_L. As the risk π is given, the bank now sets the optimal interest rate r_L from the first order condition

$$\begin{aligned}\frac{\partial \Pi_B}{\partial (1+r_L)} &= \frac{(1+r_L)-(1+r_D)(1-\kappa)}{C}L \\ &\quad + \left(\frac{((1+r_L)-(1+r_D)(1-\kappa))^2}{2C}\right. \\ &\quad \left. -(1+r+\delta)\kappa\right) \times \frac{\partial L}{\partial (1+r_L)} \\ &= 0.\end{aligned} \tag{39.15}$$

We can differentiate the above expression with respect to the risk-free rate r and obtain

$$\frac{\partial^2 \Pi_B}{\partial (1+r_L)\,\partial (1+r)} = -\kappa \frac{\partial L}{\partial (1+r_L)} > 0, \tag{39.16}$$

where the final inequality arises if we assume that $\frac{\partial L}{\partial(1+r_L)} < 0$, i. e. higher loan rates reduce the demand for loans.

We furthermore know that $\frac{\partial^2 \Pi_B}{\partial(1+r_L)^2} < 0$ as this is the second order condition for the solution to equation (39.15) being a maximum. Using the implicit function theorem we have then obtain from this and equation (39.16) that

$$\frac{\partial (1+r_L)}{\partial (1+r)} = -\frac{\frac{\partial^2 \Pi_B}{\partial(1+r_L)\partial(1+r)}}{\frac{\partial^2 \Pi_B}{\partial(1+r_L)^2}} > 0. \tag{39.17}$$

Not surprisingly, a higher risk-free rate increases loan costs, This is the result of the costs of banks increasing, namely the deposit rates are increasing as per equation (39.7) as well as the cost of equity raise as we see from equation (39.8).

Optimal capital ratio Using equation (39.7) in equation (39.11) and inserting from equation (39.8), as well as replacing $D = (1-\kappa)L$ and $K = \kappa L$, into the profit function of equation (39.9), we get

$$\Pi_B = \left(\pi (1+r_L) - (1+r+\delta\kappa) - \frac{1}{2}C\pi^2\right)L. \tag{39.18}$$

The optimal capital ratio κ, is then obtained from the first order condition

$$\frac{\partial \Pi_B}{\partial \kappa} = \left(-\delta + \frac{\partial \pi}{\partial \kappa}((1+r_L) - C\pi)\right)L = 0. \tag{39.19}$$

As the solution to this equation is a maximum, we will have that $\frac{\partial^2 \Pi_B}{\partial \kappa^2} < 0$.

From equation (39.13), we now obtain that

$$\frac{\partial^2 \pi}{\partial (1+r_L)\,\partial \kappa} = \frac{(1+r_L)^2 - 2C(1+r)(1-\kappa)}{\left((1+r_L)^2 - 4C(1+r)(1-\kappa)\right)^{\frac{3}{2}}} > 0 \tag{39.20}$$

and from equation (39.15) we get that

$$\begin{aligned}\frac{\partial^2 \Pi_B}{\partial (1+r_L) \partial \kappa} &= \frac{\partial \pi}{\partial \kappa} L + \frac{\partial L}{\partial (1+r_L)} \left(\frac{\partial \pi}{\partial \kappa} (1 + r_L - C\pi) - \delta \right) \\ &= \frac{\partial \pi}{\partial \kappa} L,\end{aligned} \tag{39.21}$$

where we used equation (39.11) to replace π and the final equality arises due to equation (39.19) eliminating the second term. With $\frac{\partial^2 \Pi_B}{\partial (1+r_L)} < 0$ due to the solution being a maximum, we have from the implicit function theorem that

$$\frac{\partial (1+r_L)}{\partial \kappa} = -\frac{\frac{\partial^2 \Pi_B}{\partial (1+r_L)\partial \kappa}}{\frac{\partial^2 \Pi_B}{\partial (1+r_L)^2}}. \tag{39.22}$$

The sign of $\frac{\partial^2 \Pi_B}{\partial (1+r_L)\partial \kappa}$ is identical to that of $\frac{\partial (1+r_l)}{\partial \kappa}$ as the sign of the denominator is negative. Further, from equation (39.21), the numerator and $\frac{\partial \pi}{\partial \kappa}$ have the same sign.

Using equation (39.12) we easily obtain

$$\frac{\partial \pi}{\partial \kappa} = \frac{1+r}{\sqrt{(1+r_L)^2 - 4C(1+r)(1-\kappa)}} > 0. \tag{39.23}$$

A higher capital ratio reduces the risks taken as the bank exposes themselves to larger losses from loans not being repaid, inducing them to limit these risks. Using this result we easily obtain that

$$\begin{aligned}&\frac{\partial (1+r_L)}{\partial \kappa} > 0 \\ &\frac{\partial \kappa}{\partial (1+r)} = \frac{\partial \kappa}{\partial (1+r_L)} \frac{\partial (1+r_L)}{\partial (1+r)} > 0,\end{aligned} \tag{39.24}$$

where for the inequality in the second equation, we have used the result in equation (39.17). That the loan rate increases as the capital ratio rises, is because the higher equity costs, compared to deposits, as can be seen comparing equations (39.7) and (39.8), increase the costs to banks, who in turn increase the loan rate to compensate for these higher costs. We also notice that higher risk-free rates will increase the capital ratio banks hold. The reason for this effect is that a higher risk-free rate increases the loan rate as we see in equation (39.17), which induces banks to reduce monitoring in order to save costs. However, less monitoring implies a higher risk, i. e. a lower value for π, which in turn causes the deposit rate and the return on equity to rise, see equations (39.7) and (39.8), making profit margins even smaller, giving even more incentives to reduce monitoring and increase risks. By increasing their capital, banks reduce increase their exposure to these risks, making taking this additional risk less attractive. Given this larger exposure of banks to the risk they are taking, banks have a higher incentive to increase monitoring, counteracting the reduced monitoring from the lower profit margins and due to the lower risk taken, increasing profit margins again.

Total effect We see two effects from increasing risk-free rates. Firstly, we notice that it increases the capital ratio, which from equation (39.23) would reduce the risks taken. The second effect is that from equation (39.17) we know that the loan rate increases, but so does the deposit rate from equation (39.7). Given the sensitivity of lenders to the loan rate, we assumed that $\frac{\partial L}{\partial(1+r_L)} < 0$, the effect on the deposit rate will be higher as we assumed they are only compensated such that they break even and hence no increase in deposits is observed when increasing deposit rates. This will reduce overall profits. To recover these profits, banks will monitor less and thus increase risks. It is now that this second effect from monitoring dominates the effect of increasing the capital ratio, as can be seen from equation (39.13). Therefore, an increase in the risk-free rate will increase the risks taken by the bank.

Summary If the central bank increases the risk-free rate, it does not only affect loan and deposit rates, but also the banks' incentives of risk-taking. Increasing interest rates will reduce profit margins as due to the sensitivity of borrowers to loan rates, loan rates cannot be increased to the same extend as deposit rates have to be. Banks will react to these lower profits margins by cutting back on the monitoring of loans, causing risks to increase. This effect is only partially reduced by increasing the capital they hold in reaction. Thus banks are increasing risks as risk-free rates rise, but do so equipped with a higher capital ratio.

Reading Dell'Ariccia, Laeven, & Marquez (2014)

39.3 Bank bailouts

If a bank fails, it may be bailed out by the government or central bank. While commonly this is done t protect depositors, we are here concerned about the continued provision of loans to companies. If a bank was to fail, these loans would seize, and it is this scenario that a bailout seeks to avoid. We assume that banks can give loans that are repaid with probability π_H charging interest r_L^H, or provide loans that are repaid with probability $\pi_L < \pi_H$ at interest rate r_L^L. Alternatively, they can or invest into a risk free asset yielding a fixed return of r. Banks finance their loans L using deposits D and equity $K = \kappa L$, where κ denotes the capital ratio. If loans are not repaid, the bank cannot meet its obligations to depositors as banks have limited liability.

Choice of low-risk loans Depositors require their expected returns to be equal to the risk-free rate as that is the alternative investment opportunity they have. If the bank is bailed out with probability p and depositors repaid their initial deposit without interest, then if depositors know the type of loan the bank provides, deposit rates r_D are such that

$$\left(\pi_i\left(1+r_D^i\right)+\left(1-\pi_i\right)p\right)D=(1+r)\,D, \tag{39.25}$$

where the first term denotes the return if the loan is repaid and subsequently the bank can repay their deposits, and the second term the payment to depositors following a bailout if the loan is not repaid. We can solve this equation for

$$1+r_D^i=\frac{1+r-\left(1-\pi_i\right)p}{\pi_i}, \tag{39.26}$$

where $\frac{1-\pi_i}{\pi_i}p$ can be interpreted as a subsidy to the bank, arising from the bailout. We furthermore assume that low risk loans are profitable as we have $\pi_H\left(1+r_H^H\right)\geq 1+r_D^H$, while high risk loans are not profitable due to $\pi_L\left(1+r_L^L\right)<1+r_D^L$.

Bank profits are then given by

$$\begin{aligned}\Pi_B^i&=\pi_i\left(\left(1+r_L^i\right)L-\left(1+r_D^i\right)D\right)\\&=\pi_i\left(\left(1+r_L^i\right)-\left(1+r_D^i\right)(1-\kappa)\right)D,\end{aligned} \tag{39.27}$$

using that $D=(1-\kappa)\,L$. For the investment into the risk-free asset we set $\pi_i=1$ and $r_L^i=r$. If $\Pi_B^H>\Pi_B^L$, the bank will grant the loan with the lower risk rather than the higher risk. Using equation (39.27) after inserting equation (39.26), we easily obtain the requirement that

$$\kappa\geq\kappa^*=1-\frac{\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)}{p\left(\pi_H-\pi_L\right)}. \tag{39.28}$$

Thus if a bank has sufficient capital, it will choose the low risk loan. Similarly, if $\left((1+r)-\left(1+r_D\right)(1-\kappa)\right)D\leq\Pi_B^H$, the bank will grant the low-risk loan rather than invest into the risk-free asset. Here, from equation (39.26) with $\pi_i=1$, we get $1+r_D=1+r$. Hence we require that

$$\kappa\leq\kappa^{**}=1-\frac{(1+r)-\pi_H\left(1+r_L^H\right)}{p\left(1-\pi_H\right)}. \tag{39.29}$$

If banks have not too high capital, they will provide the low-risk loan. For low risk loans to be feasible, we need that $\kappa^*<\kappa^{**}$, which we assume to be fulfilled here.

We can rewrite equations (39.28) and (39.29) as

$$\begin{aligned}p\leq p^*&=\frac{\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)}{(1-\kappa)\left(\pi_H-\pi_L\right)},\\1+r\leq 1+r^*&=\pi_H\left(1+r_L^H\right)+(1-\kappa)\,p\left(1-\pi_H\right).\end{aligned} \tag{39.30}$$

Exogenous shock Let us now assume that the return of the borrower, R_H, in the low risk company is subject to an exogenous shock $0\leq\xi\leq 1$ such that $1+R_H=\left(1+R_H^0\right)(1-\xi)$. We do not assume that a similar shock affects the high-risk company. A reasoning can be that in times of crises it is often the traditional and well-established companies that are affected, while more innovative and technologically advanced companies often see such crises as a chance to offer new products.

The company profits are then

$$\Pi_C^H = \pi_H \left((1 + R_H) - \left(1 + r_L^H \right) \right) L \tag{39.31}$$

and a bank extracting all company surplus sets

$$1 + r_L^H = 1 + R_H = \left(1 + R_H^0 \right) (1 - \xi) . \tag{39.32}$$

If we have a lower bound for the interest rate at $\underline{r}$, and the regulator seeks to avoid bailouts, i. e. $p = 0$, by reducing the risk-free rate, then using equation (39.32), we see that the second line in equation (39.30) is fulfilled if $1 + \underline{r} \leq 1 + r^*$. This can be solved for

$$\xi \leq \underline{\xi} = 1 - \frac{1 + \underline{r}}{\pi_H \left(1 + R_H^0 \right)} . \tag{39.33}$$

Commonly it would have been assumed that the lower bound for interest rates would be zero, thus $\underline{r} = 0$, however, the financial crisis of 2008/9 has shown that negative interest rates are not only theoretically possible, but have been implemented in a number of countries. Therefore, no general lower limit on interest rates seem to exist, even though large negative value were not used during the aforementioned financial crisis.

Thus for small shocks $\xi < \underline{\xi}$, the interest rate r is lowered until it reaches its lower bound $\underline{r}$. Banks do not need to be bailed out in such a case as they continue to provide low-risk loans. The lower risk-free rate reduces the attractiveness of the risk-free asset and induces the bank to provide low-risk loans.

If, on the other hand, $\xi > \underline{\xi}$, the subsidy of a bailout needs to be added to ensure banks provide these low-risk loans. This ensures the deposit rates remain sufficiently low to provide banks with sufficient profits from low-risk loans. If we set $1 + r^* = 1 + \underline{r}$ in the second line of equation (39.30), we get using equation (39.32) that the probability of a bailout needs to be set at

$$p = \frac{\left(1 + \underline{r} \right) - \pi_H \left(1 + R_H^0 \right) (1 - \xi)}{(1 - \kappa)(1 - \pi_H)} . \tag{39.34}$$

Nor surprisingly, we see that as the shock ξ increases, the probability of a bailout also increases. The reason is that the lower loan rate that can be charged by the bank reduces the profitability of the low-risk loan; to maintain the profitability the deposit rate needs to reduce further, which can be achieved by a higher likelihood of a bailout. To avoid the bank providing high-risk loans, we need that $p \leq p^*$ as in the first line of equation (39.30). Using equation (39.34), this is fulfilled as long as

$$\xi \leq \overline{\xi} = 1 - \frac{\frac{\pi_H - \pi_L}{1 - \pi_H} \left(1 + \underline{r} \right) + \pi_L \left(1 + R_L \right)}{\left(1 + \frac{\pi_H - \pi_L}{1 - \pi_H} \right) \pi_H \left(1 + R_H^0 \right)} , \tag{39.35}$$

with $1 + r_L^L = 1 + R_L$, where banks extract all surplus from high-risk borrowers. For any shocks $\xi > \overline{\xi}$, we need to set $p = p^*$ and the probability of a bailout is reducing as the shock increases even more as we can see from inserting equation (39.32) into equation (39.30). This result arises because high shocks reduce the profits from low-risk loans and combined with a reduced deposit rate from the expected bailout, would otherwise induce them to provide high-risk loans.

Thus, for small shocks $\xi \leq \underline{\xi}$, monetary policy is sufficient to induce banks to provide socially desirable loans, for larger shocks this needs to be supplemented by bailouts to subsidize the deposit rates banks have to pay. Even larger shocks $\xi > \overline{\xi}$ then require a lower subsidy, that is bailout probability, as not to induce banks to provide risky loans.

The lowest probability of a bailout is 0, thus we require $\pi^* \geq 0$ in equation (39.30). This implies that

$$\xi \leq \overline{\overline{\xi}} = 1 - \frac{\pi_L \left(1 + R_L\right)}{\pi_H \left(1 + R_H^0\right)}. \tag{39.36}$$

Hence if $\xi > \overline{\overline{\xi}}$, banks will not provide low-risk loans. The shock to the low-risk companies is so big that their expected returns are less than that of high-risk companies, making the provision of low-risk loans less profitable than high-risk loans.

Summary We see that monetary policy can support the decisions of banks to provide low-risk loans in the event of shocks to the profitability of companies. For small shocks a reduction of interest rates will be sufficient. Larger shocks require the introduction of potential bailouts, where the likelihood of a bailout is initially increasing in the size of the shock and then decreasing. For even larger shocks, even a bailout cannot induce banks to provide low-risk loans.

Reading Acharya, Lenzu, & Wang (2021)

39.4 Liquidity regulation as monetary policy

Banks are commonly required to hold a minimum amount of cash reserves. While this ensures that banks are not overly vulnerable to liquidity shortages, such cash reserves also have the effect of limiting the amount of lending on the one hand and on the other hand reducing the requirement of the bank to seek additional central bank funding if they face a liquidity shortage. We will consider here how the central bank can affect the funding costs of banks in the interbank market through such liquidity requirements. It is thus that liquidity requirements may supplement monetary policy decision conducted through liquidity provision to banks or changing the interest rates at which banks can obtain loans from the central bank or deposit funds with them.

We assume that banks have to hold a minimum cash reserve $\underline{R}$; this cash reserve will be composed of their current cash reserves, R, which are affected by a random and unknown liquidity shock S. Such a liquidity shock might be arising from the unexpected withdrawal of deposits, but it might also be a negative liquidity shock in that the bank obtains additional deposits. Thus the net cash reserves are given by $R - S$, where S can be positive or negative. Banks can increase (decrease) these cash reserves by borrowing from (lending to) the interbank market, B, at an interbank rate r_B, and by borrowing $M + \hat{M}$ from the central bank at its interest rates r_M and $\hat{r}_M$. The central bank is assumed to provide loans of any size, but will only apply the more favourable discount rate r_M to lending that is covered by collateral C the bank can provide, denoted M; any lending exceeding the collateral the bank can provide, denoted $\hat{M}$, will be charged a penalty rate $\hat{r}_M > r_M$.

In order to determine the required borrowing from the central bank, we need to determine the liquidity shortage relative to the minimum reserve requirements prior to this borrowing, which is given by $\underline{R} - (R - S + B)$. The bank will borrow from the central bank at the discount rate r_M if this shortage is less than the collateral they have available, thus for an amount of $\min\left\{C; \underline{R} - (R - S + B)\right\}$, provided there is a shortage, which gives us $\max\left\{0; \min\left\{C; \underline{R} - (R - S + B)\right\}\right\}$. If the liquidity shock is so large that the amount the bank has to borrow from the central bank exceeds the collateral available, the bank will borrow the largest possible amount at the discount rate, C, and then borrow the remainder at the penalty rate $\hat{r}_M$. The expected borrowing at the discount rate is thus given by

$$M = \int_{R+B+\underline{R}}^{R+B+C-\underline{R}} \left(S - \left(R + B + C - \underline{R}\right)\right) dF(S) + \int_{R+B+C-\underline{R}}^{+\infty} C dF(S), \tag{39.37}$$

where $F(\cdot)$ denotes the distribution function of the liquidity shock. The first term denotes the case where the amount of collateral is sufficient for the bank to obtain the full amount required from the central bank at the discount rate. The liquidity shock needs to be sufficient large such that the existing cash reserves, $R + B$ cannot meet the liquidity requirements $\underline{R}$ as $S > R + B - \underline{R}$. The liquidity shock must not be so large that the liquidity requirements cannot be met by using the discount rate alone, $S < R + B + C - \underline{R}$. If the liquidity shock is larger, then the bank will obtain a loan from the central for the full value of the collateral and the remainder will then be raised at the penalty rate, hence

$$\hat{M} = \int_{R+B+C-\underline{R}}^{+\infty} \left(S - \left(R + B + C - \underline{R}\right)\right) dF(S). \tag{39.38}$$

The profits of the bank are now given by the repayment of the loans L they have provided ar loan rate r_L, where the repayment is completed with probability π, and from which the remaining depositors, $D - S$ are repaid with interest r_D. In addition, the bank has to pay or obtains interest r_B on their interbank loans B and pay interest on their borrowing from the central bank. We thus have

$$\Pi_B = \pi (1 + r_L) L - (1 + r_D) (D - S) - r_B B - r_M M - \hat{r}_M \hat{M}. \tag{39.39}$$

Banks will be anticipating that they may face a liquidity shock and will therefore seek to rebalance their liquidity by obtaining or making interbank loans. Maximizing their profits over such interbank lending and borrowing, banks obtain the first order condition that

$$\begin{aligned} \frac{\partial \Pi_B}{\partial B} &= -r_B + \left(F\left(R + B + C - \underline{R}\right) - F\left(R + B - \underline{R}\right)\right) r_M \\ &\quad + \left(1 - F\left(R + B + C - \underline{R}\right)\right) \hat{r}_M \\ &= 0. \end{aligned} \tag{39.40}$$

In obtaining this first order condition, we have used that

$$\begin{aligned} \frac{\partial M}{\partial B} &= -\left(F\left(R + B + C - \underline{R}\right) - F\left(R + B - \underline{R}\right)\right), \\ \frac{\partial \hat{M}}{\partial B} &= -\left(1 - F\left(R + B + C - \underline{R}\right)\right), \end{aligned}$$

which has been obtained using the Leibniz integral rule. The first order condition can now be solved for

$$r_B = \hat{r}_M - F\left(R + B - \underline{R}\right) r_M - F\left(R + B + C - \underline{R}\right) (\hat{r}_M - r_M). \tag{39.41}$$

We can now easily get that

$$\begin{aligned} \frac{\partial r_B}{\partial \underline{R}} &= f\left(R + B - \underline{R}\right) r_M + f\left(R + B + C - \underline{R}\right) (\hat{r}_M - r_M) > 0, \\ \frac{\partial r_B}{\partial C} &= -f\left(R + B + C - \underline{R}\right) (\hat{r}_M - r_M) < 0, \end{aligned} \tag{39.42}$$

where the inequalities arise due to the penalty rate exceeding the discount rate, $\hat{r}_M > r_M$; $f(\cdot)$ represents the density function.

These last two relationships show how central banks can use liquidity requirements to complement or even replace their monetary policy. If the central bank, as the regulator of banks, would increase the liquidity requirements by increasing $\underline{R}$, the equilibrium rate in the interbank market would increase. This is equivalent to a tightening of monetary policy as the costs of banks will increase, which will be reflected in higher loan and deposit rates in the future. The higher liquidity requirements will make it more likely that the bank will have to obtain a loan from the central bank, which increases their costs, that will be reflected in a higher interbank rate as banks providing such loans will be requiring a higher compensation for giving up liquidity, while banks seeking such loans are willing to pay a higher interest rate to avoid having to rely on central bank borrowing.

Similarly, the central bank may restrict the use of collateral more, reducing C, either by increasing collateral requirements for loans or by reducing the type of assets that can be used as collateral. In this case, again the costs of banks are increased and this amounts to the equivalent of tightening monetary policy. The increased costs are here arising because banks can obtain fewer loans are the discount rate, but will have to rely more on the higher penalty rate.

We have thus seen that central banks have other tools than conventional monetary policy to affect the borrowing costs of banks and hence the loan and deposit rates they set. Tightening (loosening) liquidity requirements has the same effect as tightening (loosening) monetary policy as this policy increases (reduces) the borrowing costs to banks. In the same way can the use of collateral be more (less) restricted and increase (decrease) the borrowing costs of banks by making them more (less) reliant on borrowing at the penalty rate. It is noteworthy that for this policy to become effective, banks do not have to borrow from the central bank, and if they do, the costs they are facing for doing so in the form of the interest rate charged, were here assumed to be unaffected; it is the impact on the interbank rate that increases the borrowing costs of banks and through which this policy is effective.

Reading Monnet & Vari (2023)

39.5 Liquidity provision and interest rates

A common assumption in monetary policy is that the injection of liquidity into the banking system by central banks will reduce interest rates. The argument commonly is that banks having more cash reserves due to the funding obtained from the central bank will demand less funds in the interbank market or provide more interbank loans, reducing the interest rate there; this will in turn reduce the funding costs of banks causing loan and deposit rates to fall. We will re-evaluate this relationship for the case that banks anticipate that such central bank funding, while provided now, will be reduced in the near future.

A bank provides a loan L_0 to a company, who seeks to repay this loan with interest r_L, and holds back cash reserves R. This loan will be repaid with certainty, but the company might become distressed and require an additional loan L_1 to complete the investment and repay both loans, provided this additional investment is successful. The probability of the company becoming distressed is $1 - \pi$. Whether the company can become distressed in the first place depends on the state of the economy; a well-performing economy will allow the company to repay the loan without additional investment, while in a poorly-performing economy an additional investment would be required for a distressed company. The economy is poorly-performing with probability p. With this additional loan, the original loan can always be repaid.

If the economy is performing poorly, the company seeks an additional loan L_1 at a loan rate r_L^1 if it is distressed; thus an additional loan is requested with probability $p\,(1 - \pi)$. This additional loan can be funded by raising additional funds from long-term capital E_1 and interbank loans B on which interest r_B is payable. Providing the additional loan will then give the bank additional profits from this loan.

In the case the company is not distressed, the bank does not require funds to provide an additional loan to the company, but can provide an interbank loan $\hat{B}$ to other banks whose companies are in distress; to do so they may raise additional long-term capital. We assume that the long-term capital is raised before the bank knows about the distress of the company, but after they know the state of the economy. Hence, they will not raise additional long-term capital if the economy is well-performing, but would do so in a poorly-performing economy and as they do not know whether the company is distressed or not, they will raise the same amount in both cases. However, we assume that only a fraction γ of banks would provide interbank loan as they are not sure about the quality of other banks and whether they will be able to repay their interbank loans.

Using the above arguments, the profits of the bank are then given by

$$\begin{aligned}\Pi_B = & \left(1 + r_L^0\right) L_0 + R - M - (1 + r_D) D - E_0 - \frac{1}{2} c E_0^2 \\ & + p (1 - \pi) \left(r_L^1 L_1 - r_B B - \frac{1}{2} c E_1^2\right) \\ & + p \pi \gamma \left(-r_B \hat{B} - \frac{1}{2} c E_1^2\right).\end{aligned} \tag{39.43}$$

We assume that raising long-term capital is costly and increases in the amount raised. The source to fund the initial loans are the deposits D on which interest r_D is payable, and long-term capital E_0. The bank will also have obtained central bank funding M, which for simplicity we assume is provided at not cost to the bank with zero interest, and held back reserves R. Providing the initial loan will be costly to the bank and increasing in the loan size, for example due to more stringent due diligent requirements for larger loans; these costs are to be financed upfront and reduce the amount that can be lent out. We thus require that

$$L_0 + \frac{1}{2} c L_0^2 + R = D + M + E_0. \tag{39.44}$$

The interbank loan sought if the company is distressed is given by

$$B = L_1 + D + \lambda M - E_1 - R. \tag{39.45}$$

Banks require total funds of L_1 to provide the additional loan and need to repay depositors. This arises from the distress of the company, which makes depositors concerned about the failure of the loan and they seek to withdraw these. In addition, we assume that the central bank will withdraw a fraction λ of the funding they have provided. This funding need is reduced by the amount of long-term capital that is obtained, E_1, and the cash reserves held, R.

Finally, if the company is not in distress, the bank does not require funds to provide an additional loan nor will depositors be concerned about the ability of the bank to repay deposits and hence none are withdrawn. This leaves the bank with a need to finance the withdrawn central bank funding, less the long-term capital raised and the cash reserves held. Thus

$$\hat{B} = \lambda M - E_1 - R. \tag{39.46}$$

The bank will now seek to determine the loan amount and long-term capital raised that maximises their profits. If we solve equation (39.44) for the deposits D and insert the resulting expression into equations (39.43) and (39.45), we easily obtain

$$\begin{aligned}
\frac{\partial \Pi_B}{\partial L_0} &= \left(1 + r_L^0\right) - \left(1 + r_D + p\left(1 - \pi\right) r_B\right)\left(1 + cL_0\right) = 0, \\
\frac{\partial \Pi_B}{\partial E_0} &= r_D - cE_0 + p\left(1 - \pi\right) r_B = 0, \\
\frac{\partial \Pi_B}{\partial E_1} &= \left(p\left(1 - \pi\right) + p\pi\gamma\right)\left(r_B - cE_1\right) = 0.
\end{aligned} \tag{39.47}$$

The interbank market needs to clear such that the amount of interbank borrowing is identical to the amount of interbank lending. We therefore require that

$$p\left(1 - \pi\right) B = p\pi\gamma\hat{B} \tag{39.48}$$

as interbank loans are sought if the economy is performing poorly and the company in distress; interbank loans are provided if the company is not in distress and the bank is willing to provide a loan. Inserting for the interbank loans from equations (39.45) and (39.46), and solving the final first order condition in equation (39.47) for E_1, we obtain that

$$r_B = c\left(\lambda M - R + \frac{1 - \pi}{1 - \pi - p\pi}\left(L_1 + D\right)\right), \tag{39.49}$$

We now easily see that as long as $\lambda > 0$, and hence central bank funding will be withdrawn in the future, the interbank loan rate is increasing in the central bank funding M. This is because the withdrawal of funds in the future makes it more difficult for banks to provide the funding for additional loans and they therefore rely more on interbank loans. Here the higher central bank funding will increase the amount of loans provided as it is not optimal for banks to retain all such funding as cash reserves given that cash reserves do not generate any profits; this will then cause a potential liquidity shortage in the future as this funding is withdrawn.

We have thus seen that increasing central bank funding to banks with the threat of reducing this funding in the near future will not reduce the interbank loan rate, and hence funding costs of banks. Increased funds available from the central bank will be invested into loans and increase the reliance on interbank loans to cover any liquidity shortage, leading to a rise in the interbank loan rate and hence funding costs for banks. Permanently increasing the funding to banks, $\lambda = 0$, would not affect the interbank loan rate as all central bank funding would be used to provide additional loans and hence no effect on interest rates will be observed. It is thus the use that banks make of the additional funds to provide loans that drives this result; banks not retaining the funds to increase their liquidity will be counterproductive as this will actually reduce the liquidity in the market if these funds are withdrawn, leading to an increase interbank loan rates.

We have seen from equation (39.44) that an increase in central bank funding, M, will increase the provision of loans and from the first order conditions (39.47) we see that the resulting increase in the interbank rate, r_B, will limit the increase in the loan provision, while also reducing the amount of initial long-term-capital, E_0, further limiting the increase in loans that can be provided. It is thus that increasing central bank funding will increase the provision of initial loans, but this is limited by increasing interbank loan rates and hence funding costs for any liquidity shortfalls. Due to higher central bank funding, banks are exposed to more liquidity risk from the larger size of the initial loans that result in potentially larger liquidity demands for additional loans. It is therefore that central bank funding increases not only the provision of loans, but also the exposure to liquidity risk, which is partly offset by higher interbank rates.

Reading Acharya & Rajan (2024)

39.6 International spillover of monetary policy

Banks are often operating in multiple countries, in particular are they providing loans across borders. In order to provide basic services to their customers in other countries, banks will maintain cash reserves in their home market as well as the foreign market. However, for their funding such banks often rely mainly on their home market. Monetary policies across borders will often vary and banks will be facing different interest rates in the countries they are operating in; these different interest rates may affect the decisions of banks in which countries they provide loans as monetary policy changes. We will here investigate how banks react to different monetary policy decisions by changing the allocation of loans between countries.

Let us assume a bank provides loans in their home market as well as a foreign market, L_H and L_F, respectively, at loan rates r_L^H and r_L^F. The loans are repaid with probability π_H in the home market and π_F in the foreign market. In each country the bank maintains cash reserves, R_H and R_F, to allow customers to make and receive payments and use other associated services. The interest rate paid on such reserves are given by r^H and r^F for the home and foreign market. The bank fully finances their loans and cash reserves through deposits D from the home market, paying interest r_D, and equity E.

The foreign loans are provided in a foreign currency and thus their value is subject to changes in the exchange rate e. The total exposure of the bank to the foreign currency is thus $\pi_F \left(1 + r_L^F\right) L_F + \left(1 + r^F\right) R_F$ and we assume that the bank hedges an amount $\overline{F}$ of this exposure at the current forward rate $\overline{e}$ and the remaining exposure F is left unhedged. However, the bank expects the exchange rate to remain at its current level e. We thus have for the foreign exposure that

$$\pi_F \left(1 + r_L^F\right) L_F + \left(1 + r^F\right) R_F = \overline{e}\overline{F} + eF. \tag{39.50}$$

When providing loans, the liabilities of the bank consist of deposits and equity, with assets being the home and foreign loans as we all as the reserves held in both countries. Thus we require that

$$D + E = L_H + eL_F + R_H + eR_F. \tag{39.51}$$

Finally, banks are subject to minimum capital requirements by regulators applying a capital ratio κ such that the amount of equity must be at least a fraction κ of the loans they have provided in the home and foreign markets. An additional capital requirement is applied for the unhedged exposure to currency risk at a capital ratio $\hat{\kappa}$. The capital requirements are thus

$$E = \kappa \left(L_H + eL_F\right) + \hat{\kappa} eF. \tag{39.52}$$

The profits of the bank now consist of the repaid domestic loans and cash reserves as well as their foreign counterparts; from these positions the bank needs to repay its depositors. Thus we obtain the bank profits as

$$\Pi_B = \pi_H \left(1 + r_L^H\right) L_H + \left(1 + r^H\right) R_H + \overline{e}\overline{F} + eF - \left(1 + r_D\right) D. \tag{39.53}$$

We additionally assume that as banks provide more loans in the home or foreign markets, the probability of the loans being repaid reduces as the amount of lending increases. While a larger loan amount will increase the amount that is repaid, the higher risk makes this repayment increase less than the loan amount. More formally if we define $\Psi_j = \frac{\pi_j\left(1+r_L^j\right)L_j}{\partial L_j} > 0$ and $\frac{\partial \Psi_j}{\partial L_j} < 0$.

Banks will seek to maximize their profits by choosing the optimal loan amounts for home and foreign loans, home and foreign cash reserves, as well as the amount of hedged and unhedged foreign exposure and the optimal amount of deposits. When conducting this maximization, the constraints in equations (39.50) to (39.52) need to be considered. With ξ_i denoting the corresponding Lagrange multipliers, we get the first order conditions as

$$
\begin{aligned}
\frac{\partial \Pi_B}{\partial L_H} &= \Psi_H - \xi_2 - \kappa \xi_3 = 0, \\
\frac{\partial \Pi_B}{\partial L_F} &= -e\xi_2 - e\kappa\xi_3 + \Psi_F \xi_1 = 0, \\
\frac{\partial \Pi_B}{\partial R_H} &= 1 + r^H - \xi_2 = 0, \\
\frac{\partial \Pi_B}{\partial R_F} &= -e\xi_2 + \left(1 + r^F\right)\xi_1 = 0, \\
\frac{\partial \Pi_B}{\partial D} &= -\left(1 + r_D\right) + \xi_2 = 0, \\
\frac{\partial \Pi_B}{\partial \overline{F}} &= \overline{e} - \xi_1 = 0, \\
\frac{\partial \Pi_B}{\partial F} &= e - \hat{\kappa}\xi_3 - \xi_1 = 0.
\end{aligned}
\tag{39.54}
$$

Combining the third and fifth condition, we obtain

$$
\xi_2 = 1 + r^H = 1 + r_D \tag{39.55}
$$

and combining the first and third condition yields

$$
\Psi_H = 1 + r_D + \kappa\xi_3. \tag{39.56}
$$

If now use conditions 2, 3 and 6, we obtain

$$
\Psi_F \frac{\overline{e}}{e} = 1 + r^H + \kappa\xi_3 \tag{39.57}
$$

and from conditions 3, 4, and 6 it is

$$
1 + r^H = \frac{\overline{e}}{e}\left(1 + r^F\right). \tag{39.58}
$$

This result recovers the covered interest rate parity. Using in turn equations (39.57) and (39.58) we can show that

$$
\Psi_F = 1 + r^F + \frac{e}{\overline{e}}\kappa\xi_3 = 1 + r^F + \frac{1 + r^F}{1 + r^H}\kappa\xi_3. \tag{39.59}
$$

Using this last result we can easily see that

$$
\frac{\partial \Psi_F}{\partial\left(1 + r^H\right)} = \frac{\partial \Psi_F}{\partial L_F}\frac{\partial L_F}{\partial\left(1 + r^H\right)} = -\frac{1 + r^F}{\left(1 + r^H\right)^2}\kappa\xi_3 \tag{39.60}
$$

and hence

$$
\frac{\partial L_F}{\partial\left(1 + r^H\right)} = -\frac{1 + r^F}{\frac{\partial \Psi_F}{\partial L_F}\left(1 + r^H\right)^2}\kappa\xi_3 \geq 0, \tag{39.61}
$$

where the positive sign arises from our assumption that $\frac{\partial \Psi_F}{\partial L_F} < 0$. The inequality is strict if $\xi_3 > 0$, which can be interpreted as the capital requirements being binding. Thus, if capital requirements are binding, an increase in the home interest rate will make it optimal for the bank to provide more loans in the foreign market. This result rests on the assumption that the marginal product Ψ_F remains unchanged by this interest rate change, which is reasonable to assume given these companies are operating in different countries. The effect of higher interest rates in a country is therefore not only that the lending in that country reduces, this arises due to the constraint in equation (39.51), but it will also expand the lending in the other country, provided it has not increased interest rates, too.

We can now combine equations (39.55) to (39.58) and obtain

$$\Psi_H = \frac{1 + r^H}{1 + r^F} \Psi_F. \tag{39.62}$$

If the marginal repayments of companies are not changing as interest rates change, thus Ψ_j remains constant, we can rewrite this expression as

$$\ln \Psi_H - \ln \Psi_F \approx -\left(r^H - r^F\right). \tag{39.63}$$

We thus observe that as the home interest rate increases, the differences in the log marginal repayment rates reduce, which with equation (39.61) implies that foreign loans increase and domestic loans reduce.

We have therefore established that the monetary policy in one country, here a change in the interest rate, will not only affect the provision of loans in that country, but will have implications for the provision of loans in other countries. This spillover of loan provisions through internationally operating banks will have to be considered by policymakers in other countries; if they would want to prevent more loans being provided, for example to avoid an overheating of the economy, they might want to raise their own interest rates to counter this effect. On the other hand, if they seek to stimulate economic growth and would like loan provision to increase, they might not have to lower their interest rates as much, but will be supported by the rising interest rates in another country.

Readings Bräuning & Ivashina (2020), Cao (2022, Ch. 9.3)

Conclusions

We have seen that monetary policy does not only directly affect the costs of borrowing for companies through its effect on loan rates, but also has implications for the incentives of banks; such incentives need to be considered when conducting monetary policy. Not only do lending constraints, for example arising from the limited availability of collateral or regulatory limits, restrict the amount that can be lent, but they make the impact of interest rate changes more pronounced as the higher value of future repayments makes any constraints more binding and will therefore reduce (increase) lending more than the interest rate increase (decrease) would suggest. However, higher interest rates are also providing banks with incentives to provide higher risk loans. The reason for this is that while deposit rates increase in line with central bank policy as depositors have alternative investment opportunities, the demand for loans will reduce if loan rates are increased. This makes passing on interest rate increases more difficult, reducing the profits of banks. Banks will seek to compensate for that by reducing the monitoring of loans and thereby making them more risky. Hence higher interest rates are associated with higher risks for banks.

If companies are affected by adverse shocks to their profitability, they become less attractive for banks to be granted loans as the loan rate that can be charged will have to be reduced. Lowering interest rates can support banks in continuing to provide such loans as this reduces deposit rates, maintaining the profitability of banks. Large shocks would need a subsidy to banks, such as potential bailouts if loans fail, as that would reduce the deposit rates even further due to the lower risks to depositors. Too high shocks would not allow for sufficient subsidies as lending to these companies is not sufficiently attractive and no subsidy or monetary policy can induce banks to grant these loans.

Interest rates cannot only be affected by monetary policy tools, such as the setting of interest rate setting and liquidity provisions by the central bank, but also through liquidity requirements as this determines the degree to which banks can rely on interbank markets to provide additional liquidity. Central banks providing additional liquidity to banks will not necessarily reduce the interest rates if it is anticipated that any liquidity provided will be withdrawn in the near future; it is therefore that short-term liquidity injections may not only be without much impact on interest rates, it may actually have the opposite of the intended effect. We have also seen that interest rate changes in one country can have an impact on the provision of loans in another country, making the international coordination of monetary policy decisions beneficial.

Banks play not only an important role in transmitting central bank decisions on monetary policy, but are an active part of the economy. The models discussed here, however, show that monetary policy involves complex interactions between monetary policy on the one hand and the decision by banks on the other hand. It is therefore for any monetary policy decision to take into account these incentives for banks to balance the intended effect of the policy decision against any secondary effects that might reduce the welfare.

Chapter 40

Economic fluctuations

Of course, banks are not only affected by monetary policy, but also by market-wide expectations and sentiments. Chapter 40.1 will show that these have an impact on the capital ratio banks operate with, and thus indirectly the amount of lending banks are willing to give. To this effect we do not model the decision of individual banks, but take an aggregate view of how much lending the market overall can support. In the same way, chapter Chapter 40.2 will provide a reasoning for such leverage changes to be moving in a similar way to the business cycle of the economy.

Banks provide loans not only for investment by companies, but also for consumption. In doing so, chapter 40.3 shows that banks can eliminate consumption risk to consumers. But as they do so, they have to react to the demands of consumers and will do so by varying the amount of loans available for investment, causing investments to fluctuate. If the success of investments is sufficiently sensitive to the amount of investment in an economy, such as the amount of competition it generates, then a steady-state equilibrium, while it exists, can never be reached and investment fluctuations continue for ever. We thus see that banks are not passive in that they react to macroeconomic decisions, but they actively influence macroeconomic conditions through their decision-making.

40.1 Bank leverage

Let us assume there are two types of borrowers, the investment of the first type succeeds with probability π_H and being charged a loan rate of r_L^H, and the other has a success rate of $\pi_L < \pi_H$ and play loan rate r_L^L. The expected return to a lender from the low risk borrower is higher than that of a high risk borrower with $\pi_H \left(1 + r_L^H\right) \geq 1 \geq \pi_L \left(1 + r_L^L\right)$, making the high-risk borrower undesirable as it does not cover its initial investment. The type of borrower is unknown, we have a probability p that a borrower is of type H. Moreover, the value of p is different for across individuals and we assume it to be uniformly distributed in $[0; 1]$.

We are here concerned with the amount of lending banks can connduct in such an economy.

Optimal capital ratio The value of a loan, for a given belief p is then given by

$$P = p\pi_H \left(1 + r_L^H\right) + (1 - p)\,\pi_L \left(1 + r_L^L\right). \tag{40.1}$$

Those individuals that are optimistic about the composition of the borrowers, that is if they belief that $p \geq \overline{p}$, for some threshold $\overline{p}$, will use their initial wealth as equity in a bank to lend. Given the uniform distribution of p the fraction of individuals acting as banks will be $1 - \overline{p}$ and with aggregate wealth K we get total funds available to the bank of $(1 - \overline{p})\,K + D$, where D are deposits collected by the bank. Those less optimistic would prefer to borrow money to consume now rather than lend. The wealth of these individuals is $\overline{p}K$ and the price they pay for each loan is given by P such they can raise $\overline{p}KP$. The amount of loans given has to equal those demanded, thus $(1 - \overline{p})\,K + D = \overline{p}KP$ for market clearing. We assume that even a risky loan does allow the bank to repay its depositors. With the loan amount L, we have $\pi_L \left(1 + r_L^L\right) L \geq D$ and the bank will lend as much as possible to make this equation an equality. We have that loans can be financed by deposits and equity, thus $L = K + D$ and we use a capital ratio κ with $K = \kappa D$, such that $L = (1 + \kappa)\,D = \frac{1+\kappa}{\kappa}K$. Inserting all these relationships, we obtain from market clearing that the price of the loan must be

$$P = \frac{1 - \overline{p} + \pi_L \left(1 + r_L^L\right) \frac{1+\kappa}{\kappa}}{\overline{p}}. \tag{40.2}$$

For the marginal individual with $p = \overline{p}$, this expression has to equal the value of the loan. Inserting $\overline{p}$ for p in equation (40.1) and setting equal to equation (40.2), gives us

$$\begin{aligned} \mathcal{L} = {} & \overline{p}^2 \left(\pi_H \left(1 + r_L^H\right) - \pi_L \left(1 + r_L^L\right)\right) \\ & + \overline{p} \left(1 + \pi_L \left(1 + r_L^L\right)\right) - \left(1 + \pi_L \left(1 + r_L^L\right) \frac{1 + \kappa}{\kappa}\right) \\ = {} & 0. \end{aligned} \tag{40.3}$$

For the leverage of a bank, $\frac{L}{K}$, we have that the value of the assets is PL and the equity $PL - D$, thus after inserting that $D = \pi_L \left(1 + r_L^L\right) L$, we get

$$\frac{1 + \kappa}{\kappa} = \frac{P}{P - \pi_L \left(1 + r_L^L\right)}. \tag{40.4}$$

We can solve this for κ and after inserting for P from equation (40.1), we get

$$\kappa = \overline{p} \left(\frac{\pi_H \left(1 + r_H^H\right)}{\pi_L \left(1 + r_L^L\right)} - 1\right). \tag{40.5}$$

We can now solve this expression for $\overline{p}$ and insert back into equation (40.6), which then becomes

$$\begin{aligned}\mathcal{L} = & \frac{\kappa^2\pi_L^2\left(1+r_L^L\right)^2}{\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)} \\ & +\frac{\kappa\pi_L\left(1+r_L^L\right)\left(1+\pi_L\left(1+r_L^L\right)\right)}{\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)} \\ & -\left(1+\pi_L\left(1+r_L^L\right)\frac{1+\kappa}{\kappa}\right) \\ = & \, 0.\end{aligned} \tag{40.6}$$

We can use this expression to derive the equilibrium capital ratio κ of the banks.

Changes in expected returns We now investigate the impact changes of the expected returns have. Firstly we consider that $\pi_H\left(1+r_L^H\right)$ increases, i.e. individuals are more optimistic about the low risk borrowers only. We can easily obtain that

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\pi_H\left(1+r_L^H\right)} = & -\frac{\kappa^2\pi_L^2\left(1+r_L^L\right)^2}{\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)^2} \\ & -\frac{\kappa\pi_L\left(1+r_L^L\right)\left(1+\pi_L\left(1+r_L^L\right)\right)}{\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)^2} < 0, \\ \frac{\partial\mathcal{L}}{\partial\pi_L\left(1+r_L^L\right)} = & \frac{\pi_L^2\left(1+r_L^L\right)^2}{\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)^2}\kappa^2 \\ & +\frac{2\pi_L\left(1+r_L^L\right)\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)}{\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)^2}\kappa^2 \\ & +\frac{\left(1+2\pi_L\left(1+r_L^L\right)\right)\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)}{\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)^2}\kappa \\ & +\frac{\pi_L\left(1+r_L^L\right)\left(1+\pi_L\left(1+r_L^L\right)\right)}{\left(\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)\right)^2}\kappa \\ & -\frac{1+\kappa}{\kappa} \lesseqgtr 0, \\ \frac{\partial\mathcal{L}}{\partial\kappa} = & \frac{2\kappa\pi_L^2\left(1+r_L^L\right)^2}{\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)} \\ & +\frac{\pi_L\left(1+r_L^L\right)\left(1+\pi_L\left(1+r_L^L\right)\right)}{\pi_H\left(1+r_L^H\right)-\pi_L\left(1+r_L^L\right)} \\ & +\pi_L\left(1+r_L^L\right)\frac{1}{\kappa^2} > 0.\end{aligned}$$

The signs of the first and final expressions are obvious and the sign of the second expression can be either positive or negative. For very small values of κ, implying very highly leveraged banks, this expression would be negative, but if the expected returns of the two types of borrowers are not too different, the final term, $\frac{1+\kappa}{\kappa}$, will not turn this expression negative. Hence we work with the assumption that $\frac{\partial \mathcal{L}}{\partial \pi_L \left(1+r_L^L\right)} > 0$.

The implicit function theorem gives us then that

$$\frac{\partial \kappa}{\partial \pi_H \left(1+r_L^H\right)} = -\frac{\frac{\partial \mathcal{L}}{\partial \pi_H \left(1+r_L^H\right)}}{\frac{\partial \mathcal{L}}{\partial \kappa}} > 0, \tag{40.7}$$

$$\frac{\partial \kappa}{\partial \pi_L \left(1+r_L^L\right)} = -\frac{\frac{\partial \mathcal{L}}{\partial \pi_L \left(1+r_L^L\right)}}{\frac{\partial \mathcal{L}}{\partial \kappa}} < 0.$$

Here we see that with more positive expectations about low-risk borrowers, the capital ratio is increasing. The reason is that with higher expected returns from loans to these borrowers, lending becomes more attractive and hence more individuals decide to become bankers, increasing the capital available to banks. This also decreases the demand for lending, it becomes more expensive to borrow and there are less borrowers available as many have become bankers. On the other hand, higher expected returns by high-risk borrowers decrease the capital ratio. Here the constraint that $\pi_L \left(1+r_L^L\right) L \geq D$ drives the results. With $\pi_L \left(1+r_L^L\right) \leq 1$, this constraint limits the amount of lending that can be conducted and increasing the expected returns will make this constraint less binding, allowing lending to grow. With increased lending, the capital ratio will fall. This is effect is stronger than the attractiveness of becoming a banker increasing as outlined for low-risk borrowers. Only once the leverage is already high, will this effect dominate and a further increase in the expected returns will result in lower capital ratios.

Summary If the market is optimistic about the returns of low risk borrowers, the capital ratio will be high, thus we would expect banks to exhibit a low leverage in times of strong economic growth and the opposite during times of recessions. The opposite would be the case for high-risk borrowers, here being positive about their prospects would result in higher leverage as lending constraints become less binding. We might thus see a reduction in lending during times of high economic growth is banks lend mainly to low-risk borrowers, but an expansion of loans if they lend to high-risk borrowers.

Reading Geanakoplos (2010) and Cao (2022, Chapter 8.6)

40.2 Procyclical leverage

Without any restrictions, banks will provide loans as is optimal for their profits and it is often observed that banks readily provide loans during times an economy is performing well, but are much more reluctant to do so during recessions. It is thus that loans are more widely available in times of high economic growth and less available during recessions, which could easily amplify the economy cycle. We will analyse here how such procyclical behaviour of banks can be explained.

Let us assume a company can finance its investment either through a bank loan L or issue a bond B. If the company decides to issue a bond, its profits are given by

$$\Pi_C = I(B) - (1 + r_B) B, \tag{40.8}$$

where $I(B)$ denotes the outcome the company can produce when making an investment using the bond B; from this outcome the bond is repaid, on which interest r_B is charged. Companies will seek to maximize their profits by choosing the optimal amount of lending and the first order condition $\frac{\partial \Pi_C}{\partial B} = 0$ easily solves for

$$\frac{\partial I(B)}{\partial B} = 1 + r_B. \tag{40.9}$$

This expression shows that the bond rate is determined by the marginal productivity of the company. We assume here that a larger amount of borrowing increases the outcome the company obtains from its investment, $\frac{\partial I(B)}{\partial B} > 0$, but that this increase is diminishing, $\frac{\partial^2 I(B)}{\partial B^2} < 0$. If banks compete with bonds for depositors, they need to offer the same deposit rate as the bonds, thus we find that $r_D = r_B$.

A bank will charge a loan rate r_L on its loan of size L and finance these loans with deposits D and equity E such that $L = D + E$. We assume that banks have additional costs of c when providing loans to account for the increased monitoring of loans and improved due diligence assessments of borrowers. Banks will maximize their profits from lending, $\Pi_B = (1 + r_L) L - (1 + r_D) D$, ignoring these additional costs in their maximization, but will provide loans only if they can recover their monitoring costs and hence we require that $\Pi_B \geq cL$. Maximizing the profits of banks over the optimal amount of lending, subject to this constraint, we obtain the first order condition as

$$\frac{\partial \mathcal{L}}{\partial L} = r_l - r_D + \xi (r_L - r_D - c) = 0, \tag{40.10}$$

where ξ represents the Lagrange multiplier and we used that $D = L - E$. This expression then solves for

$$r_L - r_D = \frac{\lambda}{1 + \lambda} c. \tag{40.11}$$

Combining this result with the binding condition that $\Pi_B = cL$, we finally obtain that

$$\frac{L}{E} = \frac{\lambda}{1 + \lambda} \frac{1}{E} + \frac{1 + r_D}{c}. \tag{40.12}$$

We now see that banks will have a higher leverage, $\frac{L}{E}$ if the deposit rate $1 + r_D$ is high. Having established above that the deposit rate is equal to the bond rate and that this was determined by the marginal productivity of the company, we can conclude that in times of high productivity of companies the leverage of banks will be higher. Such higher productivity we typically observe when the economy is performing well, but productivity is low during recessions or times of economic stress. It is therefore that banks have a high leverage if the economy is performing well and a lower leverage if the economy is not performing well.

The conclusion from this model is that banks will provide loans in a procyclical manner. A large amount of loans is given during times of high economic growth and loans will be more difficult to obtain during recessions as banks reduce leverage. Such a behaviour of banks might be not desirable from an economic policy point of view as the provision of loans during times an economy performs well would increase growth even more due to increased investments, leading to an overheating economy; the reduction in loan provision during recessions might prolong or deepen them. For this reason a regulator may want to tighten the capital requirements of banks during economic expansions and limit the leverage they can take and thereby support the central bank in its efforts to reduce economic growth to a more sustainable level. In contrast to this in times of recessions, capital requirements might be loosened. Changing capital requirements over the business cycle is often referred to as macroprudential regulation is its aim is to manage the risks of banks but also support general macroeconomic policies.

Reading Gersbach & Rochet (2017), Freixas & Rochet (2023, Ch. 9.4.3)

40.3 Cyclical investments

Banks provide consumers with an alternative investment to providing money to companies and bearing the risks of such investments, deposits. Similarly, companies are provided with an alternative funding source to monies obtained from consumers, namely loans. Let us, however, aggregate consumers and companies; we can do this by assuming that consumers hold the shares of companies and thus obtain any profits they make. Furthermore, each consumer-company considers two time periods, t and $t+1$, in which they can consume. Consumption at t is financed by their wages, w_t, any loans that are made by banks, L_t, and reduced by the investment that will be conducted at $t+1$, I_{t+1}, and any money retained as deposits in the bank, D_t. Thus

$$C_t = w_t + L_t - I_{t+1} - D_t. \tag{40.13}$$

The investment is successful with probability π and then generates a return R. With deposit rate r_D and loan rate r_L, consumption in this state is then after the loan has been repaid

$$C_{t+1}^H = (1 + r_D)\, D_t + (1 + R)\, I_{t+1} - (1 + r_L)\, L_t. \tag{40.14}$$

If the investment is not successful, we assume that it yields no funds and due to limited liability of the company, the loans also do not need to be repaid, such that

$$C_{t+1}^{L} = (1 + r_D)\, D_t. \tag{40.15}$$

Individuals, as owners of the company are not liable for any outstanding loans due to limited liability and can thus retain their deposits.

We assume wages are only paid in advance to production at time t. As, however, at $t + 1$ another investment is launched by another consumer-company, wages are paid in every time period. It is only that consumer-companies, have a life span of two time periods.

The value of the total production of the company, V, can now the divided up into the profits the company makes, $\frac{\partial V}{\partial \hat{I}_t}\hat{I}_t$, and the wages to be paid, w_t, where $\hat{I}_t = \pi I_t$ denotes the fraction of successful investments only. Hence

$$V = \frac{\partial V}{\partial \hat{I}_t}\hat{I}_t + w_t, \tag{40.16}$$

where $\frac{\partial V}{\partial \hat{I}_t} > 0$ and $\frac{\partial^2 V}{\partial \hat{I}_t^2} < 0$. The total production V is paid out in wages w_t and profits. These profits are given by the marginal product $\frac{\partial V}{\partial \hat{I}_t}$ of the successful capital investment and the successful investment itself. Hence any excess returns beyond the marginal product is paid out as wages.

We furthermore assume that $\frac{\partial \pi}{\partial I_{t+1}} < 0$ and $\frac{\partial^2 \pi}{\partial I_{t+1}^2} > 0$, hence the more is invested the lower the success rate as the company has to seek more and more risky investments, with the increase in risk reducing as investment becomes larger. The utility of consumption is given by $u(\cdot)$ and we assume as is common that $\frac{\partial u}{\partial C_i} > 0$ and $\frac{\partial^2 u}{\partial C_i^2} < 0$.

Finally we define for convenience

$$\eta = \frac{\partial \pi}{\partial I_{t+1}} \frac{I_{t+1}}{\pi} \tag{40.17}$$

as the elasticity of the success rate with $-1 \leq \eta \leq 0$ and $\frac{\partial \eta}{\partial I_{t+1}} = \frac{\partial^2 \pi}{\partial I_{t+1}^2}\frac{I_{t+1}}{\pi} + \frac{\partial \pi}{\partial I_{t+1}}\frac{1}{\pi}$ being negative for small investments and positive for large investments.

The total utility of a consumer is then

$$U_C = u(C_t) + \pi u\left(C_{t+1}^{H}\right) + (1 - \pi)\, u\left(C_{t+1}^{L}\right), \tag{40.18}$$

with C_t, C_{t+1}^{H}, and C_{t+1}^{L} given by equations (40.13)-(40.15). The utility consists of the consumption in time period t and time period $t+1$, where its level will depend on whether the investment of the company is successful or not. We neglect discounting between the two time periods for simplicity.

In this framework, we firstly evaluate the equilibrium investment by companies in the absence of banks as a benchmark, before introducing banks to see the impact their presence has.

Absence of banks If there are no banks, then neither loans nor deposits are available, thus $L_t = D_t = 0$, implying from equations (40.13)-(40.15) that

$$\begin{aligned} C_t &= w_t - I_{t+1}, \\ C_{t+1}^H &= (1+R)\, I_{t+1}, \\ C_{t+1}^L &= 0. \end{aligned} \tag{40.19}$$

If, for simplicity, we normalize $u(0) = 0$, then from equations (40.18) and (40.19) we have

$$U_C = u(C_t) + \pi u\left(C_{t+1}^H\right) \tag{40.20}$$

and the first order condition for optimal investment becomes

$$\begin{aligned} \frac{\partial U_C}{\partial I_{t+1}} &= \frac{\partial u(C_t)}{\partial C_t}\frac{\partial C_t}{\partial I_{t+1}} + \pi \frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H}\frac{\partial C_{t+1}^H}{\partial I_{t+1}} \\ &\quad + \frac{\partial \pi}{\partial I_{t+1}} u\left(C_{t+1}^H\right) + (1-\pi)\frac{\partial u\left(C_{t+1}^L\right)}{\partial C_{t+1}^L}\frac{\partial C_{t+1}^L}{\partial I_{t+1}} \\ &\quad - \frac{\partial \pi}{\partial I_{t+1}} u\left(C_{t+1}^L\right) \\ &= -\frac{\partial u(C_t)}{\partial C_t} + \pi(1+R)\frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H} + \frac{\partial \pi}{\partial I_{t+1}} u\left(C_{t+1}^H\right) \\ &= 0, \end{aligned} \tag{40.21}$$

where we used equation (40.19) to obtain the derivatives of consumption with respect to investments and noted that $u(0) = 0$. In order to obtain the dynamics of investment in this optimum we can now differentiate both sides with respect to I_t and get

$$\begin{aligned} \frac{\partial^2 U_C}{\partial I_{t+1}\partial I_t} &= -\frac{\partial^2 u(C_t)}{\partial C_t^2}\left(\frac{\partial w_t}{\partial I_t} - \frac{\partial I_{t+1}}{\partial I_t}\right) \\ &\quad + (1+R)\frac{\partial \pi}{\partial I_{t+1}}\frac{\partial I_{t+1}}{\partial I_t}\frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H} \\ &\quad + \pi(1+R)^2 \frac{\partial^2 u\left(C_{t+1}^H\right)}{\partial\left(C_{t+1}^H\right)^2}\frac{\partial I_{t+1}}{\partial I_t} \\ &\quad + \frac{\partial^2 \pi}{\partial I_{t+1}^2}\frac{\partial I_{t+1}}{\partial I_t} u\left(C_{t+1}^H\right) \\ &\quad + (1+R)\frac{\partial \pi}{\partial I_{t+1}}\frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H}\frac{\partial I_{t+1}}{\partial I_t} \\ &= 0. \end{aligned} \tag{40.22}$$

This expression can be solved for

$$\frac{\partial I_{t+1}}{\partial I_t} = \frac{\frac{\partial^2 u(C_t)}{\partial C_t^2}\frac{\partial w_t}{\partial I_t}}{\Psi} > 0, \tag{40.23}$$

where $\Psi = \frac{\partial^2 u(C_t)}{\partial C_t^2} + 2(1+R)\frac{\partial u(C_{t+1}^H)}{\partial C_{t+1}^H}\frac{\partial \pi}{\partial I_{t+1}} + \pi(1+R)^2\frac{\partial^2 u(C_{t+1}^H)}{\partial C_{t+1}^H} + \frac{\partial^2 \pi}{\partial I_{t+1}^2}u\left(C_{t+1}^H\right)$. Solving equation (40.16) for w_t and differentiating, we obtain

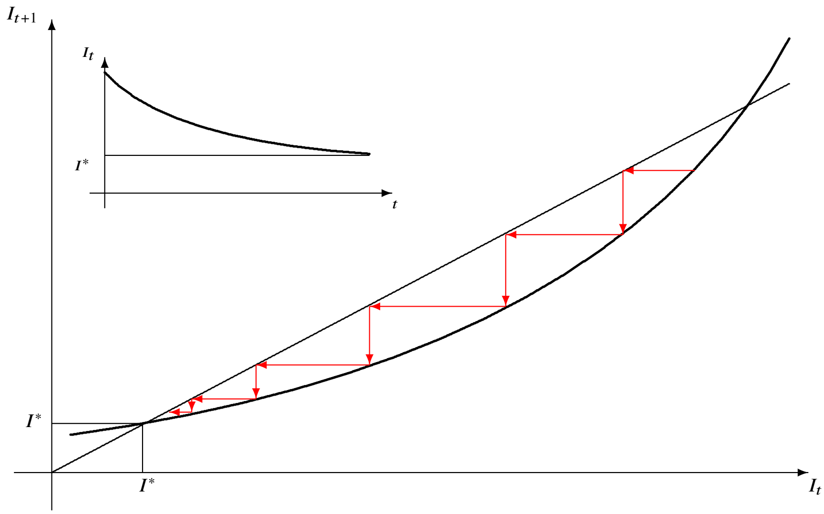
$$\frac{\partial w_t}{\partial I_t} = \pi\frac{\partial V}{\partial \hat{I}_t} - \pi\frac{\partial^2 V}{\partial \hat{I}_t^2}\hat{I}_t > 0, \tag{40.24}$$

due to $\frac{\partial V(\hat{I}_t)}{\partial \hat{I}_t} > 0$, $\frac{\partial^2 V(\hat{I}_t)}{\partial \hat{I}_t^2} < 0$ and noting that $\hat{I}_t = \pi I_t$. As $\frac{\partial^2 u(C_t)}{\partial C_t^2} < 0$ the numerator in equation (40.23) is negative. The denominator Ψ is also negative, because the first term is negative by assumption on the utility function, the second term is negative as we assumed $\frac{\partial \pi}{\partial I_{t+1}} < 0$ and marginal utility is positive, the third term again arises from the properties of the utility function to be negative, and the final term is negative as $\frac{\partial^2 \pi}{\partial I_{t+1}^2} < 0$ by assumption and given that $C_{t+1}^H \geq 0$, we have $u\left(C_{t+1}^H\right) \geq u(0) = 0$. Therefore, $\frac{\partial I_{t+1}}{\partial I_t} > 0$.

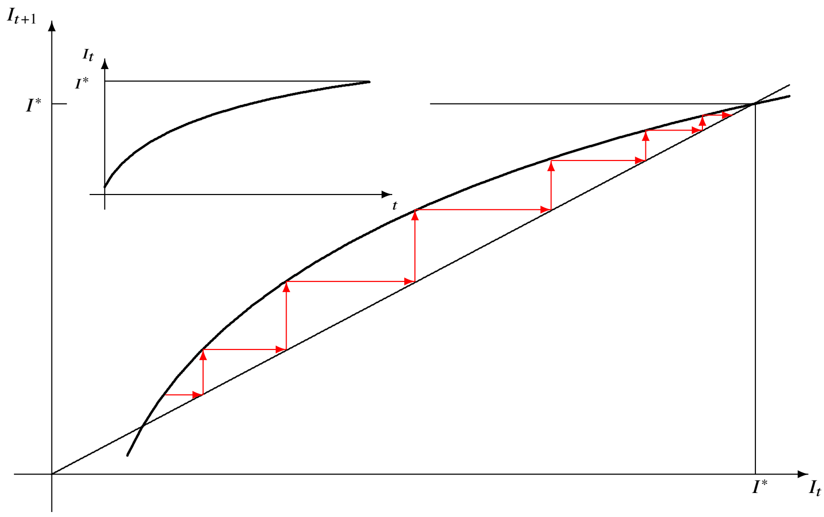
Thus there is a positive relationship between investments in time periods t and $t+1$, an increase in investments in time period t will induce an increase in time period $t+1$. This is because a larger amount of wages is paid out in time period t, that can then be invested in the coming time period. Figure 40.1 shows two possible shapes of this relationship. The yop panel covers the case where $\frac{\partial^2 I_{t+1}}{\partial I_t^2} > 0$ while the lower panel shows the results for $\frac{\partial^2 I_{t+1}}{\partial I_t^2} < 0$. Which shape applies will depend on the properties of the utility function $u(\cdot)$, the investment returns R and the probability of investment success π. We clearly see that starting with an initial investment, I_0, the investment will either be constantly increasing or decreasing, in both cases eventually converging to a steady state I^* where $I_t = I_{t+1}$, representing the equilibrium steady state. The inset figures show the evolution of the investment level over time and we see that the investment levels slowly approach the steady state. The location of the steady state will depend on the slope the relationship between I_t and I_{t+1}, if it is increasing, the lower point will be the steady sate while for decreasing slopes, it will be the higher level of investment. As investment is risky and consumption might be reduced to nil in time period $t+1$ if the investment fails, a higher degree of risk aversion by consumer-companies would lead to lower levels of investment.

More complex solutions are possible if $\frac{\partial I_{t+1}}{\partial I_t}$ has non-monotonous changes to such that it crosses the line $I_t = I_{t+1}$ multiple times. In this case several steady states can exist and which one is obtained would depend on the starting point I_0. In the case that the line $I_t = I_{t+1}$, is only crossed once and that crossing is not a steady state, then the dynamics would imply a run away to infinite investment or no investment at all. Suitable constraints on parameters can avoid this situation.

We see that without banks, an equilibrium level of investment is reached monotonically over time. Depending on the risk aversion of consumer-companies the optimal



(a) Dynamics with $\frac{\partial^2 I_{t+1}}{\partial I_t^2} > 0$



(b) Dynamics with $\frac{\partial^2 I_{t+1}}{\partial I_t^2} < 0$

Fig. 40.1: Equilibrium dynamics without banks

level of investment will be either high or low. We can now introduce banks and see how this affects the investment decision.

The influence of banks If we introduce banks into this model, we know that in the absence of equity and cash reserves loans L_t and deposits D_t need to be identical and hence the bank profits are given by

$$\Pi_B = \pi\left(1 + r_L\right) L_t - \left(1 + r_D\right) D_t = \left(\pi\left(1 + r_L\right) - \left(1 + r_D\right)\right) L_t, \tag{40.25}$$

assuming that banks have other resources to repay deposits.

To make banks viable we need $\Pi_B \geq 0$. Hence maximizing equation (40.18) subject to equation (40.25), we get the Lagrangian function as

$$\mathcal{L} = U_C + \lambda\left(\pi\left(1 + r_L\right) - \left(1 + r_D\right)\right), \tag{40.26}$$

where λ denotes the Lagrange multiplier. The first order conditions for its maximum are then given by

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial D_t} &= -\frac{\partial u\left(C_t\right)}{\partial C_t + \pi\left(1 + r_D\right)} \frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H} && (40.27)\\
&\quad + \left(1 - \pi\right)\left(1 + r_D\right) \frac{\partial u\left(C_{t+1}^L\right)}{\partial C_{t+1}^L} = 0, \\
\frac{\partial \mathcal{L}}{\partial I_{t+1}} &= -\frac{\partial u\left(C_t\right)}{\partial C_t} + \frac{\partial \pi}{\partial I_{t+1}} u\left(C_{t+1}^L\right) && (40.28)\\
&\quad + \pi\left(1 + R\right) \frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H} - \frac{\partial \pi}{\partial I_{t+1}} u\left(C_{t+1}^L\right) \\
&\quad + \lambda\left(1 + r_L\right) \frac{\partial \pi}{\partial I_{t+1}} = 0, \\
\frac{\partial \mathcal{L}}{\partial L_t} &= \frac{\partial u\left(C_t\right)}{\partial C_t} - \pi\left(1 + r_L\right) \frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H} = 0, && (40.29)\\
\frac{\partial \mathcal{L}}{\partial\left(1 + r_D\right)} &= \pi L_t \frac{\partial u\left(C_{t+1}^H\right)}{\partial C_{t+1}^H} - \lambda \pi = 0. && (40.30)
\end{aligned}$$

Solving equation (40.29) for $\frac{\partial u(C_t)}{\partial C_t}$ and inserting into equation (40.27), we get

$$\begin{aligned}
&\pi\left(\left(1 + r_D\right) - \left(1 - r_L\right)\right) \frac{\partial u\left(C_{t+1}\right)}{\partial C_{t+1}^H} && (40.31)\\
&\quad + \left(1 - \pi\right)\left(1 + r_D\right) \frac{\partial u\left(C_{t+1}^L\right)}{\partial C_{t+1}^L} = 0.
\end{aligned}$$

As $L_t \geq 0$ and $\frac{\partial u(C_{t+1}^H)}{\partial C_{t+1}^H} > 0$, we see in equation (40.30) that $\lambda > 0$ and hence the constraint is binding such that $1 + r_D = \pi\left(1 + r_L\right)$. Inserting this into equation (40.31), we easily obtain that

$$\frac{\partial u\left(C_{t+1}^{H}\right)}{\partial C_{t+1}^{H}}=\frac{\partial u\left(C_{t+1}^{L}\right)}{\partial C_{t+1}^{L}} \tag{40.32}$$

implying that $C_{t+1}^{H} = C_{t+1}^{L} = (1 + r_D) D_t$ as the marginal utilities are identical. Hence banks provide consumer-companies with insurance against the uncertainty of investments. While in the absence of banks, consumer-companies were exposed to the risk of investments, this is not the case in the presence of banks.

From equations (40.14) and (40.15) we then get by setting them equal that

$$L_t = \frac{1+R}{1+r_L} I_{t+1}. \tag{40.33}$$

If we use that $C_{t+1}^{H} = C_{t+1}^{L}$, in equation (40.28) and then insert from equation (40.29) for $\frac{\partial u(C_t)}{\partial C_t}$ and from equation (40.30) for λ, we get after simplifying that

$$-\pi\left(1+r_L\right)+\pi\left(1+R\right)+\left(1+r_L\right) L_t \frac{\partial \pi}{\partial I_{t+1}}=0. \tag{40.34}$$

Inserting from equation (40.33) and using the definition in equation (40.17), this can be solved for

$$1+r_L=\left(1+R\right)\left(1+\eta\right). \tag{40.35}$$

As we assumed that $-1 \leq \eta \leq 0$ and from equation (40.33) we obtain that

$$L_t = \frac{I_{t+1}}{1+\eta} > I_{t+1}. \tag{40.36}$$

Hence consumer-companies would obtain a loan larger than the investment they seek to make. As investors are risk averse, they will not invest the full amount but retain some fraction as a safe deposit. This strategy allows for certain consumption in time period $t + 1$.

Using $C_{t+1}^{H} = C_{t+1}^{L} = (1 + r_D) D_t$ from equation (40.15) in equation (40.27), we obtain

$$\frac{\partial \mathcal{L}}{\partial D_t}=-\frac{\partial u\left(C_t\right)}{\partial C_t}+\left(1+r_D\right)\frac{\partial u\left(C_{t+1}^{L}\right)}{\partial C_{t+1}^{L}}=0, \tag{40.37}$$

and then differentiating this expression we get

$$\frac{\partial^2 \mathcal{L}}{\partial D_t^2} = \frac{\partial^2 u(C_t)}{\partial C_t^2} + (1+r_D)^2 \frac{\partial^2 u\left(C_{t+1}^L\right)}{\partial \left(C_{t+1}^L\right)^2} < 0, \tag{40.38}$$

$$\frac{\partial^2 \mathcal{L}}{\partial D_t \partial W_t} = -\frac{\partial^2 u(C_t)}{\partial C_t^2} > 0,$$

$$\frac{\partial^2 \mathcal{L}}{\partial D_t \partial (1+r_D)} = \frac{\partial u\left(C_{t+1}^L\right)}{\partial C_{t+1}^L} + (1+r_D) D_t \frac{\partial^2 u\left(C_{t+1}^L\right)}{\partial \left(C_{t+1}^L\right)^2}$$

$$= (1-z) \frac{\partial u\left(C_{t+1}^L\right)}{\partial C_{t+1}^L} > 0,$$

$$\frac{\partial^2 \mathcal{L}}{\partial D_t \partial (L_t - I_{t+1})} = -\frac{\partial^2 u(C_t)}{\partial C_t^2} > 0.$$

where the signs follow from the properties of the utility function. We make the additional assumption here that the relative risk aversion

$$z = -C_i \frac{\frac{\partial^2 u}{\partial C_i^2}}{\frac{\partial u}{\partial C_i}} < 1, \tag{40.39}$$

meaning that consumers are not too risk averse.

Using the four relationships in equation (40.38), the implicit function theorem then implies

$$\frac{\partial D_t}{\partial w_t} > 0, \tag{40.40}$$

$$\frac{\partial D_t}{\partial (1+r_D)} > 0,$$

$$\frac{\partial D_t}{\partial (L_t - I_{t+1})} > 0.$$

As we need $D_t = L_t$ at all times, we also need that $\frac{\partial L_t}{\partial I_{t+1}} = \frac{\partial D_t}{\partial I_{t+1}}$. From equation (40.33) we obtain after replacing $1 + r_L$ as given in equation (40.35), that

$$\frac{\partial L_t}{\partial I_{t+1}} = \frac{1 + \eta - \frac{\partial \eta}{\partial I_{t+1}} I_{t+1}}{(1+\eta)^2}, \tag{40.41}$$

where a value exceeding 1 can be found if $\frac{\partial \eta}{\partial I_{t+1}} < 0$, which we assumed to be the case for small investments, while for larger investments, this value will be less than 1. Note that we also assumed that $\eta > -1$. The change in D_t is now given by the total difference

$$\frac{\partial D_t}{\partial I_{t+1}} = \frac{\partial D_t}{\partial w_t}\frac{\partial w_t}{\partial I_t}\frac{\partial I_t}{\partial I_{t+1}} + \frac{\partial D_t}{\partial (1+r_D)}\frac{\partial (1+r_D)}{\partial I_{t+1}} + \frac{\partial D_t}{\partial (L_t - I_{t+1})}\left(\frac{\partial L_t}{\partial I_{t+1}} - 1\right). \tag{40.42}$$

Setting equation (40.42) equal to $\frac{\partial L_t}{\partial I_{t+1}}$, we get after a few manipulations that

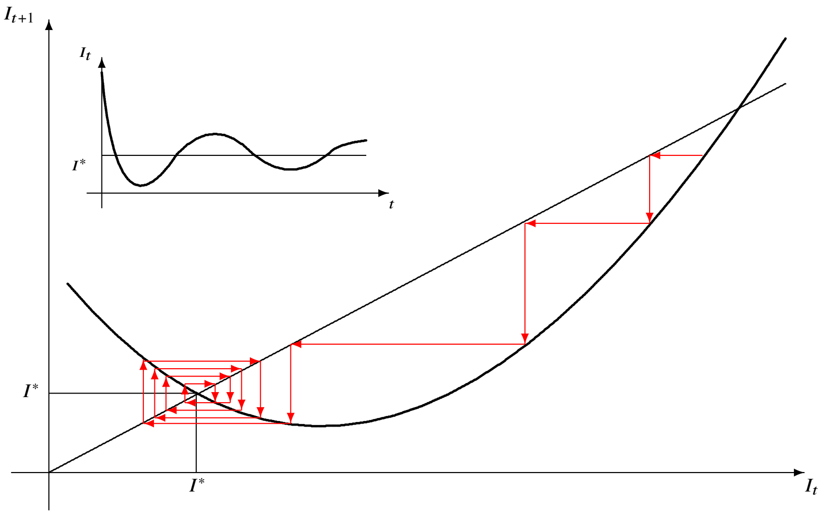
$$\frac{\partial I_{t+1}}{\partial I_t} = -\frac{\frac{\partial D_t}{\partial w_t}\frac{\partial w_t}{\partial I_t}}{\frac{\partial D_t}{\partial (L_t - T_{t+1})}\left(\frac{\partial L_t}{\partial I_{t+1}} - 1\right) + \frac{\partial D_t}{\partial (1+r_D)}\frac{\partial (1+r_D)}{\partial I_{t+1}} - \frac{\partial L_t}{\partial I_{t+1}}}. \tag{40.43}$$

From equations (40.24) and (40.40), the numerator will be positive, but the denominator will change sign at some investment level $\widetilde{I}_t$. Using that $\frac{\partial L_t}{\partial I_{t+1}} = \frac{\partial D_t}{\partial I_{t+1}}$ to eliminate the last two terms in the denominator of equation (40.43), we see that the sign of the denominator depends on whether $\frac{\partial L_t}{\partial I_{t+1}}$ is above or below 1. We have argued with equation (40.41) that for large investments this expression will be less than 1 and for small investments above 1. Hence the sign of the denominator is negative for large investments, giving $\frac{\partial I_{t+1}}{\partial I_t} > 0$ and positive for small investments, thus $\frac{\partial I_{t+1}}{\partial I_t} < 0$.

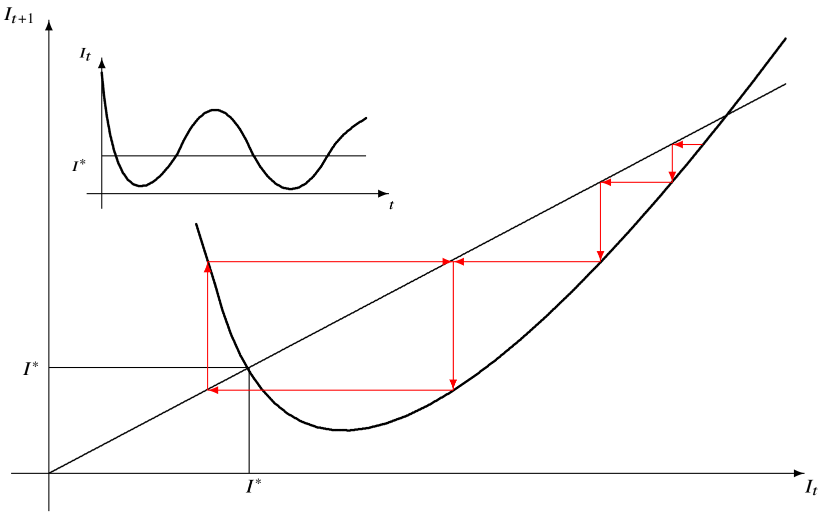
We can now graphically analyze the resulting dynamics of investments over time. Figure 40.2 shows in the upper panel the case where at the lower crossing point I^*, we have $\frac{\partial I_{t+1}}{\partial I_t} > -1$. We note that the upper crossing point is also steady state, but unstable in that any small reduction in investment, results in convergence to the lower steady state $I*$. As we see from the dynamics and the inserted small panel showing the time evolution of investments, the steady state will be reached with ever diminishing cycles. In the absence of banks the adjustment to the steady state was monotonous, but not fundamentally different. A high-investment economy is, however, not a sable equilibrium.

A far more interesting case emerges if at I^* we have $\frac{\partial I_{t+1}}{\partial I_t} < -1$ as shown in hew lower panel of figure 40.2. Here I^* is a steady state that can never be reached and investments will fluctuate cyclically forever around this steady state. Depending on the exact specification of the utility and production functions it may result in periodic behavior or chaotic behavior. We thus observe a investment patterns that, even in the long run, are continue to fluctuate.

The observed periodic behaviour arises from an income effect. This arises because not the entire loan is invested, but some fraction, depending on the elasticity, is held back to finance consumption. This elasticity now increases (η gets closer to -1) in the level of investment, thus by equation (40.36) less of the loan is consumed. As the increased investment would increase the risks to the bank (the consumers are unaffected due to $C_{t+1}^H = C_{t+1}^L$), they will stipulate low future investment and a higher use of the loan for consumption. As then the investment is lower, the risks are low and hence banks may seek higher returns by increasing loans (and risks), reversing the previous change, without ever reaching a steady state. The presence of banks allows



(a) Diminishing cycles



(b) Persistent cycles

Fig. 40.2: Equilibrium dynamics with banks

this income affect by making the relationship between loans and investments not only imperfect, but even negative as is possible in equation (40.41) for sufficiently large investments. This relationship causes the never-ending cycles in investment.

For completeness, we have to consider the case that the change in slope happens such that $\widetilde{I_t}$ is located to the left of the line $I_t = I_{t+1}$. In this case, the negative slope does not affect the equilibrium and steady state as both crossing points are exhibiting positive slopes. The some equilibria and convergence process will be observed as in the upper panel of figure 40.1.

Summary Banks smooth the consumption pattern by allowing consumer-companies to maintain the same consumption level, regardless of the outcome of investments. At times banks for this purpose have to reduce loans for investments and instead provide loans for consumption, making investments more attractive as less loans given, which reduces the risks associated with them.This low risk makes them subsequently more attractive as the profits are high due to the low risk, causing loans for investments to increase, increasing risks and making these loans less attractive again. This process continues in a never-ending cycle. Without banks, no such buffer is available and consumers have to bear the consumption risk themselves. It is thus that risks to the level of consumption have been eliminated by banks, but they have introduced varying levels of investment as a consequence.

We have shown that banks can induce periodic behaviour in investment by allowing loans to be used partially for consumption rather than investment.

Reading Banerji, Bhattacharya, & Long (2004)

Conclusions

We have seen that the optimal behaviour of banks can have profound implications for the wider economy. Their procyclical behaviour in the provision of loans makes any economic fluctuations even more pronounced by providing more loans in a well-performing economy and less loans in a poorly-performing economy. Of course, such procyclical behaviour works counter to the aim of policy makers exacerbating fluctuations and hence additional measures, like stricter capital regulations in time of high economic growth, might be required. Such macroprudential regulation may limit the ability of banks to provide loans in times of high economic growth and hence limiting growth itself, especially if companies with high growth prospects are not able to secure loans.

However, banks are not only reacting to changing economic conditions. The way loans are provided, banks may actually cause fluctuations in the amount of investments observed in an economy. By seeking to balance loans for consumption and investment, banks will continuously change the amount of loans provided for investment. With investment being an important element of economic growth, this will lead to fluctuating growth rates in the economy caused by banks.

Having seen this impact of banks on macroeconomic conditions, it is clear that any policy decisions need to take into account the impact they have not only banks itself, but also how banks can affect the wider economy. It might well be that a reduction in the risks banks are taking is desirable, but at the same time any regulatory intervention might affect the amount of lending banks are providing, affecting economic growth.

An adverse effect on economic growth may well eliminate the social benefits of reduced risks banks are taking.

Chapter 41

Credit rationing

In chapter 40.1 we had already discussed credit rationing; it was defined as a situation in which at a given loan rate, the demand for loans exceeds the amount of loans that a bank is willing to provide and where an increase of the loan rate to balance demand and supply is not optimal for the bank. We will now re-visit credit rationing, but rather than taking a look at the decision to provide individual loans and the size of such loans, we will instead focus on the aggregate supply of loans in the economy.

We assume that companies make investments I, financed though a combination of bank loans $L \geq I$ and equity E, such that $I = L + E$. The expected investment yields a return of R if it is successful and no return otherwise, where success is achieved with probability π. This probability as well as the return in the case of a successful investment are not known in advance to either the bank or the companies; however, it is known that the expected outcome, $\pi (1 + R) I$ has a distribution function $F (\cdot)$.

The bank will obtain the outcome of the investment if the companies cannot repay its loan in full and if the outcome is sufficiently high, will be repaid the loan, where we know that the highest possible loan rate is given by $\overline{r}_L$ for companies to demand loans; this loan rate is the loan rate at which companies would break even and hence for higher loan rates no loans would be demanded. If we assume that loans are financed fully by deposits with a deposit rate r_D and the loan rate is r_L, the bank profits are given by

$$\begin{aligned}
\Pi_B = & \int_0^{(1+r_L)L} \pi (1 + R) LdF (\pi (1 + R) L) \qquad (41.1) \\
& + \int_{(1+r_L)L}^{(1+\overline{r})L} (1 + r_L) LdF (\pi (1 + R) L) - (1 + r_D) L \\
= & \int_0^{(1+r_L)L} \pi (1 + R) LdF (\pi (1 + R) L) \\
& + (F ((1 + \overline{r}_L) L) - F ((1 + r_L) L)) (1 + r_L) L - (1 + r_D) L.
\end{aligned}$$

Using the Leibniz integral rule, we easily obtain that the optimal loan rate and loan amount are given by

$$\begin{aligned}\frac{\partial \Pi_B}{\partial (1+r_L)} &= \left(F\left(\left(1+\overline{r}_L\right)L\right) - F\left(\left(1+r_L\right)L\right)\right)L > 0, \qquad (41.2)\\ \frac{\partial \Pi_B}{\partial L} &= \left(F\left(\left(1+\overline{r}_L\right)L\right) - F\left(\left(1+r_L\right)L\right)\right)\left(1+r_L\right)\\ &\quad + \frac{1+r_L}{1+\overline{r}_L} - \left(1+r_D\right).\end{aligned}$$

The first term is positive as $\overline{r}_L \geq r_L$ and hence the term in bracket must be positive. The second term will be negative for some $L \geq \hat{L}$. This is because if the amount lend is very small, then $\left(1+\overline{r}_L\right)L \approx \left(1+r_L\right)L$ and hence $F\left(\left(1+\overline{r}_L\right)L\right) \approx F\left(\left(1+r_L\right)L\right) \approx 0$, while the second term will be less than 1 due to $r_L \leq \overline{r}_L$ and hence the second and final term will be jointly negative. Similarly, for very large bank loans, we have $F\left(\left(1+\overline{r}_L\right)L\right) \approx F\left(\left(1+r_L\right)L\right) \approx 1$, and the first term vanishes again, making the expression negative for $L > \hat{\hat{L}}$. For intermediate sizes of bank loans, this expression might well be positive as long as $F\left(\left(1+\overline{r}_L\right)L\right)$ is sufficiently larger $F\left(\left(1+r_L\right)L\right)$. Hence the expression is positive if $\hat{L} < L \leq \hat{\hat{L}}$.

Assuming that banks are competing such that $\Pi_B = 0$, we can use the implicit function theorem to get

$$\frac{\partial\left(1+r_L\right)}{\partial L} = -\frac{\frac{\partial \Pi_B}{\partial L}}{\frac{\partial \Pi_B}{\partial(1+r_L)}}, \qquad (41.3)$$

which is positive for $L \leq \hat{L}$ and $L > \hat{\hat{L}}$ and negative for $\hat{L} < L \leq \hat{\hat{L}}$. Figure 41.1 shows this relationship between the loan rate and the amount of loans offered. We clearly see that the loan rate is not monotonically increasing in the amount of loans offered, but downward slowing for an intermediate range of loan rates. This is the case because as loan rates are increased, the amount the companies need to repay their loans will also increase; such an increased repayment be possible for some outcomes and banks reduce the loan amount to avoid too many companies defaulting.

We can interpret this result as follow. As banks increase the loan rate, they will make larger profits from these loans, giving them an incentive to increase the provision of loans. However, as they increase the loans, they have to provide loans to ever more risky companies as the least risky companies have been selected first, given they provide the highest profits. Increasing the risks reduces the repayments the banks receive. In addition, the higher loan rate will also rule out lending to the least risky companies as their low risks are most likely yielding low investment returns, making the higher loan rates unprofitable for them. As a consequence, the bank has to rely on increasingly risky loan portfolio. While initially the increase in revenue from higher loan rates will dominate, once more risky loans have to be provided, the additional loans will be less and less likely be repaid, and this may then actually reduce the profits of the banks and they will reduce the loan rate to re-capture some of the less risky companies. Once they have reached a high level of

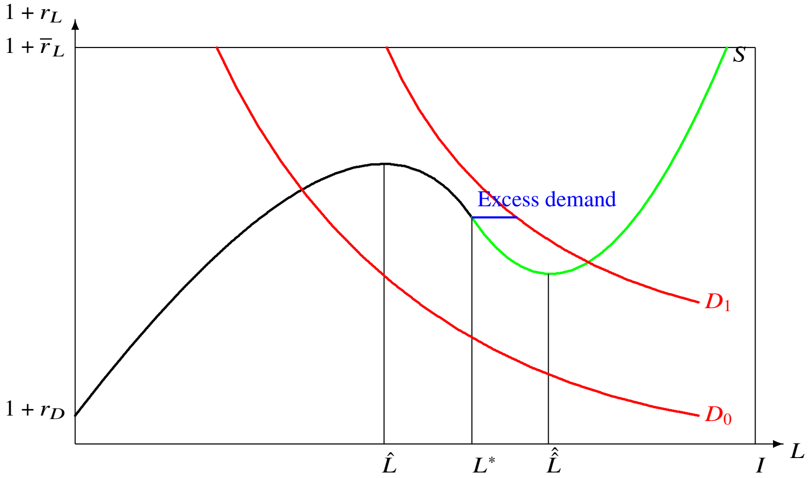


Fig. 41.1: Credit rationing due to uncertain outcomes

risk due to having provided a large amount of loans, these risks are not increasing much further and increasing the loan rate will increase profits again, and banks will increase the provision of loans in response.

Banks will maximize their profits by choosing the optimal loan repayment, $(1 + r_L)\, L$. The first order condition $\frac{\partial \Pi_B}{\partial (1+r_L) L} = 0$ solves for

$$1 + r_D = \left(F\left(\left(1 + \overline{r}_L\right) L\right) - F\left(\left(1 + r_L\right) L\right)\right)\left(1 + r_L\right), \tag{41.4}$$

where we used that $\frac{\partial L}{\partial (1+r_L) L} = \frac{1}{\frac{\partial (1+r_L) L}{\partial L}} = \frac{1}{1+r_L}$. Inserting this optimal solution into equation (41.2), we easily obtain that $\frac{\partial \Pi_B}{\partial L} = \frac{1+r_L}{1+\overline{r}_L} > 0$ and hence the amount of loans, L^*, will be such that $\hat{L} < L^* \leq \hat{\hat{L}}$. Providing more loans would reduce profits to the bank and they would therefore not be doing so, thus there will be no supply of loans beyond L^*. This has direct implications for the equilibrium loan amount.

If the loan demand is low, indicated by D_0 in figure 41.1, then an equilibrium can easily be found where demand equals supply. However, if the demand increases to D_1, we see that demand and supply only meet at a point which would require a loan amount exceeding the optimal loan size for the bank, L^*, which they therefore would not offer; this area of the loan supply is indicated in green. Banks would only offer a loan of size L^*. However, at this point, the demand for loans exceeds that of the supply of loans, causing loans to be rationed.

In times of low demand, an equilibrium can be reached in which the demand for loans and their supply are matched, even though the bank supplies less than their optimal amount of loans. They would not be able to provide their optimal ;loan amount, L^*, as this would necessitate a loan rate that would not be profitable. The supply curve S in figure 41.1 represents the line in which bank profits are equal and any point below this line would cause the bank to make losses. With the demand at L^* requiring a lower loan rate, the bank would make a loss. Thus the equilibrium would be at the point demand and supply equal. If the demand is high, D_1, demand and supply are equal only for a loan size $L > L^*$, but as the bank would not offer loans above L^*, this cannot be an equilibrium. Banks will offer their optimal loan amount L^* and competition between banks ensures that the loan rate associated with this loan offer is not raised; this results in an excess demand for loans as more companies would want to obtain a loan at that loan rate. The competition between banks prevents them from raising the loan rate to a level where the demand for loans by companies would be L^*. The result is an equilibrium with credit rationing; not all companies that would demand a loan at that loan rate are allocated a loan, even though they would be willing to pay a higher loan rate. If banks are less competitive, the supply curve would shift upwards as banks will be able to make some profits, this might alleviate credit rationing, although if the demand would increase further, credit rationing would emerge again.

We thus see that in times of high demand for loans, credit rationing may occur and not all companies can secure a loan at the prevalent loan rate. Such credit rationing emerges from the uncertainty of the investments companies conduct and hence the uncertainty about the repayment of the loan to banks. Providing more companies with loans increases the risks banks are taking, making defaults more likely than with a lower loan rate. In order to reduce defaults, banks may lower the loan rate and thereby lower their risks by attracting more companies with lower risks, balancing these two aspects to maintain their profitability.

Such credit rationing has macroeconomic implications in that during times of high demand for loans, not all companies can obtain a loan; this will reduce the investment that companies can make and will thus reduce economic growth. This might be particularly relevant if the economy is emerging from a recession and companies are seeking loans for investments into innovations, the lower investments due to credit rationing will prolong the time for the economy to recover and limit future economic growth.

Readings Stiglitz & Weiss (1981), Arnold & Riley (2009)

Review

Banks can create money as a deposit is created every time a loan is provided. Such deposit creation will be limited by the effect of banks having to pay higher deposit rates such that the increased deposits are held within the banking system and the subsequent increase in loan rate. Having to ensure that loans will be repaid by borrowers at the given loan rate, will further limit the amount of loans a bank can provide, and thus limiting the ability of banks to create money.

We have seen that banks can have significant impact on macroeconomic outcomes and they themselves are also influenced by the general macroeconomic conditions and monetary policy decisions which are aimed at affecting interest rates only. Raising interest rates can lead to banks increasing their risk-taking, but might also make banks less likely to lend, as fewer companies can afford to pay the higher loan rates. It is thus that changing interest rates do not only affect the demand for loans, but will also affect their supply by banks. Furthermore, injecting additional liquidity into the banking system may actually increase interest rates as banks invest the additional liquidity into long-term loans, but then face liquidity shortages if the liquidity injection is expected to be reversed. Thus conducting monetary policy will have side effects that may make the use of monetary policy tools less effective than envisaged by the central bank and may distort the allocation of loans within the economy. When conducting their monetary policy central bank will have to consider these secondary effects to assess the overall welfare implications of their policy decisions.

Having seen that banks are affected by monetary policy, we also explored how banks themselves may affect macroeconomic outcomes. Generally banks are more willing to give loans during times of economic expansion as during those times, borrowers are most likely to repay their loans and the high productivity allows banks to charge high loan rates. It is therefore that banks generally provide loans at times when investment is already high and reduce the provision of loans during recessions, exacerbating any economic fluctuations. Furthermore, in times of high demand for loans, but when uncertainty about the future prospects of companies is high, credit rationing might be imposed by banks seeking to maximize their profits, leaving companies short of loans to conduct investments they seek to make. We may therefor see that banks may hinder the efforts of economic policy-makers to reduce economic fluctuations. It is even that when balancing the demand of consumers and investors, banks are instigating fluctuations in investments through their provisions of loans.

Banks are not operating in an economic vacuum but they are affected by macroeconomic events, such as monetary policy decisions, not only through the impact these have on their borrowers, but these events may directly affect their decision-making. On the other hand, the decisions by banks to grant loans will have macroeconomic consequences as more loans usually result in more investment and enhanced economic growth. It is thus that banks are in complex interactions with the wider economy, being affected by it and affecting it themselves. Having a good understanding of such interactions between banks and macroeconomic conditions will allow not only for a better understanding of banks and the development of the economy, but also the selection of the most appropriate tools to manage the economy, such a monetary policy, but also the regulation of banks.

Epoligue
The valuation of banks

Valuing any company will involve to determine their expected future cash flows and discount them at an appropriate rate to the present value. This process will face the inevitable task of determining these expected future cash flows and thus assessing the company's prospects. In addition, in order to determine the appropriate discount rate to apply to these future cash flows, the risks of the company need to be assessed. This process is not fundamentally different when seeking to value a bank. However, there are added complications when assessing the value banks; firstly banks by their very nature are highly leveraged and the deposits financing their assets, which are mainly loans, will reflect any risk of banks much more accurately than the more long-term financed non-bank companies. In addition, banks are subject to capital requirements that will restrict the way they conduct their business or they will have to raise additional equity if needed; this regulatory constraint will affect the way banks operate and thus their value. In addition, banks are subject to liquidity requirements and may face bank runs, increasing the likelihood of a bank failing, even though it has not accumulated losses that make it insolvent. A further aspect to consider is the spread of losses through contagion, systemic risk, which for banks is much more important than any contagion in traditional supply chains of non-financial companies. When valuing a bank, these additional aspects need to be taken into account.

We will now develop a model that takes into account some of those features that are unique to banks to obtain the value of a bank to its owners. We will commence by looking at the value of the portfolio of loans the bank has provided and how this determines thye profits banks make.

The value of loan portfolios Let us consider a bank that has raised deposits D and hold equity E; we can then define the leverage of the bank as $\kappa = \frac{D}{E}$, representing the size of deposits (debt) relative to the equity of the bank. Using these funds, the bank provides loans L and retains a fraction ρ of their deposits as cash, $C = \rho D = \kappa\rho E$. The total amount of loans given can then be determined as $L = D + E - C = (1 + k(1 - \rho))E$. This bank grants N individual loans of size L_i, such that $L = NL_i$; these loans are repaid with probability π and any defaults by companies are assumed to have a correlation ν. We assess the loans after one time period, assuming that interest is paid in each time period and borrowers may default in every time period.

Let us define

$$\mathbf{1}_\pi = \begin{cases} 1 \text{ with probability } \pi \\ 0 \text{ with probability } 1 - \pi \end{cases}, \tag{E.1}$$

as the an indicator for the ability of a borrower to repay its loan, including interest r_L. A value of 1 indicates that the loan is repaid and a value of 0 that it is not repaid. The value of the portfolio of N loans at the end of the time period will then be given by

$$V_L = \sum_{i=1}^{N} \mathbf{1}_\pi (1 + r_L) L_i = (1 + r_L) \frac{L}{N} \sum_{i=1}^{N} \mathbf{1}_\pi . \tag{E.2}$$

This value of the loan portfolio is also the amount of loans that are repaid at the end of the time period.

The variable $\mathbf{1}_\pi$ is a Bernoulli variable and if we assume that there are a sufficiently large number of loans, N, provided, the sum of these Bernoulli variables will converge towards a Poisson distribution with parameter πN, with the assumption that these variables are independent; we assumed defaults to be correlated and will make an adjustment to the variance of outcomes below. Then, if πN is sufficiently large, this Poisson distribution can be approximated by a normal distribution with a mean and variance of πN. In order to account for the correlation between defaults, the variance of the loan value has to be adjusted by a factor $1 + \nu (N - 1)$ as can easily be verified; the expected value of the variable is not affected by the correlation. With $\mathcal{N} \left(\mu; \sigma^2\right)$ denoting the normal distribution with mean μ and variance σ^2, we have the distribution of these loan repayments given by

$$V_L \sim \mathcal{N} \left(\pi (1 + r_L) L; \pi (1 + \nu (N - 1)) (1 + r_L)^2 \frac{L^2}{N} \right) . \tag{E.3}$$

Banks use these loan repayments to repay their depositors, including the interest they have been promised, r_D. The profits of the bank at the end of the time period are then given by the amount the bank can retain from these repayments after repaying all deposits. We thus have

$$\Pi_B = V_L - (1 + r_D) D. \tag{E.4}$$

Having established the distribution of the value of the loan portfolio as well as the profits of the bank, we can now consider the regulatory constraints bank have to operate in.

Bank recapitalisation If the repayment of loans is low, a low value for V_L, the bank may make a loss, $\pi_B < 0$, and this loss will reduce the equity of the bank and thus may cause it to break its minimum capital requirements. We implement minimum capital requirements as the maximum leverage that banks can have; the maximal leverage is defined implicity through the minimum equity a bank has to hold, E^*, given its level of deposits, thus $E^* = \frac{D}{\kappa^*} = \frac{\kappa}{\kappa^*} E$. The minimum amount of equity is no longer maintained due to the low amount of loan repayments, V_L, if $E^* > E = V_L + C - D$, or $V_L < D + E^* - C$, . By inserting for the left-hand side of this condition, we easily obtain that the loan repayments have to be such that

$$V_L < H = \kappa \frac{1 + \kappa^* (1 - \rho)}{\kappa^*} E. \tag{E.5}$$

Knowing that the loan repayments are normally distributed as outlined in equation (E.3), we can easily obtain the probability of the bank facing such a breach of capital requirements, which will be given by

$$\begin{aligned}\hat{p} &= \Phi\left(\left(\frac{\kappa}{\kappa^*}\frac{1+\kappa^*(1-\rho)}{1+\kappa(1-\rho)}-\pi\right)\sqrt{\frac{N}{\pi(1+\nu(N-1))}}\right) \\ &\approx \Phi\left(\left(\frac{\kappa}{\kappa^*}\frac{1+\kappa^*(1-\rho)}{1+\kappa(1-\rho)}-\pi\right)\sqrt{\frac{1}{\nu}}\right),\end{aligned} \tag{E.6}$$

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution and we have inserted for the loan amount that $L = (1 + k(1-\rho))E$ and conducted the normalisation $\frac{V_L-\pi(1+r_L)L}{\sqrt{\pi(1+\nu(N-1))(1+r_L)^2\frac{L^2}{N}}}$. The final approximation is valid for a large number of banks, N, provided the correlation of loans is not too close to zero.

We have thus determined the first main difference to non-financial companies in that banks are subject to minimum capital requirements, or a maximum leverage, and if this condition is not met, the banks will have to take measures to ensure the regulatory requirements are met again. Thus banks will have to reduce their leverage, either by reducing the amount of deposits and consequently lending, or by increasing equity. With banks committing to provide loans for the long-term, the most feasible way is to raise additional equity and we consider this approach here. Banks will have to raise additional equity from new or existing investors. This increased equity will reduce the relative stake of the current owners of the bank and we assume that their share of profits reduces to a fraction $\lambda < 1$ of their initial holdings. In order to keep the analysis simple, we here assume that this fraction λ is fixed, regardless of the additional equity required. We might justify such an approach by asserting that if banks have to raise additional equity they will only raise an amount to allows them to meet the minimum capital requirements exactly, but that due to the costs involved in such a step, banks will raise a larger amount that will well cover any shortfalls.

We will have to take into account this possible dilution in the stake of existing owners when assessing the value of the bank to their current owners. Of course, losses can be larger than merely cause banks to not meet minimum capital requirements, they could fail by not being able to meet their current obligations to depositors. we will consider this case next.

Bank failure If the loan repayments were sufficiently low, the bank would not be able to repay its depositors in full, thus if $V_L < (1 + r_D)D - C = (1 + r_D - \rho)\kappa E$ the assets of the bank, consisting of the loan repayments and any cash reserves, $V_L + C$, will not allow banks to repay all deposits. Such a default by the bank would similarly have to be considered for a non-financial company, however, the deposit rate required would reflect the risk to depositors of not being repaid and with p denoting the probability of the bank not being able to meet these obligations, deposits are only attracted if $(1 - p)(1 + r_D)D \geq (1 + r)D$. Non-financial companies have in place more long-term financing and these terms are not easily reflect the current financial situation of the company, but the financial situation at the time the debt was obtained; as deposits in banks can be withdrawn at any time, they would reflect the current risks much better and faster than in non-financial companies. Hence, the expected return from deposits have to exceed the return of an alternative risk-free investment yielding r. From this requirement we obtain the deposit rate as

$$1 + r_D = \frac{1 + r}{1 - p}, \tag{E.7}$$

assuming that depositors are not paid more than required to attract their funds. Inserting this relationship into the condition for the bank defaulting, we obtain

$$V_L < K = \left(\frac{1 + r}{1 - p} - \rho\right) \kappa E \tag{E.8}$$

and when again using the normality of the loan repayments, we obtain the probability of the bank defaulting as

$$\begin{aligned} p &= \Phi\left(\left(\frac{\left(\frac{1+r}{1-p} - \rho\right)\kappa}{(1 + r_L)(1 + \kappa(1 - \rho))} - \pi\right)\sqrt{\frac{N}{\pi(1 + \nu(N - 1))}}\right) \\ &\approx \Phi\left(\left(\frac{\left(\frac{1+r}{1-p} - \rho\right)\kappa}{(1 + r_L)(1 + \kappa(1 - \rho))} - \pi\right)\sqrt{\frac{1}{\nu}}\right), \end{aligned} \tag{E.9}$$

where the final approximation is again valid for a large number of banks, N, and correlations ν not too close to zero. This expression can only be solved numerically for the default rate p as it is included on both sides of this equation.

Given that banks have to recapitalise if they do not meet the capital requirements, which will be less stringent than the bank failing, losses will not accumulate over time and hence the failure of a bank is very unlikely. Evaluating realistic parameter constellations, we see that any solution will be very close to $p = 0$. We assume here that failing banks are not recapitalised to avoid their failure as doing so would impose an instant loss on those investors providing additional funds. Without changing our arguments below, we could assume that banks are recapitalised if failing, but that the bank's current owners will lose their entire stake in the rescued bank.

Bank failures can also emerge from bank runs or contagion due to other banks failing. We do not model these aspects separately here, but including such a possibility would increase the probability of default of the bank, p, above the level implied from loan defaults; using a separate analysis we could determine such a probability of bank failure and add it to the probability obtained here.

Having addressed the possibility of bank failure and recapitalisation, we can now assess the capital requirements and how they relate to the type of loans the bank provides and thereby provide more endogenous capital requirements for the bank.

Regulatory requirements Bank regulations impose two restrictions on banks that lead to minimum capital requirements. The first such requirement is that the leverage κ must not exceed a maximum value κ_L. The second requirement is based on the risk of the bank; banks have to hold a fraction ω of their risky assets in equity. With the risky assets here being loans, banks would require to hold minimum equity of $E \geq \omega\pi(1+r_L)L = \omega(1+\kappa(1-\rho))\pi(1+r_L)E$, which solves for

$$\kappa \leq \kappa_R = \frac{1-\omega\pi(1+r_L)}{\omega(1-\rho)\pi(1+r_L)}. \tag{E.10}$$

With banks having to meet both capital requirements, we have can set the capital requirement of the bank as $\kappa^* = \min\{\kappa_L, \kappa_R\}$.

In addition, banks are also subject to liquidity requirements. However, we do not model these here explicitly as they do not have direct bearing on the value of the bank, but assume that to meet such regulation, banks hold a fraction ρ of their deposits as cash reserves.

Loan rates The loan rate the bank charges should reflect the risks banks are taking; a higher risk, reflected in a lower repayment rate π, should result in higher loan rates. For modelling purposes, let us assume that the Capital Asset Pricing Model applies to loan rates. Defining the Sharpe ratio of the market loan portfolio as $\theta = \frac{\pi^*\left(1+r_L^*\right)-(1+r)}{\sqrt{\pi^*(1-\pi^*)}}$, noting that the variance of loans with repayment rate π is given by $\pi(1-\pi)$. If we assume that the correlations between the loans the bank has given is identical to the correlation of their loans portfolio with the market loan portfolio, the Capital Asset Pricing Model can be written as $\pi(1+r_L)-(1+r) = \nu\sqrt{\pi(1-\pi)}\theta$, which solves for the loan rate to be given by

$$1+r_L = \frac{1+r}{\pi} + \nu\sqrt{\frac{1-\pi}{\pi}}\theta. \tag{E.11}$$

If we assume that market participants are risk-neutral, then they will set the loan rate such that $\pi^*\left(1+r_L^*\right) = 1+r$ and hence $\theta = 0$; however, in general, loans will carry a risk premium and $\theta > 0$.

We can now have all elements to determine the value of the bank to their current owners.

Bank value As discussed, if a bank does not meet its minimum capital requirements, then it needs to be recapitalised and we assumed that in this case the existing owners of the bank will only retain a fraction λ of the bank for the future. If the bank defaults by not being able to repay its depositors, we assume that the bank is liquidated and bank owners do not obtain any payments. A recapitalisation of the bank implies that the those providing this capital will take a fraction $1-\lambda$ of the bank's future profits, leaving a fraction λ to the exiting bank owners. It is thus that the future value of the bank to current owners, V_{t+1}, is allocated to the current owners is given by

$$V_{t+1} = \begin{cases} (1+g)\,V_t & \text{with probability } 1-\hat{p} \\ \lambda\,(1+g)\,V_t & \text{with probability } \hat{p}-p \\ 0 & \text{with probability } p \end{cases} \tag{E.12}$$

, where we assumed that banks grow at a steady rate g, for example through the expansion of lending and deposits. If the bank does meet the capital requirements, $1-\hat{p}$, the full current value of the bank is taken forward, including any growth of the bank, and if the bank defaults, p, the current owners completely lose the value of the bank. In the intermediate case that the bank needs to be recapitalised but does not default, $\hat{p}-p$, the current owners will only retain a fraction λ of their current holdings, including any growth of the bank. The capital requirements are set such that a bank meeting these requirements will be able to repay its depositors and hence $p < \hat{p}$, implying that $K < H$ and hence $\kappa^* < \frac{1-p}{r+p(1-\rho)}$.

The current profits by the bank are given by $\Pi_B = V_L - (1+r_D)\,D$ and we can then get the value of the bank to its current owners as

$$V_t = \Pi_B + \frac{(1-\hat{p})\,V_{t+1} + (\hat{p}-p)\,\lambda V_{t+1}}{1+r_V}, \tag{E.13}$$

where r_V denotes the weighted average cost of capital of the bank at which future values are discounted. The current owners firstly obtain the profits the bank has made in that time period and with probability $1-\hat{p}$ the owners retain full ownership of the bank and its future profits as the bank continues to operate unchanged, while with probability $\hat{p}-p$ the bank needs recapitalisation and the owners retain a fraction λ of these future profits and hence the value of the bank. If we assume that profits, and hence bank value grow at a constant rate g, then we will have $V_{t+1} = (1+g)\,V_t$ and we can solve for the bank value as the steady-state equilibrium, which becomes

$$V_t = \frac{1+r_V}{(1+r_V)-(1+g)\,(1-\hat{p}+(\hat{p}-p)\,\lambda)}\Pi_B. \tag{E.14}$$

We now need to determine the weighted average cost of capital. in company valuation, the price-earnings ratio of a company is given by $\frac{V_t}{\Pi_B} = \frac{1+g}{r_E-g}$, where r_E denotes the cost of equity. The weighted average cost of capital is given by

$$r_V = \frac{E}{E+D}r_E + \frac{D}{D+E}r_D = \frac{r_E+\kappa r_D}{1+\kappa} \tag{E.15}$$

and inserting this expression as well as equalising the two price-earnings ratios, we can solve for the cost of equity to be

$$r_E = -\frac{\kappa\,(1+r_D)-2g}{2} + \sqrt{\frac{(\kappa\,(1+r_D)-2g)^2}{4}+\xi}, \tag{E.16}$$

with

$$\xi = g\,(1+\kappa\,(1+r_D)) + (1+g)\,(\kappa r_D - (1+\kappa)\,g \\ + (1+\kappa)\,(1+g)\,(\hat{p}-(\hat{p}-p)\,\lambda))\,.$$

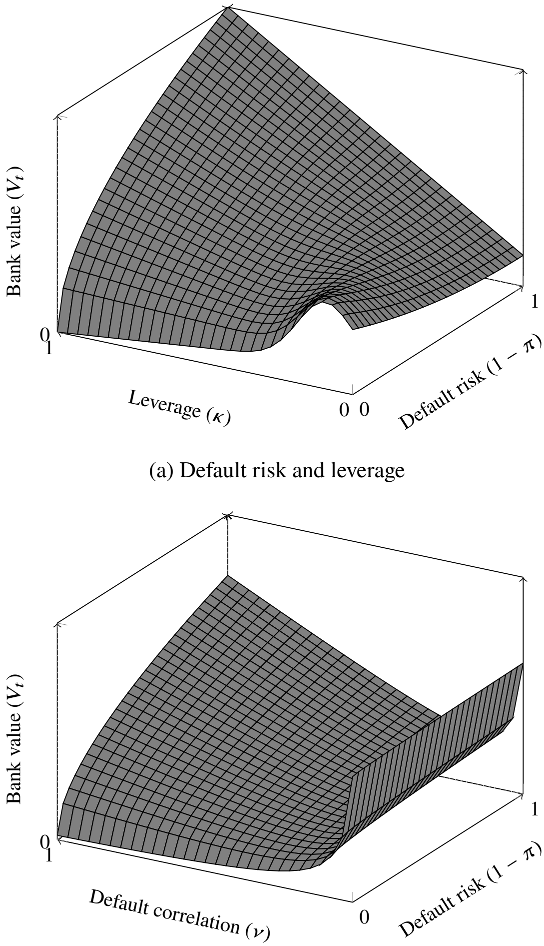
Using this expression we can obtain the cost of capital and thence the value of the bank.

Properties of the bank value Banks mainly have control over three parameters, the riskiness of the loans they provide, π, their correlation, μ, and the leverage they employ, κ, therefore will focus our analysis on these variables. Changing the remaining parameters will not materially change the results. An analytical assessment of the influence these three parameters of interest have on value of the bank is not feasible, given the complexity of the resulting expressions. Instead we selected realistic parameters and for those of less interest and conducted a visual analysis of the results.

The bank values for different parameter combinations are illustrated figure E.1. When selecting two parameters of the three and determining the optimal value of the third parameter, we see that mostly the remaining free parameter is selected at its minimum or maximum value, or very close to it; for this reason we will refer to the value of the parameter that maximizes the bank value as being either high or low. Figure E.2 provides the key relationship between these variables as established from the graphical illustration.

We clearly can determine that low risks, $1-\pi$, low leverage, κ, and low correlation, ν, are reinforcing themselves, as doe the corresponding high values. This is consistent with the observation that the highest bank values are located at the position where parameters are either the lowest or the highest. We also see that for all but the lowest default correlation, which might be difficult to find, a low default correlation will make it optimal to provide high-risk loans. It therefore follows that it will be optimal for banks to choose the highest possible default rate, the highest default correlation, and the highest possible leverage; choosing high default rates and high default correlations together imply that the bank chooses the highest possible risk for their loan portfolio. While, depending on parameter constellations, it might be that the bank value is higher if choosing a low-risk loan portfolio, low default risk and low correlations, in combination with a low leverage, this will not be optimal as in this case increasing the default rate would be optimal.

This result requires some explanation. What we observe is that as we increase the initially low values of any of these three parameters, loan risk, leverage, and correlation, the value of the bank reduces. As the loan risk increases, it has multiple effects; it increases the loan rate the bank charges and if the Sharpe ratio θ is sufficiently high, this increased loan rate more than compensates the bank for the higher default rate on loans, increasing bank profits and hence, everything else being equal, the bank value. However, the increased risk the bank is exposed to also increases the required rate of return on equity and thus the weighted average cost of capital, counteracting the increase in the bank value from higher profits. With low leverage the effect of a risk increase is very marginal as these two effects cancel each other out. Similarly, the higher risk will increase the possibility of the bank requiring a recapitalisation, but the higher returns from loans offset this effect mostly. The overall effect for small loan risks is to increase bank values slightly and then as risks increase, the bank value falls. For higher loan risks or higher default correlations,



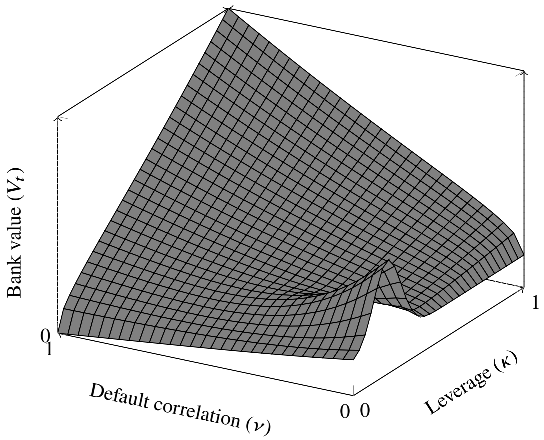
(a) Default risk and leverage

(b) Default risk and default correlation

Fig. E.1: Bank value as a function of default risk, default correlation, and leverage

the risk premium makes loans more profitable and the bank profits will increase, causing the value of the bank to increase as well.

If we increase the leverage from a low value, the probability of the bank requiring a recapitalisation, and thus imposing losses of $1-\lambda$ on the bank owners, will increase and reduce the value of the bank to their owners. However, once the leverage has increased, the probability of a recapitalisation remains approximately stable and does therefore not affect the bank value for higher leverages. An increase in the correlation of loan defaults will increase the risk of the loan portfolio as diversification is reduced, this increases the probability of the bank requiring a recapitalisation and hence the loan value would fall.



(c) Default correlation and leverage

Fig. E.1: Bank value as a function of default risk, default correlation, and leverage (ctd.)

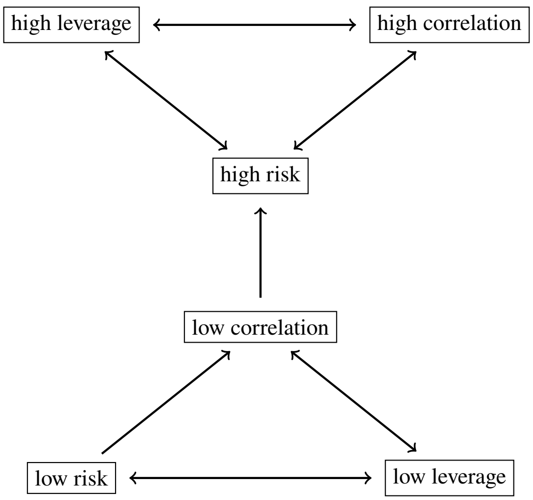


Fig. E.2: Optimal parameter values to maximize bank value

Once the values for the leverage and correlation have increased, the probability of recapitalisation does not increase much further, allowing other factors to dominate. If the loans the bank provides are profitable, a higher leverage allows for more loans to be provided and with the bank profits increasing, will the bank value increase. This is enhanced by the weighted average costs of capital reducing due to the increasing use of depositors that have lower costs than equity; deposit rates are only affected by the risk of the bank defaulting and we had argued above that this probability is so small that it barely affects the deposit rate. The exception to this finding that a higher leverage increases the bank value is for very safe banks. This is because the cost of equity will increase substantially due to the higher probability of a recapitalisation and due to the low risk of the loans and subsequently the low loan rate and profitability of loans, the higher leverage does not result in a significant increase of profits, causing the overall bank value to reduce.

However, once these parameters have increased in value, a further increase will increase the bank value rather than reduce it. The probability of recapitalisation remains stable as we increase any of these parameters, which also implies that the cost of equity remains stable, but loans become more profitable due to the increased risk from higher default risk, but also a higher default correlation as we had assumed that loans have this correlation also with the market portfolio of loans, increasing profits and hence the bank value. An increase in leverage will increase the loan amount and thus bank profits similarly. This result requires the risk premium for loans, the Sharpe ratio θ to be sufficiently high; if this risk premium is low, loans do not increase in profitability sufficiently to overcome the adverse movement of the other variables.

We also observe that the capital requirements are entirely driven by κ_R, the risk-based capital requirements rather than the maximum leverage, κ_L, provided we choose a realistic value for this parameters. The maximum leverage that banks are supposed to have is found to be close to 10, which is a value we approximately observe in large and sound real banks.

Summary We have seen not only that the valuation of banks is more complex than the valuation of other, non-financial companies, but that banks maximizing their value will also seek to take high risks, as evidenced by high default risks, high default correlations, and high leverage. The complexity of the valuation of banks arises from the interplay of risk and return in banks, for loans as well as deposits, in addition to the regulatory constraints that impose capital requirements and can lead to banks requiring a recapitalisation. These aspects have to be taken into account when seeking to assess the future earnings and risks of a bank. In addition, banks will have other lines of business that rely on fee income, such as account and service fees, as well the expansion of their business, which needs to be taken into account in a more conventional way.

The observation that, given a sufficiently high risk premium on loan rates, banks maximize their value if they seek the highest risk justifies the regulation banks are subjected to. These regulations impose limits on the leverage through capital requirements, supported by regulation on the concentration of risk, the default correlation, and the risks bank take on loans. This is achieved, for example by imposing higher capital requirements on more risky loans than a comparable low-risk loan. Given the importance of banks for the economy and the widespread effect that their failure would have, the results on the value of banks suggest that these regulations are imposed to limit the risk-taking by banks, which without such constraints would be higher.

Reading Krause (2025b)

Final reflections: The complexity of banking

The conduct of core banking activities seem to be fairly straightforward. To grant a loan the bank, or its employees, assesses the ability of the potential borrower to repay the loan, including any agreed interest, and if this assessment shows that the criteria set for granting a loan have been met, it is approved. Such an assessment of the creditworthiness requires knowledge and skills, but is not an inherently complex decision. Managing deposits seems to be an even less demanding task as banks only need to ensure to offer conditions that are sufficiently attractive for depositors, given the conditions offered by other banks and the presence of alternative investment opportunities.

Lending decisions However, this supposed simplicity of conducting a banking business neglects many additional aspects that banks need to consider in their decision-making. such additional concerns will not necessarily be impacting the assessment of creditworthiness as conducted by employees assessing loan applications, nor will they directly be reflected in deposit conditions; however, they should affect the guidelines under which employees operate. We have seen that the type of loans a borrower applies for can reveal information about their own assessment of the risks they are taking with the proceeds of the loan, in chapter 9 we argued that borrowers taking low risks would be willing to provide collateral, while those taking higher risks would not be offering collateral. Thus a borrower applying for a loan and willing to provide collateral might be offered better loan conditions than, from the bank's assessment, an otherwise comparable applicant who is not willing or only reluctantly willing to consider the provision of collateral; these better loan conditions are not only the result of the bank facing lower risks due to the collateral provided, but is also due the information the offer of a collateral provides to the bank.

Furthermore, offering loans at high loan rates to account for the risks the banks has identified, might provide incentives for the borrower accepting such a loan to choose even more risky investments or strategically default as we have discussed in chapters 7 and VIII. Banks might react to this possibility with credit rationing if a smaller loan would entice the borrower to reduce the risk they are taking or to not strategically default. It is thus that the loan conditions a bank offers its borrowers will affect their behaviour; this reaction by borrowers needs to be taken into account when making loan offers. While such considerations might be beyond the duties of employees assessing loan applications, they should be taken into account when providing guidelines for the assessment of loan applications to ensure that the characteristics of the incentives loan conditions give to borrowers are included in the overall assessment.

The complexity of deciding on loan applications does not stop at this point, but as was analysed in chapter 11, more strategic long-term concerns might determine lending decisions and the conditions offered. It might be worthwhile to offer new borrowers more favourable loan conditions than existing borrowers; the aim would be to attract these borrowers to the bank and then through learning more about them by repeated interactions become more informed about the risks these borrowers are actually presenting to the bank. having this informational advantage compared to competing banks would allow a bank to recover in the future the losses made from initially very favourable loan conditions. The benefits of offering loans in such a strategic manner would need to be communicated to decision-makers in the bank such that they can offer initially more favourable loan conditions than their risk-assessment suggests. But is not only for new borrowers that loans might offered at conditions that are not aligned with the risk assessment; borrowers in financial distress might be granted loans as the bank hopes to recover some of the losses from the profits the new loan will generate. Here, a bank might approve a loan to a borrower that might otherwise be deemed to be not creditworthy. Again, such strategic behaviour by the bank needs to be communicated to decision-makers.

Deposit taking Banks will compete with each other for deposits and will also face competition from other investment opportunities, such a risk-free government securities, risky investment such as stocks, or the investment into property, amongst many other alternatives. It is, however, not only the conditions to attract depositors that need to be considered. As deposits typically can be withdrawn at any time without notice, banks face the risk of bank runs as discussed in chapter 15. We have seen that banks do not only have to be solid in that they are able to repay depositors if they decide to withdraw their funds, but they must be believed to have this ability. It is thus that banks have to establish a reputation for being able to meet the demand of depositors for any such withdrawals. It is therefore especially important for bank, more so than for non-financial companies, to project an image of reliability, stability, and being trustworthy. Banks are retaining some of the deposits they obtain as the form of a liquidity reserve in order to meet the demand from those depositors that withdraw their funds. When setting deposit rates banks have to be aware that higher deposit rates make banks more vulnerable to bank runs as the amount depositors withdraw are higher due to the accumulated interest. Thus high deposit rates might attract depositors, but they also make the bank more vulnerable to a bank run.

As occasionally their liquidity reserve will not be sufficient to meet the demand by depositors, while at other times their liquidity reserves are in excess of what is required, banks have resorted to provide liquidity reserves to each other in the form of interbank loans. Such loans, as we have seen in chapter 16, will allow banks to withstand unexpected deposit withdrawal and knowing that banks can obtain additional liquidity reserves from other banks will therefore reduce the threat of a bank run as depositors will feel re-assured that their deposits are safe. However, the provision of interbank loans may fail at critical times and thus cannot be taken as granted in all circumstances. If a bank fails, either due to the withdrawal of deposits or through high defaults on loans they have provided, interbank loans can be detrimental to the banks having provided such loans as we have seen in chapter 32. The losses of one bank will impose losses on the provider of interbank loans and they might face losses as a consequence that could cause them to fail. Thus interbank loans might be able to reduce the threat of bank runs and failures from deposit withdrawals, but they expose the banking system as a whole to systemic risk and the lending bank to losses.

Many countries operate implicit or explicit deposit insurance with the aim to prevent bank runs due to depositors losing trust into the stability of individual banks or the banking system as a whole. Even without considering the ability of the deposit insurance to cover all deposits in cases where a major bank fails, chapter 18 has pointed out that the very presence of deposit insurance can have consequences for the provision of loans, namely banks might be willing to provide more risky loans than they would in the absence of deposit insurance. Here the reduction of risks to depositors will lead to an increase of risks for the economy overall. While this is of no direct consequence for banks, it shows how a well-intended measure can have negative side effects.

Given the importance of deposits, they make up the majority of funding for the provision of loans, it is essential that banks are able to attract deposits by offering good conditions, but they also need to retain the trust of depositors to avoid a bank run. For this reason banks might not want to provide loans that are too risky, even if such loans were highly profitable; depositors, even if adequately compensated through a higher deposit rate, might not be willing to accept such risks and withdraw funds. Also, when providing loans with higher risks, banks might lose the trust of depositors more easily as they might worry more about a downturn in economic conditions affecting the loan default of such a banks that of a bank known to provide low-risk loans. We therefore see that the lending policy of banks is not only important for their profitability and needs to take into account the incentives to borrowers, but it also impacts the ability to attract and retain deposits. Having adequate safeguards in the form of liquidity reserves and equity holding, as discussed in chapter 26

On first sight it seems that when obtaining and retaining deposits banks play a rather passive role in that they set conditions that are sufficiently attractive. The lending policy of banks, however, affect the ability of banks to retain existing deposits in particular; while a bank might seek to attract additional deposits in order to increase the amount of loans they can grant, losing existing deposits will lead to the bank failing as loans are in most cases provided on a long-term basis and cannot be called in to meet the demand of depositors withdrawing their funds. Instead banks need to retain the trust of depositors, taking into account the risks of the loans they are granting, but also the risks of other banks to which they have provided interbank loans. As interbank loans a bank has obtained might not be extended by a bank facing a liquidity shortage itself, the stability of banks granting interbank loans are equally important.

Bank behaviour Banks play a central role in the economy and their presence brings considerable advantages. Banks allow that lending is conducted efficiently and most importantly that while they allow depositors seeking to access their funds at any time, the loan they provide are long-term and offer a stable financial environment for borrowers to conduct investments. We have discussed these advantages in chapters 3 and 4. While banks are subject to bank runs and systemic risk, such events are not frequently observed and thus do not reduce the benefits of banks for the wider economy substantially.

With the central role of banks in the economy in terms of meeting the demands for borrowers to obtain long-term finance for their investment and depositors to access their funds if needed, banks also have a high responsibility to discharge their duties responsibly and thereby avert severe negative consequences for the economy. It is therefore essential that employees and banks have incentives to provide their services fairly and equally to all customers. In chapter 28 we discussed the possibility of malpractice in banks and how the impact of any consequences can be managed, even if it cannot be eliminated. We saw the importance of incentives to banks and individual employees to not engage in activities that are seen to be detrimental to the bank or society as a whole. Having a robust system in place to prevent any abuse of their central position in the economy will be essential for a wider trust in the banking system and therefore also for the stability of deposits in banks. Of course, employees, need to be incentivised to conduct their role to the best of their ability and chapter 27 discussed some policies bank might want considering to that effect.

While incentives for employees to work to the best of their ability and to not engage in malpractice should be a universal desire of any company, regulator, or society as a whole, it is especially important in the case of banks. The inefficient provision of loans detrimentally affects economic growth and hence the overall welfare in the economy. Given the large amount of money involved in banking, the potential benefits from malpractice are significant and in many cases substantially higher than in non-financial companies. It is thus even more important for banks to reduce the risk of malpractice and ensure their employees working as efficiently as possible.

Regulation The tight regulation banks face, in particular with respect to capital requirements, will determine many of the bank policies on lending. With banks having a given amount of equity available to them, they will have to restrict lending such that they meet the regulatory constraints. These restrictions on their lending activity will have to be reflected in the lending policies that are applied to decide on loan applications. However, regulatory constraints also affect the incentives of banks to provide loans and may well induce banks to pursue a policy that is seen as more risky by the bank itself or their depositors, even though due to the way regulations are enacted, they are seen as less risky in that context. While it may be tempting for banks to exploit this misalignment of actual risk and risk as defined by relation, it may well affect their ability to retain deposits. Thus banks will be restricted not only regulation but also by the need to give the perception of being safe for depositors. This will often limit the ability of banks to use the way regulations are implemented to their advantage; in the management of the bank, this aspect needs to be considered carefully.

Compliance with an often complicated regulatory framework, in some instances regulation might even contradict each other, is essential for banks to retain their reputation with depositors. being in breach of regulatory requirements, even if of no real impact on the risks the bank faces, can erode the trust the public, and hence depositors have in banks. This can have the consequence of banks facing deposit withdrawals and akin to a bank run. Banks also need to be careful when seeking to take advantage of regulatory loopholes, gaps, or inconsistencies as any increase in the risk the bank takes as a result reduce the perception of the bank being safe and may negatively affect the willingness of depositors to retain their funds, even though regulatory constraints have all been met.

But it is not only regulation that is aimed at banks that affects their behaviour, banks are also reacting to monetary policy decisions, and not always in the way the decision-maker intends this behaviour to be. In turn, banks will also affect macroeconomic outcomes through their provision of loans and these will have to be taken into account and anticipated by a policy-maker. It is thus that banking is not a passive element in economic policy-making, predominantly in monetary policy decision, but an active element that reacts to any decisions that have been taken and whose actions will affect the macroeconomic outcome.

Summary We have seen that banks have an outstanding role in the economy and it is essential for a stable economic development that the banking system is working effectively. While the task of assessing loan applications might be seen as not overly difficult, it has become apparent that the decisions to grant a loan and the conditions attached to it, such a loan rate and collateral requirements, are affecting the decisions of the borrower after they have obtained the loan. These decisions need to be anticipated by banks they will affect the ability of the borrower to repay the loan. this is in contrast to non-financial companies who after selling a product or service will normally not be affected by the way it used by their customers. It is not only that these consequences of the offer conditions need to be considered when making the loan offer, they will also affect the funding the bank receives from depositors. Banks need to maintain a reputation with depositors for being a safe place for their funds; when seeking to make their bank safer by being able to withstand deposit withdrawals better through the use of interbank loans, banks expose themselves to systemic risk, which in turn will negatively affect their ability to retain deposits.

It is therefore that decisions banks make are highly interconnected and one measure to reduce risks on one aspect can increase risks elsewhere. It is this interconnection of decisions and risks, along with the consequences on borrowers' behaviour of any decisions banks make, that transforms banking from a seemingly straightforward business of assessing the creditworthiness of loan applicants to a complex network of highly interconnected consequences and decisions.

Appendix

In this book I assume that readers are familiar with the economic ideas, and the associated mathematics, as is covered in typical first year programmes of economics degrees. I therefore assume that readers are aware of utility functions and the concept of equilibrium, along with budget constraints. It is also necessary to be aware of perfect and imperfect competition in markets; specifically, competitive equilibria, oligopolistic models of competition and monopolies. Strategic interaction as introduced with basic non-cooperative game theory is also required, along with the concept of the Nash equilibrium.

The mathematical tools associated with such basic economic ideas include the rules of differentiation and integration as well as the maximization and minimization of functions. Knowledge of linear algebra, in particular the use of matrices and vectors is also beneficial. I also assume that readers have elementary knowledge of statistics. They should be aware of the properties of probabilities and probability distributions, including specific knowledge of normal, binomial and uniform distributions, along with an understanding of means, variances, covariances, and correlations. Knowledge of basic linear multivariate regressions is also helpful.

While most models in investment banking do not go beyond such basic concepts, on occasions, I require more advanced knowledge of some specific economic theories as well as mathematical tools. While a detailed understanding of these is not necessary to appreciate the models and their results, it is unavoidable to require such knowledge for the derivation of the results and sometimes their interpretation. Those economic theories and mathematical tools that are not commonly taught at an introductory level but needed in model here, are included in this appendix. The treatment will be such that economic theories are not presented overly mathematical, instead the focus will be on gaining an understanding of the basic idea. For mathematical tools, mainly the results are stated without formal derivations or proffs.

For reference this appendix also include some basic notations of the valuation of securities and assessment of credit risk, which is a key expertise required for investment bankers.

Appendix A

Mathematical tools

A.1 Implicit functions

Let us assume we have a function f of $n \geq 2$ variables, x_i, with $f(x_1, x_2, \ldots, x_n) = 0$. Such a function might emerge in our context mainly from a first order condition, in which case f would be the first derivative of the objective function, or a constraint to an optimization problem. We would often be interested in the trade-off between two variables, x_i and x_j, that allows this function to remain at zero; we would thus would seek to find $\frac{\partial x_i}{\partial x_j}$.

To achieve this goal, consider the total differential of the function f, which is given by

$$df(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i} dx_i, \tag{A.1}$$

where $\frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i}$ denotes the partial derivative with respect to x_i. The total differential can be interpreted as a linear approximation of the change of the function f, represented by $df(x_1, x_2, \ldots, x_n)$, through the changes of all the variables, dx_i, who have a marginal influence of $\frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i}$, which can be interpreted as the slope of the function f in the direction of variable x_i.

Let us now assume that we are only interested in two variables, x_i and x_j. We set all the other changes of variables in equation (A.1) equal to zero and obtain

$$\begin{aligned} df(x_1, x_2, \ldots, x_n) = & \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i} dx_i \\ & + \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_j} dx_j. \end{aligned} \tag{A.2}$$

From the condition that $f(x_1, x_2, \ldots, x_n) = 0$, we know that its value cannot change, thus we need $df(x_1, x_2, \ldots, x_n) = 0$. If we insert for $df(x_1, x_2, \ldots, x_n)$ from equation (A.2), we can solve this expression for

$$\frac{dx_i}{dx_j} = -\frac{\frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_j}}{\frac{\partial f(x_1,x_2,\ldots,x_n)}{\partial x_i}}. \tag{A.3}$$

This expression, known as the *implicit function*, now provides us with the relationship between x_i and x_j, namely how x_i needs to change if we change x_j (marginally). The implicit function gives us therefore the trade-off between these two variables such that $f(x_1, x_2, \ldots, x_n) = 0$.

A.2 Matrix theory

Let us assume that we have a matrix $\mathbf{A}$ whose entries are all non-negative and they do not exceed one, hence for each element we have $0 \leq a_{ij} \leq 1$. Then the matrix $(\mathbf{I} - \mathbf{A})^{-1}$ only has non-negative elements, provided the largest eigenvalue of the matrix $\mathbf{A}$ is below 1. The inverse matrix can be expanded into a Neumann series, $(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=1}^{+\infty} \mathbf{A}^k$ and as all elements in $\mathbf{A}$ are non-negative, all elements in $\mathbf{A}^k$ are non-negative, making the entire expression non-negative. This series is only well-defined if it converges, which requires the largest eigenvalue to be below 1. This is fulfilled if the sum of all rows are below 1.

A.3 Constrained optimisation

In many cases we will face constraints when maximizing or minimizing our objective function. We might have to apply a constraint due to limited resources available (budget constraint) or have to meet other constraints to induce certain behaviour (incentive constraint). There might also be limits on the maximum number of units of a good available from production, that the demand for a good cannot be negative, or if seeking an optimal probability that the solution must be between 0 and 1. We therefore face a constrained optimization.

The most common form of constrained optimisation in economic theory is that the constraint is an equality, and we apply the Lagrange multiplier, as detailed in appendix A.3.1. A less common and more general case arises if the constraints consist of equalities and inequalities, requiring the use of the Karush-Kuhn-Tucker conditions in appendix A.3.2. The Lagrange multiplier can be seen as a special case of the Karush-Kuhn-Tucker conditions where no inequalities exist. Nevertheless, we treat these two approaches separately as many of the interpretations of Lagrange multipliers also apply to the Karush-Kuhn-Tucker conditions and this therefore facilitates its treatment and interpretation.

A.3.1 Lagrange multiplier

Let us assume we have n decision variables x_i that we need to choose optimally by maximizing or minimizing an objective function $f(x_1, x_2, \ldots, x_n)$, such as profits, costs, or utility. In addition, we have m constraints that need to be fulfilled with equality, denoted $g_k(x_1, x_2, \ldots, x_n) = c_k$. We can add these constraints to the objective function without changing its value if we deduct c_k. When doing so, we multiply them with a constant ζ_k and obtain

$$\mathcal{L} = f(x_1, x_2, \ldots, x_n) + \sum_{k=1}^{m} \zeta_k (c_k - g_k(x_1, x_2, \ldots, x_n)). \tag{A.4}$$

This function is commonly referred to as the *Lagrangian* and ζ_k as the *Lagrange multipliers*. We now need to find the maximum for this Lagrangian and the variables to be determined are the decision variables x_i and the Lagrange multipliers ζ_k. The first order conditions are then given by

$$\begin{aligned} \forall i = 1, \ldots, n : \frac{\partial \mathcal{L}}{\partial x_i} &= \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i} \\ &\quad - \sum_{k=1}^{m} \zeta_k \frac{\partial g_k(x_1, x_2, \ldots, x_n)}{\partial x_i} = 0, \\ \forall k = 1, \ldots, m : \frac{\partial \mathcal{L}}{\partial \zeta_k} &= c_k - g_k(x_1, x_2, \ldots, x_n) = 0. \end{aligned} \tag{A.5}$$

The second set of equations recovers the constraints and we can solve the resulting equation system fir the x_i, which will provide us with the solution(s) to the optimization problem, subject to the constraints imposed.

We can also provide an interpretation of the Lagrange multipliers as the marginal effect the constraints have on the optimal value of the objective function $f(x_1, x_2, \ldots, x_n)$. We easily have that $\frac{\partial \mathcal{L}}{\partial c_k} = \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial c_k} + \zeta_k = 0$ and hence

$$\frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial c_k} = -\zeta_k. \tag{A.6}$$

We can interpret the (negative) Lagrange multiplier as the marginal product of changing the constraint. If we were to increase the value of the constraint, c_k, the effect on the objective function $f(x_1, x_2, \ldots, x_n)$ is given by $-\zeta_k$. In economic terms, the marginal product would be the price of a good. We can therefore interpret the Lagrange multipliers as the shadow prices of the constraint.

A.3.2 Karush-Kuhn-Tucker conditions

As before, let us assume we have n decision variables x_i that we need to choose optimally by maximizing or minimizing an objective function $f(x_1, x_2, \ldots, x_n)$, such as profits, costs, or utility. In addition, we have m constraints that need to be fulfilled with equality, denoted $g_k(x_1, x_2, \ldots, x_n) = c_k$, and now additionally l inequality constraints $h_s(x_1, x_2, \ldots, x_n) \leq c_s$. If the actual constraint requires non-negative solutions, a multiplication by -1 will transform the inequality as shown here.

Let us now treat these inequalities the same way as equalities and apply the Lagrange multiplier, such that

$$\mathcal{L} = f(x_1, x_2, \ldots, x_n) + \sum_{k=1}^{m} \zeta_k \left(c_k - g_k(x_1, x_2, \ldots, x_n)\right) + \sum_{s=1}^{l} \zeta_{m+s} \left(c_{m+s} - h_s(x_1, x_2, \ldots, x_n)\right) \tag{A.7}$$

and the first order conditions become

$$\begin{aligned} \forall i = 1, \ldots, n : \frac{\partial \mathcal{L}}{\partial x_i} &= \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i} \\ &\quad - \sum_{k=1}^{m} \zeta_k \frac{\partial g_k(x_1, x_2, \ldots, x_n)}{\partial x_i} \\ &\quad - \sum_{l=1}^{s} \zeta_{m+l} \frac{\partial h_l(x_1, x_2, \ldots, x_n)}{\partial x_i} = 0 \\ \forall k = 1, \ldots, m : \frac{\partial \mathcal{L}}{\partial \zeta_k} &= c_k - g_k(x_1, x_2, \ldots, x_n) = 0, \\ \forall s = 1, \ldots, l : \frac{\partial \mathcal{L}}{\partial \zeta_{m+s}} &= c_s - h_s(x_1, x_2, \ldots, x_n) = 0. \end{aligned} \tag{A.8}$$

In deviation from the solution of this equation system for Lagrange multipliers as explained above, we need to consider whether the condition $h_s(x_1, x_2, \ldots, x_n) \leq c_s$ is binding in that $h_s(x_1, x_2, \ldots, x_n) = c_s$. In this case we treat ζ_{m+s} as a Lagrange multiplier. If this constraint is not binding, then $\zeta_{m+s} = 0$ and the constraint can be dropped from the solution.

A solution concept would thus be to determine the value of all ζ_{m+s}. If $\zeta_{m+s} > 0$, then the constraint is binding and if $\zeta_{m+s} = 0$ then it is not binding and the constraint as well as ζ_{m+s} can be eliminated from the solution. Negative values for ζ_{m+s} cannot be observed. Using the interpretation of the Lagrange multipliers as shadow prices, we see that if this shadow price is positive, it affects the value of the objective function, implying that the constraint is binding. If the constraint is not binding, changing the constraint does not affect the outcome, hence the shadow price is $\zeta_{m+s} = 0$.

The remaining positive ζ_{m+s} are treated like Lagrange multipliers and the associated constraints as binding equalities, giving rise to solving an equation system akin to that emerging from using Lagrange multipliers.

A.4 Bayesian learning

Many models in banking involve incomplete information about the structure of the economy, causing outcomes to be perceived as random from the perspective of the bank, borrower, or depositor. It is however, often that some additional information is available to these market participants that will affect the distribution of such random outcomes. Assume we have an outcome H that has some probability of $Prob\,(H)$ to occur; alternatively an outcome L is observed. If we now get some signal s on the outcome for H, we are interested in the conditional probability of H occurring, given the signal s has been received. Conditional probabilities are defined as

$$Prob\,(H|s) = \frac{Prob\,(H \cap s)}{Prob\,(s)}, \tag{A.9}$$

where $Prob\,(\cdot)$ denotes the probability of the outcome, $Prob\,(\cdot|s)$ the probability of an outcome conditional on observing the signal s, and $\cap$ denotes that both outcome and signal have to be observed. Similarly we have

$$Prob\,(s|H) = \frac{Prob\,(H \cap s)}{Prob\,(H)}. \tag{A.10}$$

Solving equations (A.9) and (A.10) for $Prob\,(H \cap s)$, and setting these equal, we get what is often referred to as *Bayes' Theorem*:

$$Prob\,(H|s) = \frac{Prob\,(s|H)\,Prob\,(H)}{Prob\,(s)}. \tag{A.11}$$

We call $Prob\,(H|s)$ the posterior probability, i. e. the probability after observing the signal s, while $Prob\,(H)$ is the prior probability as it denotes the initial belief for the likelihood of H occurring. If we set $s = H$, thus the signal observed indicates outcome H will be obtained, then $Prob\,(s|H)$ can be interpreted as the probability that the signal s is correct. Furthermore, we have that

$$Prob\,(s) = Prob\,(s|H)\,Prob\,(H) + Prob\,(s|L)\,Prob\,(L)\,. \tag{A.12}$$

Using the expression in equation (A.9), we see that the first term is equal to $Prob\,(H \cap s)$ and the second term is $Prob\,(L \cap s)$. Adding these two probabilities, means the probability that it is either the signal s is observed with H (first term) or with L (second term), and as H and L are the only possible outcomes, it must be the probability that the signal s is observed. Using these conditional probabilities, we can now also determine conditional expected values and other moments of the distribution.

If the distributions of two variables are normally distributed, then Bayesian learning allows us to obtain an explicit expression for the distribution of the posterior probability. Let us assume that $Y = X + \varepsilon$, where X has a normal distribution with mean μ and variance σ_X^2 and ε a normal distribution with mean zero and variance σ_ε^2. X and ε are independent of each other. We can interpret Y as a noisy signal of X; If X is the true outcome and we can only observe Y, then Y will reflect the value of X, but with an added fluctuation ε, which is commonly referred to as noise. We can now use our observation of Y to make inferences about the true value of X. We then can show that the distribution of X given Y, denoted $X|Y$ is a normal distribution with mean and variance

$$E[X|Y] = \mu + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_\varepsilon^2}(Y - \mu), \tag{A.13}$$
$$Var[X|Y] = \frac{1}{\frac{1}{\sigma_X^2} + \frac{1}{\sigma_\varepsilon^2}},$$

respectively. Thus in the case of having normal distributions, the analysis of expected values and variances can be conducted much more easily than in the general case, where no such analytical solutions are easily available.

A.5 Leibniz integral rule

In some instances, we need to optimize a function Π, where the decision variable is the boundaries of an integral. This can be the case, for example, if the payoff Ψ depends not only on this value V, but also a random variable s, and, in addition, the value V determines the range on which this function Π can be achieved. we denote the lower limit by $\underline{\lambda}(V)$ and the upper limit by $\overline{\lambda}(V)$. Let now

$$\Pi(V) = \int_{\underline{\lambda}(V)}^{\overline{\lambda}(V)} \Psi(V, s)\, ds \tag{A.14}$$

be our objective function. The Leibniz integral rule now specifies that the derivative of the function with respect to the value V is given by

$$\begin{aligned}\frac{\partial \Pi(V)}{\partial V} = {} & \frac{\partial \overline{\lambda}(V)}{\partial V}\Psi\left(V, \overline{\lambda}(V)\right) \\ & - \frac{\partial \underline{\lambda}(V)}{\partial V}\Psi\left(V, \underline{\lambda}(V)\right) \\ & + \int_{\underline{\lambda}(V)}^{\overline{\lambda}(V)} \frac{\partial \Psi(V, s)}{\partial V} ds.\end{aligned} \tag{A.15}$$

If either $\underline{\lambda}(V)$ or $\overline{\lambda}(V)$ are constants or $\pm\infty$, then the associated derivative in equation (A.15) is zero and the term can be eliminated.

A.6 Dynamic programming

When individuals have not only to make decisions that are optimal at a given point of time by choosing an optimal value of the control variables, u, but that these values have to be chosen optimal over a certain period of time, they have to determine an optimal time path for the variables. In many cases an additional problem arises, that the environment determining the outcome of this optimization problem changes through a changing state variable, x, which may be influenced by the control variables.

We define a function $I(x, u, t)$ which measures the payoff at a certain point of time, t. The aim of the individual now is to maximize the payoffs he receives over time, i.e. the control problem is given by

$$\max_{u(t)} J = \int_{t_0}^{t_1} I(x, u, t)dt, \tag{A.16}$$

where t_0 and t_1 denote the starting and end point of the considerations. The state variable changes according to the differential equation

$$\frac{\partial x}{\partial t} = f(x, u, t). \tag{A.17}$$

We define the solution to the control problem by $J^*(x, t)$ and call this the optimal performance function.

The principle of optimality now requires that regardless of the current state the remaining decisions have to be optimal. Therewith at point $t + \Delta t$ with state $x + \Delta x$ the optimal performance function has to be $J^*(x + \Delta x, t + \Delta t)$. We can now write the optimal performance function as

$$J^*(x, t) = \max_{u(t)}\{I(x, u, t)\Delta t + J^*(x + \Delta x, t + \Delta t)\}, \tag{A.18}$$

which is known as the *fundamental recurrence relation*. Here $I(x, u, t)\Delta t$ denotes the payoff in the interval $]t, t + \Delta t]$. Approximating the second term in brackets by a first order Taylor series around (x, t), we get

$$J^*(x + \Delta x, t + \Delta t) = J^*(x, t) + \frac{\partial J^*}{\partial x}\Delta x + \frac{\partial J^*}{\partial t}\Delta t, \tag{A.19}$$

which gives after inserting into equation (A.30) that

$$0 = \max_{u(t)}\{I(x, u, t)\Delta t + \frac{\partial J^*}{\partial x}\Delta x + \frac{\partial J^*}{\partial t}\Delta t\}. \tag{A.20}$$

Dividing by Δt and taking the limit $\Delta t \to 0$ we get with

$$\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \frac{\partial x}{\partial t} = f(x, u, t) \tag{A.21}$$

$$-\frac{\partial J^*}{\partial t} = \max_{u(t)}\left\{I(u, x, t) + \frac{\partial J^*}{\partial x} f(x, u, t)\right\}. \tag{A.22}$$

This partial differential equation is known as the Bellman equation. Solving this equation will give the optimal performance function, given boundary conditions.

A.7 Differential equations

In some dynamic models, we might obtain an equation that consists of a function as well as its derivative, which is called a differential equation. We are often then interested in solving for this function. If we have a variable x and its function $f(x)$, then a differential equation will take the form

$$\frac{\partial f(x)}{\partial x} = a(x) f(x) + b(x), \tag{A.23}$$

where a and b are coefficients, which themselves might depend on the variable x.

We can now solve this differential equation in the case of $b(x) = 0$ by noting that the differential equation can then be rewritten as $\frac{\frac{\partial f(x)}{\partial x}}{f(x)} = a(x)$ and hence

$$f(x) = Ce^{\int a(x)}, \tag{A.24}$$

with an arbitrary constant C. Dividing the original differential equation in equation (A.23) by the solution in equation (A.24), we obtain

$$C\frac{\partial f(x)}{\partial x} e^{-\int a(x)dx} = Ca(x) f(x) e^{-\int a(x)dx} + Cb(x) e^{-\int a(x)dx}. \tag{A.25}$$

Using that $\frac{e^{-\int a(x)dx}}{\partial x} = -a(x) e^{-\int a(x)dx}$ and the product rule then gives us $\frac{\partial f(x)e^{-\int a(x)dx}}{\partial x} = \frac{\partial f(x)}{\partial x} e^{-\int a(x)dx} - a(x) f(x) e^{-\int a(x)dx}$, hence we can rewrite equation (A.25) as

$$\frac{\partial f(x) e^{-\int a(x)dx}}{\partial x} = b(x) e^{-\int a(x)dx}. \tag{A.26}$$

This now gives us the full solution to our differential equation as

$$f(x) = Ce^{\int a(x)} + e^{\int a(x)} \int b(x) e^{\int a(x)} dx. \tag{A.27}$$

The constant C can now be determined by using a boundary condition, such as an initial or final value.

A.8 Statistical tools

A.8.1 Conditional moments

Let us assume we have two variables, X and Y, and have obtained information on the value of Y. We now want to predict the value of X given our information on Y and limit ourself to a linear prediction. Hence, we seek to find the expected value of X, given our observation of Y, such that

$$E[X|Y] = \alpha_0 + \alpha_1 Y. \tag{A.28}$$

We now seek to minimize the prediction error $\mathfrak{E} = E\left[(X - E[X|Y])^2\right]$ over the parameters α_0 and α_1. The first order conditions after inserting for $E[X|Y]$ from equation (A.28) are given by

$$\begin{aligned}\frac{\partial \mathfrak{E}}{\partial \alpha_0} &= 2\alpha_0 - 2E[X] + 2\alpha_1 E[Y] = 0, \\ \frac{\partial \mathfrak{E}}{\partial \alpha_1} &= 2\alpha_1 E\left[Y^2\right] - 2E[XY] + 2\alpha_0 E[Y] = 0.\end{aligned} \tag{A.29}$$

These two equations can be solved for

$$\begin{aligned}\alpha_0 &= E[X] - \alpha_1 E[Y], \\ \alpha_1 &= \frac{E[XY] - E[X]E[Y]}{E\left[X^2\right] - E[X]^2} = \frac{Cov[X,Y]}{Var[X]},\end{aligned} \tag{A.30}$$

where $E[\cdot]$ denotes the expected value, $Var[\cdot]$ the variance, and $Cov[\cdot,\cdot]$ the covariance. Inserting this result into equation (A.28), we get for our prediction of X, given the value for Y, that

$$E[X|Y] = E[X] + \frac{Cov[X,Y]}{Var[X]}(Y - E[Y]). \tag{A.31}$$

From equation (A.28) we also obtain that

$$\begin{aligned}Var(E[X|Y]) &= \alpha_1^2 Var[Y] \\ &= \frac{Cov[X,Y]^2}{Var[X]^2} Var[Y].\end{aligned} \tag{A.32}$$

This expression represents the ex-ante variance of the conditional expected value, that prior to observing the value for Y, we know the conditional expected value which we will obtain, has a variance as indicated here.

A.8.2 Order statistics

The M^{th} largest value out of N total values, is called the M^{th} order statistic. The cumulative distribution function is given by

$$F_M(V) = \sum_{i=M}^{N} \binom{N}{M} F(V)^i (1 - F(V))^{N-i}, \tag{A.33}$$

where $F(\cdot)$ denotes the cumulative distribution function and $\binom{\cdot}{\cdot}$ the binomial coefficient. The order statistic is determined by M observations being below V and $N - M$ above the value of V. $F(V)$ then denotes the probability of making an observation below V. Seeking the M^{th} largest value, we require to have at least M observations above V, with the remainder below.

In general the resulting distribution F_M cannot be determined analytically, except for the case that V itself is uniformly distributed on the interval $[0; 1]$. In this case the M^{th} largest value is distributed like a beta distribution with parameters M and $N + 1 - M$, Such a distribution has a mean of $\frac{M}{M+N+1-M} = \frac{M}{N+1}$ and a variance of $\frac{M(N+1-M)}{(M+N+1-M)^2(M+N+1-M+1)} = \frac{M(N+1-M)}{(N+1)^2(N+2)}$. For other distributions, no further results are generally available analytically.

Appendix B

Economic models

B.1 Nash bargaining

Consider two individuals negotiating a price, for example a bank and borrower negotiating the loan rate, a bank and depositor negotiating the deposit rate, or an investment bank negotiating its fee for providing advice with a client. Both parties in this negotiation will have the option of terminating the negotiation and falling back on alternative solutions, the so-called outside option, which we assume provides utility $\hat{u}_i$ to individual i. Such alternative solution might be to not obtain a loan, not deposit monies or not seek advice, but could also include taking their custom to another bank. The utility of individual i derived from agreeing a price p is denoted by $u_i(p)$. The surplus from concluding the negotiation is then $u_i(p) - \hat{u}_i$. In many models we will assume that either the bank or their client have market power; in this case the price would be set such that $u_i(r) - \hat{u}_i = 0$ for the individual lacking market power.

An alternative modelling assumption is that both individuals seek to share the joint surplus fairly. Nash (1953) proposes that maximizing

$$\mathcal{L} = (u_1(p) - \hat{u}_1)(u_2(p) - \hat{u}_2) \tag{B.1}$$

is optimal and both parties obtain the same surplus beyond their outside option. The maximization of this expression is called the *Nash bargaining* solution. The first order condition is

$$\frac{\partial \mathcal{L}}{\partial p} = \frac{\partial u_1(p)}{\partial p}(u_2(p) - \hat{u}_2) + \frac{\partial u_2(p)}{\partial p}(u_1(p) - \hat{u}_1) = 0. \tag{B.2}$$

The joint surplus is $\mathcal{S} = (u_1(p) - \hat{u}_1) + (u_2(p) - \hat{u}_2)$ and hence the maximum is obtained if

$$\frac{\partial \mathcal{S}}{\partial p} = \frac{\partial u_1(p)}{\partial p} + \frac{\partial u_2(p)}{\partial p} = 0. \tag{B.3}$$

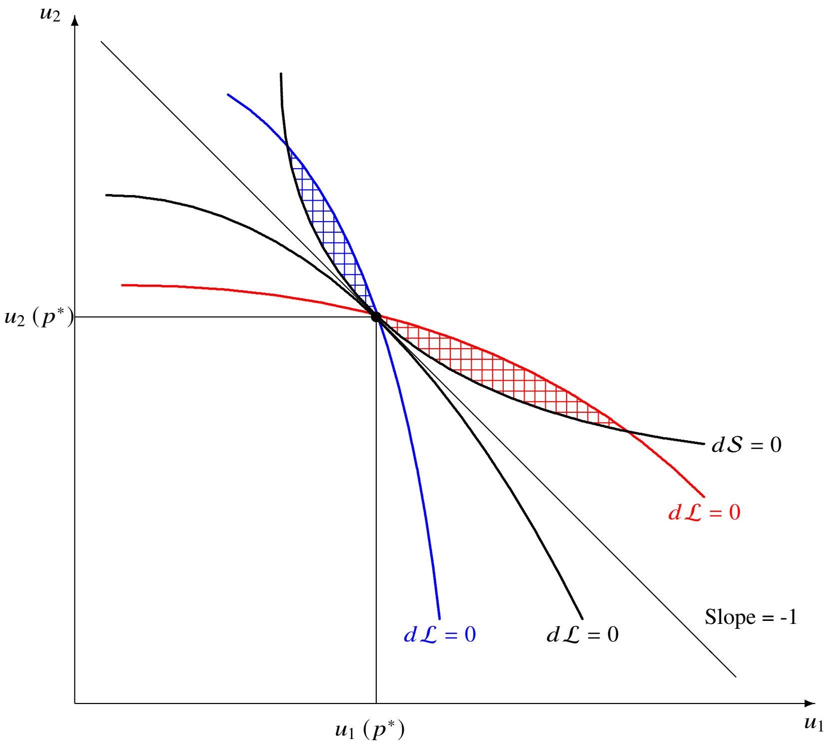


Fig. B.1: Optimality of the Nash bargaining solution

We now see from equation (B.2) that if $u_1(p) - \hat{u}_1 = u_2(p) - \hat{u}_2$, i. e. the surpluses both parties obtain are identical, then the first order conditions from maximizing $\mathcal{L}$ and $\mathcal{S}$ are identical, therefore maximizing equation (B.1) gives the outcome that maximizes the joint surplus.

We are now left to show that an unequal distribution of surpluses is not optimal. The first order conditions for maximizing the objective functions $\mathcal{L}$ and $\mathcal{S}$, equations (B.2) and (B.3) can be rewritten at the optimal price p^* as

$$\frac{\partial u_1(p^*)}{\partial u_2(p^*)} = -1, \tag{B.4}$$
$$\frac{\partial u_1(p^*)}{\partial u_2(p^*)} = -\frac{u_1(P^*) - \hat{u}_1}{u_2(p^*) - \hat{u}_2}.$$

The first equation represents the total differential $d\mathcal{L} = 0$ at the optimal price p^* and the second the total differential of $d\mathcal{S} = 0$. Suppose now the surplus is not allocated equally, thus for the second equation we have $\frac{\partial u_1(p^*)}{\partial u_2(p^*)} \neq -1$. Figure B.1 illustrates this situation. The black line illustrates the situation in which the Nash bargaining solution is optimal. Now suppose that the surplus is not allocated equally, in which case the slopes of the curve $d\mathcal{L} = 0$ will be either below or above -1 at the point where the slope of $d\mathcal{S} = 0$ is equal to -1, as indicated by the red and blue lines. We now see that in the area indicated by the red and blue hatched lines, respectively, the utility of both parties could be increased, thus the solution p^* would no longer be optimal, contradicting the optimality of p^* from maximizing the joint surplus. Hence the surplus must be divided equally and the Nash bargaining solution is optimal.

B.2 Tacit collusion

Let us consider two individuals i setting prices strategically. If both set high prices and do not aggressively compete with each other, they will both make high profits, Π^i_{CC}. This is often referred to as 'cooperate'. Alternatively they could both compete with each other and set lower prices such that their profits are $\Pi^i_{DD} < \Pi^i_{CC}$, something commonly refereed to as 'defect'. Another scenario is that only one of the two individuals seeks to compete by undercutting their competitor marginally, taking a larger share of the market and thus making higher profits $\Pi^i_{DC} > \Pi^i_{CC}$. Taking a higher markets share, leaves less or nothing for the competitor and we assume that $\Pi^i_{CD} < \Pi^i_{DD}$. Hence we have $\Pi^i_{DC} > \Pi^i_{CC} > \Pi^i_{DD} > \Pi^i_{CD}$. This is the situation of the prisoner's dilemma and the only Nash equilibrium is for both individuals to defect and hence obtain Π^i_{DD}, even though cooperating and obtaining Π^i_{CC} would be preferred by both individuals.

Let us now assume that these two individuals interact repeatedly over time. If they always cooperate, they would obtain Π^i_{CC} in each time period and with a discount factor ρ_i, the total value obtained would be $\frac{\Pi^i_{CC}}{1-\rho_i}$. If one of the individuals would defect, they would obtain Π^i_{DC} for that time period. We then assume that in reaction to the defection, the other individual will from then onwards defect, too. Thus from the subsequent time periods, both individuals obtain Π^i_{DD}, giving total value $\Pi^i_{DC} + \rho_i \frac{\Pi^i_{DD}}{1-\rho_i}$. Individuals will cooperate if $\frac{\Pi^i_{CC}}{1-\rho_i} \geq \Pi^i_{DC} + \rho_i \frac{\Pi^i_{DD}}{1-\rho_i}$, or

$$\rho_i \geq \frac{\Pi^i_{DC} - \Pi^i_{CC}}{\Pi^i_{DC} - \Pi^i_{DD}}. \tag{B.5}$$

Hence, if individuals are sufficiently patient, that is have a sufficiently high discount factor, the cooperative outcome is an equilibrium. The Nash equilibrium of the prisoner's dilemma of course remains an equilibrium, but will be inferior for both individuals. As both individuals charge higher than competitive prices, which would be the Nash equilibrium of the prisoner's dilemma, and make higher profits, this outcome is of referred to as *tacit collusion* or *implicit collusion*, even though no agreement on the behaviour has been reached; it is purely driven by economic incentives.

B.3 Risk aversion

Individuals prefer an investment with a certain outcome to an investment that has the same expected outcome but whose realised outcomes are uncertain. This is referred to as *risk aversion*. Pratt (1964) has developed a measure that allows us to quantify the strength of this risk aversion, also known as the Arrow-Pratt measure of risk aversion.

Let us assume that the outcome of an investment is V, which is random with some probability distribution and has expected value $E[V]$. If individuals are risk averse, they would prefer to receive a certain amount $E[V]$ to the random amount V. With utility function $u(\cdot)$, we can now determine a value π that makes a consumer indifferent between these two choices. Hence we need the expected utilities to be equal:

$$E[u(E[V]-\pi)] = u(E[V]-\pi) = E[u(V)], \tag{B.6}$$

where the first equality arises from the observation that no random terms are included in the arguments of the utility function.

The term $E[V]-\pi$ is also called the *cash equivalent* or *certainty equivalent* of V and π is called the *risk premium*. Approximating the right-hand side of equation (B.6) by a second order Taylor series expansion around $E[V]$, we get

$$\begin{aligned} E[u(V)] &= E\left[u(E[V]) + \frac{\partial u(E[V])}{\partial V}(E[V])(V-E[V])\right. \\ &\qquad \left. + \frac{1}{2}\frac{\partial^2 u(E[V])}{\partial V^2}(V-E[V])^2\right] \\ &= u(E[V]) + \frac{1}{2}\frac{\partial^2 u(E[V])}{\partial V^2} Var[V], \end{aligned} \tag{B.7}$$

where the second term in the original approximation vanishes as $E[V-E[V]]=0$ and for the final expression we used that $E\left[(V-E[V])^2\right] = Var[V]$ is the definition of the variance.

In a similar way we can approximate the left-hand side of equation (B.6) by a first order Taylor series around $E[V]$ and get

$$u(E[V]-\pi) = u(E[V]) + \frac{\partial u(E[V])}{\partial V}\pi. \tag{B.8}$$

Inserting equations (B.7) and (B.8) into equation (B.6), we get after solving for the risk premium π that

$$\pi = -\frac{1}{2}\frac{\frac{\partial^2 u(E[V])}{\partial V^2}}{\frac{\partial u(E[V])}{\partial V}} Var[V]. \tag{B.9}$$

We now define

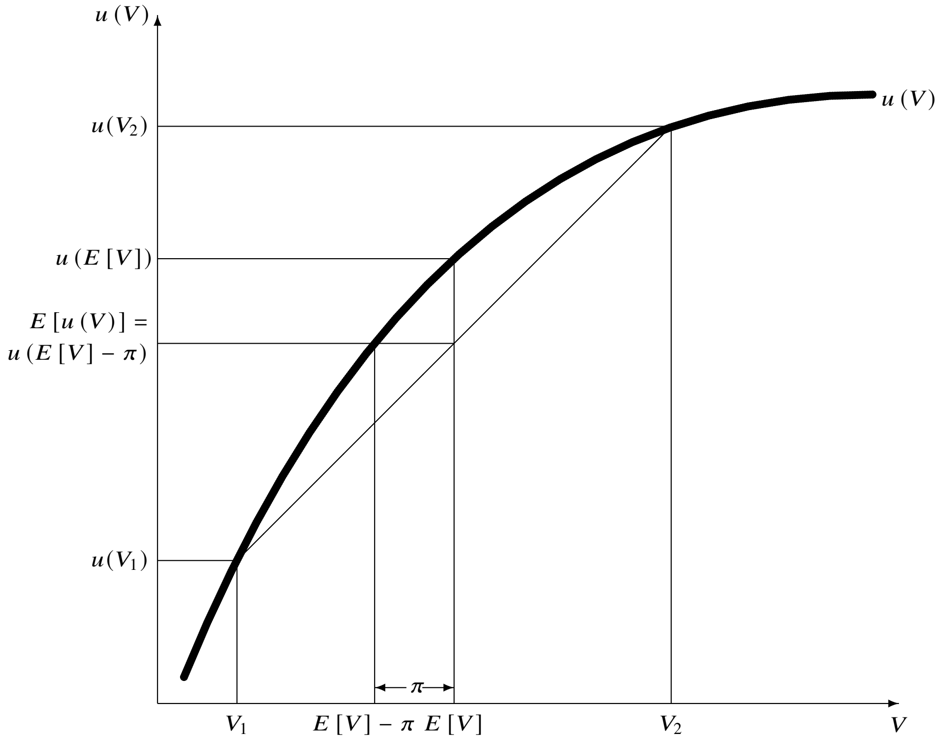


Fig. B.2: The Arrow-Pratt measure of risk aversion

$$z = -\frac{\frac{\partial^2 u(E[V])}{\partial V^2}}{\frac{\partial u(E[V])}{\partial V}} \tag{B.10}$$

as the *absolute local risk aversion*. This can be justified by noting that the risk premia π has to be larger, the more risk averse an individual is and the higher the risk. The risk is measured by the variance of V, $Var[V]$, hence the other term in equation (B.9) can be interpreted as risk aversion. By inserting equation (B.10) into equation (B.9) we get

$$\pi = \frac{1}{2} z Var[V]. \tag{B.11}$$

If we assume that individuals are risk averse we need $\pi > 0$, implying $z > 0$. It is reasonable to assume positive marginal utility, i. e. $\frac{\partial u(E[V])}{\partial V} > 0$, which then implies that we require $\frac{\partial^2 u(E[V])}{\partial V^2} < 0$ to obtain $z > 0$. This property of diminishing marginal utility is a common assumption about utility functions in economic theories. The assumption of risk aversion is therefore in line with the standard assumptions about utility functions in microeconomic theory.

These conditions imply a concave utility function, whose concavity (radius) is determined by the risk aversion. Figure B.2 visualizes these findings for the simple case of two possible outcomes, V_1 and V_2, having equal probability of occurrence.

We can now insert equation (B.11) into equation (B.6) and obtain

$$E\left[u\left(V\right)\right]=u\left(E\left[V\right]-\pi\right)=u\left(E\left[V\right]-\frac{1}{2}zVar\left[V\right]\right). \tag{B.12}$$

To simplify the analysis even further, as $\frac{\partial u(E[V])}{\partial V}>0$, we can use $E\left[V\right]-\frac{1}{2}zVar\left[V\right]$ as our objective function.

B.4 Monopolistic competition

Goods are often similar and close substitutes, but not identical. There might be small differences in certain qualities such as its design or the combination of features it offers that make a consumer to prefer one over the other. We assume that each consumer has a preference for these qualities and that each company produces a good with a given set of such qualities. The more differences there are between the preferences of the consumer and the good offered by the company, the less these goods are worth to the consumer. With each unit of difference, we assume that the value of the good, V, reduces by c.

Consumers are assumed to be evenly distributed on a line of length 1, where the location indicates their preferences for a specific quality and each company is located in a single position on this line, representing the qualities of their good. Consumers located at the some point as the company would be offered a good that exactly meets their preferences and the further the consumer is located from the company, the lower the value to the consumer would be.

We now investigate the competition between companies in such a market for differentiated goods, in terms of prices set but also the location of the companies themselves.

B.4.1 Hotelling model

Hotelling (1929) assumes that consumers are located on a straight line of length 1 and indicates their location by the distance d from the start of the line, hence $0\leq d\leq 1$. There are two companies offering a good for which they charge a price P_i and they are located in positions d_1 and d_2, respectively. Without loss of generality we assume that $0\leq d_1\leq d_2\leq 1$.

The value of the good to a consumer buying from company i will be

$$\Pi_C^i=V-P_i-c\left|d-d_i\right|, \tag{B.13}$$

where $|\cdot|$ denotes the absolute value and hence the distance of the consumer to the company. The consumer who is indifferent between buying from the two companies will be located at the point where $\Pi_C^1=\Pi_C^2$. As $d_1\leq d_2$, we will have that

$|d - d_1| = d - d_1$ and $|d - d_2| = d_2 - d$ for any $d_1 \leq d \leq d_2$. Inserting this into equation (B.13), we obtain the location of the indifferent consumer as

$$d^* = \frac{d_1 + d_2}{2} + \frac{P_2 - P_1}{2c}. \tag{B.14}$$

If $d < d_1$ or $d > d_2$, the consumer will choose the nearest company, company 1 and company 2, respectively.

Company 1 will now obtain the business of all consumers $d \leq d^*$, which given the uniform distribution of consumers, gives us a demand for their good of d^*. The demand for company 2 will be the remainder of consumers, thus $1 - d^*$, and the profits of the companies, assuming they have no further costs, are thus

$$\begin{aligned} \Pi_B^1 &= P_1 d^* \\ &= \frac{cP_1 (d_1 + d_2) + P_1 P_2 - P_1^2}{2c}, \\ \Pi_B^2 &= P_2 (1 - d^*) \\ &= \frac{cP_2 (2 - d_1 - d_2) + P_1 P_2 - P_2^2}{2c}. \end{aligned} \tag{B.15}$$

Both companies maximize their respective profits, assuming that the location has been chosen previously, and the first order conditions are given by

$$\begin{aligned} \frac{\partial \Pi_C^1}{\partial P_1} &= \frac{c (d_1 + d_2) + P_2 - 2P_1}{2c} = 0, \\ \frac{\partial \Pi_C^2}{\partial P_2} &= \frac{c (2 - d_1 - d_2) + P_1 - 2P_2}{2c} = 0, \end{aligned} \tag{B.16}$$

which can be solved for

$$\begin{aligned} P_1 &= \frac{c}{3} (2 + d_1 + d_2), \\ P_2 &= \frac{c}{3} (4 - d_1 - d_2). \end{aligned} \tag{B.17}$$

Inserting these prices into equations (B.14) and (B.15), we get the company profits as

$$\begin{aligned} \Pi_B^1 &= \frac{c}{18} (2 + d_1 + d_2)^2, \\ \Pi_B^2 &= \frac{c}{18} (4 - d_1 - d_2)^2. \end{aligned} \tag{B.18}$$

As both companies are inherently equal, we are only looking at symmetric equilibria in which $\Pi_B^1 = \Pi_B^2$ and thus companies have no desire to switch positions. Solving for this relationship using equations (B.18), we easily get that $d_1 + d_2 = 1$. Thus all locations that are symmetric from the end of the line, are giving symmetric equilibria with $P_1 = P_2 = c$ and $\Pi_B^1 = \Pi_B^2 = \frac{c}{2}$.

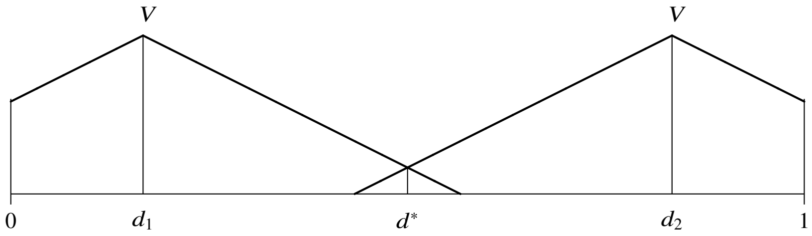


Fig. B.3: The Hotelling model

With us being interested in the competition between companies for consumers, we will normally choose $d_1 = 0$ and $d_2 = 1$ such that all consumers are subjected to the competitive forces between the two companies. Figure B.3 illustrates the Hotelling model with arbitrarily chosen (symmetric) locations of the two companies.

B.4.2 Salop's circle

Rather than assuming that consumers and companies are located on a straight line of length 1, Salop (1979) assumes the line to form a circle, thus the start and end of the line are connected. In addition, there are now $N \geq 2$ companies offering differentiated goods and the distance between bank i and companies $i+1$ is denoted d_i, where $\sum_{i=1}^{N} d_i = 1$. Of course, if $i = N$, then we set $i+1 = 1$ to complete the circle. The distance of a consumer to company i is denoted by $\hat{d}_i$. We have $\hat{d}_i + \hat{d}_{i+1} = d_i$ as the distance of a consumer located between companies i and $i+1$ to company i and company $i+1$ together, which must equal the distance between the companies. Figure B.4 illustrates this model.

The profits to a consumer selecting different companies are then given by

$$\begin{aligned} \Pi_C^i &= V - P_i - c\hat{d}_i, \\ \Pi_C^{i+1} &= V - P_{i+1} - c\left(d_i - \hat{d}_i\right), \\ \Pi_C^{i-1} &= V - P_{i-1} - c\left(d_{i-1} - \hat{d}_i\right). \end{aligned} \tag{B.19}$$

Consumers being indifferent between companies i and $i+1$ require $\Pi_C^i = \Pi_C^{i+1}$ and being indifferent between companies i and $i-1$ requires $\Pi_C^i = \Pi_C^{i-1}$. This implies that the distance to banks $i+1$ and $i-1$, respective, must be at least

$$\begin{aligned} \hat{d}_{i+1}^* &= \frac{cd_i + P_{i+1} - P_i}{2c}, \\ \hat{d}_{i-1}^* &= \frac{cd_{i-1} + P_{i-1} - P_i}{2c} \end{aligned} \tag{B.20}$$

for the consumer to prefer company i.

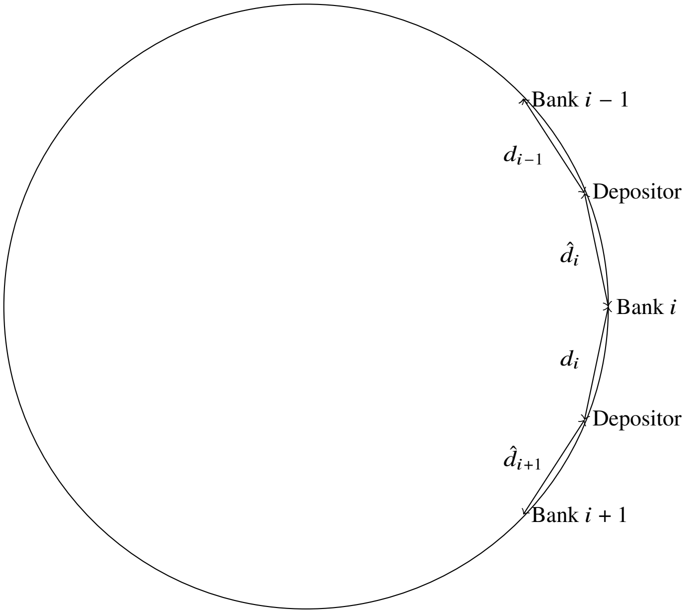


Fig. B.4: Illustration of Salop's circle

The company profits are then given by the demand of consumers located closer than $\hat{d}^*_{i-1}$ to company $i-1$ and closer than $\hat{d}^*_{i+1}$ to company $i+1$, hence

$$\Pi_B^i = \left(\hat{d}^*_{i-1} + \hat{d}^*_{i+1}\right) P_i = \frac{c\left(d_i + d_{i-1}\right) + P_{i+1} + P_{i-1} - 2P_i}{2c} P_i, \tag{B.21}$$

such that for given locations, the optimal price P_i has to fulfill the first-order condition

$$\frac{\partial \Pi_B^1}{\partial P_i} = \frac{c\left(d_i + d_{i-1}\right) + P_{i+1} + P_{i-1} - 4P_i}{2c} = 0. \tag{B.22}$$

This solves for

$$P_i = \frac{1}{4}\left(c\left(d_i + d_{i-1}\right) + P_{i+1} + P_{i-1}\right). \tag{B.23}$$

As this condition is identical for all companies, we again look only at symmetric equilibria and assume that all prices are equal, thus $P = P_{i-1} = P_i = P_{i+1}$, solving for

$$P = \frac{c}{2}\left(d_i + d_{i-1}\right). \tag{B.24}$$

Inserting this result into equations (B.20) and (B.21), we find that $\Pi_B^i = \frac{c}{4}\left(d_i + d_{i-1}\right)^2$ and requiring all companies to have equal profits to avoid swapping locations as they are inherently identical, we get from $\Pi_B^i = \Pi_B^{i+1}$ that $d_{i-1} = d_{i+1}$. Thus, the distance between any companies $i-1$ and i is the same as between companies $i+1$ and $i+2$.

While the distance between companies i and $i+1$ could be different, the symmetry would imply that there is no reason for this to be so, hence all distances between companies are identical and $d_i = \frac{1}{N}$. This implies then from equations (B.23) and (B.24) that $P_i = \frac{c}{N}$ and $\Pi_B^i = \frac{c}{N^2}$.

We note that in this model for $N = 2$ companies, the prices the companies charge for the good is identical to the case of consumers and companies being located on a line, while the profits are lower as comparison with the final results of the Hotelling model in appendix B.4.1 shows. The reason is that in the circular model here, companies face competition from both sides, while in the linear model this competition is only in the direction of the other company.

Appendix C

Credit risk assessment

Credit risk is the possibility of a bank making losses from borrowers not repaying their loans or a counterparty in a transaction not meeting their obligations in full. Such a situation is referred to as default. If the probability of default is given by p and the loss given default, the losses that are made if default occurs, are denoted λ, then for an exposure to such a risk of size L, we have the expected losses as

$$\Pi_L = p\lambda L. \tag{C.1}$$

Rather than loss given default, it also common to use the recovery rate $1 - \lambda$, which denotes the fraction of the exposure that the lender is able to obtain in the case of a default.

In terms of loans, there is a simple relationship between the loan rate r_L that should be charged, and the default rate. The loan of size L is supposed to be repaid including interest r_L, giving us an exposure of $(1 + r_L) L$. This is repaid with probability p, while with probability $1 - p$, default occurs and the lender obtains a fraction $1 - \lambda$ of their exposure. This gives us an expected repayment of

$$\begin{aligned}\hat{\Pi}_L &= p (1 + r_L) L + (1 - p) (1 - \lambda) (1 + r_L) L \\ &= (p\lambda + (1 - \lambda)) (1 + r_L) L.\end{aligned} \tag{C.2}$$

An investment of the initial loan L into a risk-free asset yielding a return of r would give us $(1 + r) L$. In equilibrium, the returns from two investments must be equal, yielding

$$1 + r_L = \frac{1 + r}{p\lambda + (1 - \lambda)}. \tag{C.3}$$

In most practical cases we find that upon default the bank will not obtain any payments from the borrowers once the liquidation process is completed, thus $\lambda = 1$, and hence upon default the entire loan is lost. This simplifies equation (C.3) to $1 + r_L = \frac{1+r}{p}$ and if we use the approximation that for small values of x we have $\ln (1 + x) \approx x$, we can rewrite this expression as

$$r_L \approx r + p. \tag{C.4}$$

The loan rate r_L is the risk-free rate r adjusted by the probability of default p. Hence the probability of default is a key parameter that credit risk assessment seeks to determine. Different methods of obtaining the probability of default have been proposed and many banks generally apply their own methodology, informed by data from past lending. The most common theoretical approaches are based on the idea of Merton (1974).

C.1 The Merton model

The total assets of a company at time t, A_t, are financed by debt L and equity E_t, such that $A_t = L + E_t$. Default occurs if at maturity of the loan, T, the assets are not sufficient to repay the debt, thus if $A_T < L$, ignoring for convenience the accumulated interest in the exposure of the bank. The probability of default is then given by $p = \text{Prob}\,(A_T < L)$. We now assume that assets grow at an expected rate of μ_A and the growth rate has a volatility of σ_A. We can then write the development of the asset price over time as a stochastic process. The change of the asset value over Δt time periods is ΔA_t and the growth rate therefore $\frac{\Delta A_t}{A_t}$. We thus have $\frac{\Delta A_t}{A_t} = \mu_A \Delta t + \varepsilon_t$, where ε_t is normally distributed with a mean of zero and variance of σ_A^2. If we take the time period becoming ever smaller, this approaches continuous time and we obtain the Brownian motion

$$\frac{dA_t}{A_t} = \mu_A dt + \sigma_A dz_t, \tag{C.5}$$

where dz_t represents a standard Wiener process, that only adds noise of variance 1 to the realised return. We can now use statistics to obtain that

$$p = \text{Prob}\,(A_T < L) = \Phi\left(\frac{\ln \frac{A_t}{L} + \left(\mu_A - \frac{1}{2}\sigma_A^2\right)(T - t)}{\sigma_A \sqrt{T - t}}\right), \tag{C.6}$$

where $\Phi\,(\cdot)$ denotes the cumulative standard normal distribution. This probability of default is the probability of default that by the end of time period T, i. e. in $T - t$ time periods, the company cannot repay the loan. If we assume that the actual default can happen at any point until time period T, then the probability of default per time period is $\frac{p}{T-t}$.

It has to be noted that in order to apply this approach, the volatility of the assets, σ_A, as well as their growth rate, need to be known. While a connection to the volatility of the observable volatility of stock prices can be made, this will only be of assistance in the determination of the credit risk from listed companies that have sufficient liquidity to ensure that prices are efficient. For the vast majority of lending decisions, no stock prices exist and therefore estimates of the asset volatility need to be obtained. This might be possible to be obtained from accounting data or other sources.

The expression $\frac{\ln \frac{A_t}{L} + \left(\mu_A - \frac{1}{2}\sigma_A^2\right)(T-t)}{\sigma_A \sqrt{T-t}}$ is also referred to as distance-to-default. With $\Phi(\cdot)$ denoting the cumulative standard normal distribution, the distance-to-default can be interpreted as the number of standard deviations the assets need to fall to trigger default. The distance to default represents the same information as the probability of default and is only a different way to express this information.

C.2 The KMV model

Building on the Merton model, the risk consultancy KMV has simplified the distance-to-default expression. Empirically we find that $\mu_A - \frac{1}{2}\sigma_A^2 \approx 0$ and hence the distance-to-default simplifies to $\frac{\ln \frac{A_t}{L}}{\sigma_A \sqrt{T-t}}$. If we only consider the probability of default during the next time period, thus setting $T = t + 1$, we only need to consider the short-term liabilities $\hat{L}$ that become due within the next year and can replace this expression in the distance to default. We can now make a further approximation that $\ln \frac{A_t}{\hat{L}} = \ln\left(1 + \left(\frac{A_t}{\hat{L}} - 1\right)\right) \approx \frac{A_t}{\hat{L}} - 1 = \frac{A_t - \hat{L}}{\hat{L}}$ and obtain

$$\frac{\ln \frac{A_t}{\hat{L}}}{\sigma_A} \approx \frac{A_t - \hat{L}}{\sigma_A \hat{L}}. \tag{C.7}$$

Using this distance to default and then applying $p = \Phi\left(\frac{A_t - \hat{L}}{\sigma_A \hat{L}}\right)$ from the Merton model, or a modified distribution that fits the actual observations better than the standard normal distribution, has the advantage that short-term defaults are considered, rather than long-term defaults that allow companies to make management decisions to avoid such default. This approach is widely used by banks.

C.3 The Vasicek model

What is unsatisfactory in the Merton model, and by extension, the KMV model, is that defaults are driven purely by a stochastic process of the assets of a company. Often we can, however, identify determinants that will affect default rates, such a macroeconomic conditions, industry specific developments or even company specific factors. As in Vasicek (2002), let us therefore assume that the value of the assets of a company, A_t are determined by

$$\ln A_t = \mu_A + \sigma_A \Psi_t, \tag{C.8}$$

where Ψ_t represents such factors and we assume that this variable is normally distributed with mean zero and variance 1. The logarithm of the assets thus has a mean of μ_A and a variance of σ_A^2. With a loan of L, neglecting interest payments in the exposure, we observe a default if at any time the assets are not sufficient to cover the loan amount, hence we require $A_t < L$. Thus

$$\begin{aligned} p &= \text{Prob}\,(A_t < L) = \text{Prob}\,(\ln A_t < \ln L) \\ &= \text{Prob}\left(\Psi_t < \frac{\ln L - \mu_A}{\sigma_A}\right) \\ &= \Phi\left(\frac{\ln L - \mu_A}{\sigma_A}\right) \end{aligned} \tag{C.9}$$

and default is caused by the variable Ψ_t being sufficiently low. $\Phi(\cdot)$ denotes the cumulative normal distribution as we had assumed that Ψ_t is standard normally distrbuted. This variable Ψ_t is used here to represent the underlying factor driving the default of a company. Let us now assume that

$$\Psi_t = \rho \Upsilon_t + \sqrt{1-\rho^2}\varepsilon_t, \tag{C.10}$$

where Υ_t is a normalised variable that is a common factor for defaults to occur in an economy and ε_t is an idiosyncratic factor that is unique and unpredictable to the company considered. We assume Υ_t and ε_t to be standard normally distributed and to be independent of each other. In this case we easily see that the Ψ_t is standard normally distributed as we had assumed in equation (C.9). Let us now assume that we know the value of Υ from observation. Such a factor might also be an aggregate of a variety of factors that we have considered to be relevant for the default of companies; banks will usually employ models that use a variety of factors to determine such values and obtain a credit score like Υ_t. Inserting equation (C.10) into equation (C.9), we obtain

$$\begin{aligned} \hat{p} &= \text{Prob}\left(\varepsilon_t < \frac{\frac{\ln L - \mu_A}{\sigma_A} - \rho \Upsilon_t}{\sqrt{1-\rho^2}}\right) \\ &= \Phi\left(\frac{\frac{\ln L - \mu_A}{\sigma_A} - \rho \Upsilon_t}{\sqrt{1-\rho^2}}\right) \\ &= \Phi\left(\frac{p - \rho \Upsilon_t}{\sqrt{1-\rho^2}}\right). \end{aligned} \tag{C.11}$$

We can now interpret p as the average probability of default of the company, and $\hat{p}$ as the probability of default adjusted to the current economic conditions. A value of $\Upsilon_t = 0$ corresponds to this average value, given that we assumed this variable to have a mean of zero. The actual probability of default $\hat{p}$ will be higher than p, because the risk of the common factor Υ_t to move adversely makes a default more likely than if default was purely driven by an idiosyncratic stochastic element.

Appendix D

Security valuation

Securities give the owner the right to future payments as specified in the conditions of the security. The challenge in determining the value of such a security is to firstly estimate these future payments, and then to use an appropriate discount rate to determine the present value of these future payments. This methodology, also referred to as the discounted cash flow methodology is the theoretically correct methodology to use, but practitioners have developed alternative methods as outlined in Rosenbaum & Pearl (2009, Chs. 1-3).

D.1 Discounted cash flow

A security promises to pay the expected amount of C_t at time t. The present value of such future payments will then be

$$V_0 = \sum_{t=0}^{+\infty} \rho^t C_t, \tag{D.1}$$

where $\rho \leq 1$ denotes the discount factor. If the future payments were guaranteed, then the risk-free rate r would be used to discount future payments and we set $\rho = \frac{1}{1+r}$. Commonly, though, payments are not guaranteed and we have to form expectations about these payments C_t. This is most prominent with shares, where C_t are the dividends the company will pay in the future. As these dividends are not guaranteed, the resulting risk for any deviations from the expectations needs to be compensated; this is achieved by applying a lower discount factor. It is common to use the Capital Asset Pricing Model (CAPM) to determine appropriate discount factors by applying the required rate of return obtained from this model, $\mu = r + \beta(\mu_M - r)$, and set $\rho = \frac{1}{1+\mu}$. Here μ_M denotes the market return and $\beta = \frac{\sigma_{iM}}{\sigma_M^2}$ the systematic risk of the shares, calculated as the ratio of the covariance of the stock with the market, σ_{iM}, and the market variance, σ_M^2.

Expectations about future dividends will be based on information the investment banker has about the prospects of the company. A common simplification is to assume that dividends remain constant, $C_t = C$, resulting in equation (D.2) to simplify to $V_0 = \frac{C}{1-\rho}$, Alternatively, a constant growth rate of g is assumed, such that $C_t = (1+g)\,C_{t-1}$. This assumption, known as the Gordon Growth Model, gives us a value of $V = \frac{1+r}{r-g}C_0$. Of course, more specific estimates can, and commonly are, used by investment banks when valuing companies. They will in particular take into account any synergies arising from mergers, but also their costs; for IPOs they would determine the impact the additional capital raised has due to the increased investment the company can make.

Discounted cash flow analysis is not only applicable to the valuation of shares, but can also be applied to bonds. The regular payments of a bond are the coupons C and the last payment at maturity T consists of this coupon payment and the repayment of the bond, L. Thus equation (D.2) becomes

$$V_0 = \sum_{t=0}^{T} \rho^t C + \rho^T L. \tag{D.2}$$

The appropriate discount factor will have to take into account the default risk of a bond. With a probability of default p and a risk-free rate r, we can set $\rho = \frac{1}{1+r_L} = \frac{1}{1+r+p}$. Here the expertise of the investment banker is required to determine the appropriate probability of default as part of their credit risk assessment.

D.2 Comparable companies analysis

Rather than relying on their own valuation of a company, investment banks might make use of market data available from other companies or bonds. For the valuation of shares and bonds it is common to identify other companies or bond issuers that have similar characteristics to the one to be considered. Investment banks look for companies that are active in the same sector and jurisdiction, have a similar size, comparable capital structure, as well as a similar market position within their industry. Having identified such companies, the investment bank then compares so called trading multiples. These often include the price-earnings-ratio or the ratio of the price and the earnings before interest, taxes, depreciation, and amortization (EBITDA). Applying the same ratio to the company to be valued, will then give its value. Underlying this methodology is the assumption that the future dividends of companies that are currently similar, will develop in a similar way and that the risks involved are also comparable, hence the same discount factor can be applied. For bonds, the comparison is easier in that the yield of bonds should be identical for otherwise comparable issuers; the only addition for bonds is that bonds of similar time to maturity have to be chosen.

The comparable company analysis is often used in addition to the discounted cash-flow methodology and might provide some insights into the assumptions the market makes about other companies' growth potential and risks that might inform the valuation of the company concerned. Investment banks might want to adjust the risks or growth prospects from those of comparable companies if this is justified by the circumstances of the company or issuer.

D.3 Precedents transactions analysis

In particular if the valuation of a company is made for a specific purpose, such as a merger or an IPO, the valuation might deviate from that of comparable companies. Such a deviation maybe the result of synergies having to be taken into account in mergers, or an acquirer might be willing to pay more for a target company as he believes he can change the prospects of the company and thus add value. This would, of course, not be included in the market price of comparable companies. In an IPO it is common to underprice the issue, which would also not be included in the price of companies already listed on an exchange. In this case, investment banks would look at similar trading multiples as in the comparable companies analysis, but select companies that recently have undergone similar transactions to the company they seek to analyse. This gives investment bankers information on how high merger premia might be and how this would translate into appropriate transaction prices in the case they are advising on. Similarly, it would give them an initial indication of a reasonable price for an IPO or a reasonable estimate for the yield of a bond, even before any book-building would commence to gain more information on the market perception of the security.

Again, similar to the comparable companies analysis, the precedents transactions analysis would serve as a benchmark to analyse the company or issuer and the results would then be compared to that of a discounted cash flow analysis for a more complete evaluation of the security.

D.4 Derivatives pricing

The valuation of derivatives is usually conducted very differently from that of shares or bonds. Most derivatives have a time to maturity, similar to bonds, and the payments made from owning derivatives often depend on the value of other securities, such as in options, futures, and swaps. However, it can be that derivatives are not based on other securities, but events happening to other entities, such as the default of companies triggering payments from credit default swaps. There are two basic methodologies to value derivatives. The first methodology seeks to balance the present value of future payments that these derivatives make with the payments the owner has to make. In the case of an option, this would mean that the payments received from exercising the option, if it is profitable to exercise the option, discounted to its present value, should equal the option premium the buyer has to pay. In the case of credit default swaps, the payments triggered by the default of the underlying entity have to the balanced against the regular payments the owner of this derivative has to make. The value would then be the amount that is regularly paid.

An alternative approach is to form a replicating portfolio. In this methodology, the value of the derivative is determined by a portfolio of securities that are available in the market and for which market prices are observable. A portfolio is constructed such that whenever the derivative makes a payment to its owner, often, but not always, at maturity, the portfolio makes the same payment or has an identical value. The market prices of the components of this portfolio are then added to give the value of the security. Often the portfolio includes short positions of securities.

Both methodologies should provide the same value of the derivative, but often one methodology is chosen in preference over the other for convenience of calculating the value. In other cases, one methodology, in most cases the replicating portfolio, is not available as securities required in such a portfolio are not traded. In other instances the reason for a specific methodology might be historical in that a certain methodology was first used to determine the value in the literature and has been maintained since.

Of central importance in derivatives pricing is the modelling of the payments the derivative makes and/or the payments the buyer makes. These payments are often assumed to have distributions that arise from stochastic processes; commonly the timing of such payments are also the result of the same or a different stochastic process. Solving for the expected payments or obtaining a replicating portfolio uses advanced mathematical and statistical methods, making the valuation of derivatives a task which is commonly left to specialists rather than investment bankers in direct contact with their clients.

D.4.1 Black-Scholes model for European Options

A Call option is an option which gives the purchaser the right to obtain the underlying asset at the strike price from the seller at maturity, European option, or at any time until maturity, American option. A Put option gives similarly the purchaser the right to sell the underlying asset to the seller of the option.

If we assume that the return on an asset is driven by a Geometric Brownian motion such that its value V evolves according to

$$\frac{dV}{V} = \mu dt + \sigma dz, \tag{D.3}$$

where μ denotes the expected return of the asset, σ its standard deviation and dz is a Wiener process representing a random error term drawn from a standard normal distribution. In this case we obtain the value of a standard European Call option on this asset with a strike price of K and a time to maturity of T time periods as

$$C = V\Phi(d) - Ke^{-rT}\Phi\left(d_1 - \sigma\sqrt{T}\right), \tag{D.4}$$

where $d = \frac{\ln\frac{V}{K}+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$, r represents the risk-free rate, and $\phi(\cdot)$ the cumulative standard normal distribution. The value of a European Put option, P, can then be obtained by applying the Put-Call parity $P = C - V + Ke^{-rT}$.

D.4.2 Barrier options

A barrier option is similar to a call option, but the option can only be exercised of a certain price level of the underlying level, the barrier H, has not been reached or it needs to be reached such that the option can be exercised. If the price of the underlying asset goes below the barrier and this causes the option to be exercisable, we have a down-and-in option, if the price has to exceed the barrier it is a up-and-in option. If the option cannot be exercised if the price falls below the barrier it is called a down-and-out option and if the barrier has to be exceeded it s a up-an-out option. Call and Put options give the right to buy or sell the underlying asset, respectively, at the stated strike price, K.

Using the assumption of a Geometric Brownian motion for the returns of the underlying asset, the value of a down-and-out Call option in the case that $H \geq K$ is given by

$$\begin{aligned} C_{DO} = {} & V\Phi\left(d_1\right) - K\Phi\left(d_1 - \sigma\sqrt{T}\right) \\ & -V\left(\frac{H}{V}\right)^{2d_3}\Phi\left(d_2\right) \\ & +K\left(\frac{H}{V}\right)^{2d_3-2}\Phi\left(d_s - \sigma\sqrt{T}\right) \end{aligned} \tag{D.5}$$

and if $H < K$ we have a down-and-in option given by

$$C_{DI} = V\left(\frac{H}{V}\right)^{2d_3}\Phi\left(d_4\right) - Ke^{-rT}\left(\frac{H}{V}\right)^{2d_3-2}\Phi\left(d_4 - \sigma\sqrt{T}\right). \tag{D.6}$$

Using the relationship that $C_{DO} = C - C_{DI}$, where C denotes the value of a standard European Call option, we can obtain the value of the down-and-out Call option. We require these supplementary variables:

$$\begin{aligned} d_1 &= \frac{\ln\frac{V}{H}}{\sigma\sqrt{T}} + d_3\sigma\sqrt{T} \\ d_2 &= \frac{\ln\frac{H}{V}}{\sigma\sqrt{T}} + d_3\sigma\sqrt{T} \\ d_3 &= \frac{r + \frac{1}{2}\sigma^2}{\sigma^2}, \\ d_4 &= \frac{\ln\frac{H^2}{VK}}{\sigma\sqrt{T}} + d_3\sigma\sqrt{T}. \end{aligned} \tag{D.7}$$

In a similar way the value of up-and-in and up-and-out Call options can be determined, as well as the equivalent put options.

References

Acharya, V. V., Lenzu, S., & Wang, O. (2021, December). *Zombie lending and policy traps.* (NBER Working Paper No. 29606)

Acharya, V. V., & Rajan, R. (2024). Liquidity, liquidity everywhere, not a drop to use: Why flooding banks with central bank reserves may not expand liquidity. *Journal Of Finance*, *79*(5), 2943-2992. doi: 10.1111/jofi.13370

Acharya, V. V., Santos, J. A. C., & Yorulmazer, T. (2010). Systemic risk and deposit insurance premiums. *Federal Reserve Bank of New York Economic Policy Review, August*, 89-99.

Acharya, V. V., & Yorulmazer, T. (2007). Too many to fail – an analysis of time-inconsistency in bank closure policies. *Journal of Financial Intermediation*, *16*, 1-31. doi: 10.1016/j.jfi.2006.06.001

Acharya, V. V., & Yorulmazer, T. (2008). Information contagion and bank herding. *Journal of Money, Credit and Banking*, *40*(1), 215-231.

Ahnert, T., Chapman, J., & Wilkins, C. (2021). Should bank capital regulation be risk sensitive? *Journal of Financial Intermediation*, *46*, 100870. doi: 10.1016/j.jfi.2020.100870

Ahnert, T., & Georg, C.-P. (2018). Information contagion and systemic risk. *Journal of Financial Stability*, *35*, 159-171. doi: 10.1016/j.jfs.2017.05.009

Ahnert, T., & Kuncl, M. (2024). Government loan guarantees, market liquidity, and lending standards. *Management Science*, *70*(7), 4502-4532.

Akerlof, G. A. (1970). The market for ”lemons”: Quality uncertainty and the market mechanis. *Quarterly Journal of Economics*, *84*(3), 488-500.

Allen, F. (1983). Credit rationing and payment incentives. *Review of Economic Studies*, *50*(4), 639-646.

Allen, F., & Gale, D. (1998). Optimal financial crises. *Journal of Finance*, *53*(4), 1245-1284.

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy*, *108*(1), 1-33.

Allen, F., & Gale, D. (2001). *Comparing financial systems*. MIT Press.

Almazan, A. (2002). A model of competition in banking: Bank capital versus expertise. *Journal of Financial Intermediation*, *11*, 87-121. doi: https://doi.org/10.1006/jfin.2001.0327

Amendola, A., Barra, C., Boccia, M., & Papaccio, A. (2021). Market structure and financial stability: the interaction between profit-oriented and mutual cooperative banks in italy. *Journal of Financial Services Research*, *60*, 235-259. doi: 10.1007/s10693-021-00360-1

Aptus, E., Britz, V., & Gersbach, H. (2020). Crisis contracts. *Economic Theory*, *70*, 121-164. doi: https://doi.org/10.1007/s00199-019-01204-9

Arnold, L. G., & Riley, J. G. (2009). On the possibility of credit rationing in the stiglitz-weiss model. *American Economic Review*, *99*(5), 2012-2021.

Arping, S. (2017). Deposit competition and loan markets. *Journal of Banking and Finance*, *80*, 108-118. doi: https://doi.org/10.1016/j.jbankfin.2017.04.006

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, *53*(5), 941-973.

Azmat, S., Azad, A. S. M. S., Bhatti, M. I., & Ghaffar, H. (2020). Islamic banking, costly religiosity, and competition. *Journal of Financial Research*, *43*(2), 263-303. doi: https://doi.org/10.1111/jfir.12207

Banerji, S., Bhattacharya, J., & Long, N. V. (2004). Can financial intermediation induce endogenous fluctuations. *Journal of Economic Dynamics & Control*, *28*, 2215-2238. doi: https://doi.org/10.1016/j.jedc.2003.10.002

Barron, J. M., Chong, B.-U., & Staten, M. E. (2008). Emergence of captive finance companies and risk segmentation in loan markets: Theory and evidence. *Journal of Money, Credit and Banking*, *40*(1), 173-192.

Bech, M. L., & Garratt, R. J. (2003). The intraday liquidity management game. *Journal of Economic Theory*, *109*, 198-219.

Bennardo, A., Pagano, M., & Piccolo, S. (2015). Multiple bank lending, creditor rights, and information sharing. *Review of Finance*, *19*, 519-570.

Berlin, M., & Mester, L. J. (1992). Debt covenants and renegotiation. *Journal of Financial Intermediation*, *2*, 95-135.

Bester, H. (1985). Screening vs rationing in credit markets with imperfect information. *American Economic Review*, *75*(4), 850-855.

Bester, H. (1995). A bargaining model of financial intermediation. *European Economic Review*, *39*, 211-228.

Bester, H., & Hellwig, M. (1987). Moral hazard and equilibrium credit rationing: An overview of the issues. In G. Namberg & K. Spremann (Eds.), *Agency theory, information, and incentives* (p. 135-166). Hamburg: Springer Verlag.

Bhattacharya, S., & Gale, D. (1987). Preference shocks, liquidity, and central bank policy. In W. A. Barnett & K. J. Singleton (Eds.), *International symposium in economic: The new approaches to monetary economics* (p. 69-88). Cambridge: Cambridge University Press.

Biglaiser, G., & Li, F. (2018). Middlemen: the good, the bad, and the ugly. *RAND Journal of Economics*, *49*(1), 3-22.

Biswas, S., & Koufopoulos, K. (2022). Bank capital structure and regulation: Overcoming and embracing adverse selection. *Journal of Financial Economics*, *143*, 973-992. doi: 10.1016/j.jfineco.2021.12.001

Blum, J. M. (2008). Why 'basel ii' may need a leverage ratio restriction. *Journal of Banking & Finance*, *32*, 1699-1707. doi: 10.1016/j.jbankfin.2007.12.003

Boissay, F., & Cooper, R. (2020). The collateral composition channel. *American Economic Journal*, *12*(1), 41-75.

Bolton, P., & Freixas, X. (2000). Equity, bonds, and bank debt: Capital structure and financial market equilibrium under asymmetric information. *Journal of Political Economy*, *108*(2), 324-351.

Bolton, P., Freixas, X., Gambacorta, L., & Mistrulli, P. E. (2016). Relationship and transaction lending in a crisis. *Review of Financial Studies*, *29*(10), 2633-2676.

Bolton, P., & Scharfstein, D. S. (1990). A theory of predation based on agency problems in financial contracting. *American Economic Review*, *80*(1), 93-106.

Bond, P., & Glode, V. (2014). The labor market for bankers and regulators. *Review of Financial Studies*, *27*(9), 2539-2579. doi: 10.1093/rfs/hht132

Boot, A. W. A., & Thakor, A. V. (2000). Can relationship banking survive competition? *Journal of Finance*, *55*, 679-713.

Boot, A. W. A., Thakor, A. V., & Udell, G. F. (1991). Credible commitments, contract enforcement problems and banks: Intermediation as credible insurance. *Journal of Banking & Finance*, *15*, 605-632.

Bouckaert, J., & Degryse, H. (1995). Phonebanking. *European Economic Review*, *39*, 229-244.

Boyer, P. C., & Kempf, H. (2020). Regulatory arbitrage and the efficiency of banking regulation. *Journal of Financial Intermediation*, *41*, 100765. doi: 10.1016/j.jfi.2017.09.002

Bräuning, F., & Ivashina, V. (2020). Monetary policy and global banking. *Journal of Finance*, *75*(6), 3055-3095. doi: 10.1111/jofi.12959

Brunnermeier, M. K., & Oehmke, M. (2013). The maturity rat race. *Journal of Finance*, *68*(2), 483-521.

Buckle, S., & Campbell, E. (2003). *Settlement bank behaviour and throughput rules in an rtgs payment system with collateralised intraday credit.* (Bank of England Working Paper No. 209)

Bunderson, S., & Thakor, A. V. (2021). Higher purpose, banking and stability. *Journal of Banking and Finance*, *forthcoming*. doi: 10.1016/j.jbankfin.2021.106138

Calomiris, C. W., Heider, F., & Hoerova, M. (2015). *A theory of bank liquidity requirements.* (Working Paper, Columbia University)

Cao, J. (2022). *The economics of banking*. Abingdon: Routledge.

Carletti, E., Dell'Ariccia, G., & Marquez, R. (2021). Supervisory incentives in a banking union. *Management Science*, *67*(1), 455-470. doi: 10.1287/mnsc.2019.3448

Carletti, E., Leonello, A., & Marquez, R. (2023). Loan guarantees, bank underwriting policies and financial stability. *Journal of Financial Economcis*, *149*, 260-295.

Carlin, B. I. (2009). Strategic price complexity in retail financial markets. *Journal of Financial Economics*, *91*, 278-287. doi: 10.1016/j.jfineco.2008.05.002

Castiglionesi, F., & Wagner, W. (2013). On the efficiency of bilateral interbank insurance. *Journal of Financial Intermediation*, *22*, 177-200.

Chakravorti, S., & To, T. (2007). A theory of credit cards. *International Journal of Industrial Organization*, 583-595.

Challe, E., Mojon, B., & Ragot, X. (2013). Equilibrium risk shifting and interest rate in an opaque financial system. *European Economic Review*, *63*, 117-133. doi: 10.1016/j.euroecorev.2013.07.003

Chan, Y.-S., & Kanatas, G. (1985). Asymmetric valuations and the role of collateral in loan agreements. *Journal of Money, Credit and Banking*, *17*(1), 84-95.

Chapman, J., Chiu, J., & Molico, M. (2013). A model of tiered settlement networks. *Journal of Money, Credit and Banking*, *45*(2-3), 327-347.

Chiappori, P.-A., Perez-Castrillo, D., & Verdier, T. (1995). Spatial competition in the banking system: Localization, cross subsidies, and the regulation of deposit rates. *European Economic Review*, *39*, 889-918. doi: https://doi.org/10.1016/0014-2921(94)00016-S

Clayton, C., & Schaab, A. (2025). Bnail-ins, optimal regulation, and crisis resolution. *Review of Financial Studies*, *forthcoming*, 10.1093/rfs/hhaf002. doi: 10.1016/j.jbankfin.2021.106300

Colliard, J.-E. (2020). Optimal supervisory architecture and financial integration in a banking union. *Review of Finance*, *24*(1), 129-161. doi: 10.1093/rof/rfz004

Cook, D. O., & Spellman, L. J. (1994). Repudiation risk and restitution costs: Toward understanding premiums on insured deposits. *Journal of Money, Credit and Banking*, *26*(3), 439-459.

Dam, K., & Chowdhury, P. R. (2021). Monitoring and incentives under multiple-bank lending: The role of collusive threats. *Journal of Economic Theory*, *197*, 105320. doi: 10.1016/j.jet.2021.105320

Dávila, E., & Walther, A. (2020). Does size matter? bailouts with large and small banks. *Journal of Financial Economics*, *136*, 1-22. doi: 10.1016/j.jfineco.2019.09.005

de Aghion, B. A. (1999). On the design of a credit agreement with peer monitoring. *Journal of Development Economics*, *60*, 79-104.

Dell'Ariccia, G., Laeven, L., & Marquez, R. (2014). Real interest rates, leverage, and bank risk-taking. *Journal of Economic Theory*, *149*, 65-99. doi: 10.1016/j.jet.2013.06.002

Dell'Ariccia, G., & Ratnovski, L. (2019). Bailouts and systemic insurance. *Journal of Banking and Finance*, *105*, 166-177. doi: 10.1016/j.jbankfin.2019.05.019

Dermine, J. (1986). Deposit rates, credit rates and bank capital: The monti-klein model revisited. *Journal of Banking & Finance*, *10*, 99-114. doi: https://doi.org/10.1016/0378-4266(86)90022-1

de Roure, C., Pelizzon, L., & Thakor, A. (2021). P2p lenders versus banks: Cream skimming or bottom fishing? *The Review of Corporate Finance Studies*, *forthcoming*. doi: 10.1093/rcfs/cfab026

Detragiache, E., Garella, P., & Guiso, L. (2000). Multiple versus single banking relationships: Theory and evidence. *Journal of Finance*, *55*(3), 1133-1161.

Dewatripont, M., & Tirole, J. (1994). *The prudential regulation of banks*. Cambridge, MA: MIT Press.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies*, *51*, 393-414.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, *91*(3), 401-419.

Diamond, D. W., & Rajan, R. G. (2000). A theory of bank capital. *Journal of Finance*, *55*, 2431-2465.

Diamond, D. W., & Rajan, R. G. (2005). Liquidity shortage and banking crises. *Journal of Finance*, *60*(2), 615-647. doi: 10.1111/j.1540-6261.2005.00741.x

Diamond, D. W., & Rajan, R. G. (2011). Fear of fire sales, illiquidity seeking, and credit freezes. *Quarterly Journal of Economics*, *126*, 557-591. doi: 10.1093/qje/qjr012

Ding, K., Hill, E., & Perez-Reyna, D. (2021). Optimal capital requirements with noisy signals on banking risk. *Economic Theory*, *71*, 1649-1687. doi: 10.1007/s00199-020-01310-z

Donaldson, J. R., Piacentino, G., & Thakor, A. (2018). Warehouse banking. *Journal of Financial Economics*, *129*(2), 250-267.

Donze, J., & Dubec, I. (2006). The role of interchange fees in atm networks. *International Journal of Industrial Organization*, *24*, 29-43.

Dow, J., & Han, J. (2018). The paradox of financial fire sales: The role of arbitrage capital in determining liquidity. *The Journal of Finance*, *73*(1), 229-274. doi: 10.1111/jofi.12584

Dreyfus, J.-F., Saunders, A., & Allen, L. (1994). Deposit insurance and regulatory forbearance: Are caps on insured deposits optimal? *Journal of Money, Credit, and Banking*, *26*, 412-438.

Eisert, T., & Eufinger, C. (2019). Interbank networks and backdoor bailouts: Benefiting from other banks' government guarantees. *Management Science*, *65*(8), 3673-3693. doi: 10.1287/mnsc.2017.2968

Faria-e-Castro, M., Paul, P., & Sánchez, J. M. (2024). Evergreening. *Journal of Financial Economics*, *153*, 103778.

Feess, E., & Hege, U. (2012). The basel accord and the value of bank differentiation. *Review of Finance*, *16*, 1043-1092. doi: 10.1093/rof/rfr002

Flynn, S., Ghent, A., & Tchistyi, A. (2024). The imitation game: How encouraging renegotiation makes good borrowers bad. *Review of Financial Studies*, *37*, 3648-3709. doi: 10.1093/rfs/hhae060

Foote, E. (2014). Information asymmetries and spillover risk in settlement systems. *Journal of Banking & Finance*, *42*, 179-190.

Freixas, X., & Gabillon, E. (1999). Optimal regulation of a fully insured deposit banking system. *Journal of Regulatory Economics*, *16*(2), 111-134. doi: 10.1023/A:1008139900120

Freixas, X., & Parigi, B. M. (1998). Contagion and efficiency in gross and net interbank payment systems. *Journal of Financial Intermediation*, *7*, 3-31.

Freixas, X., Parigi, B. M., & Rochet, J.-C. (2000). Systemic risk, interbank relations, and liquidity provision by the central bank. *Journal of Money, Credit, and Banking*, *32*(3), 611-638.

Freixas, X., & Rochet, J.-C. (2008a). *Microeconomics of banking* (2nd ed.). Cambridge, MA: MIT Press.

Freixas, X., & Rochet, J.-C. (2008b). *Microeconomics of banking* (2nd ed.). MIT Press.

Freixas, X., & Rochet, J.-C. (2023). *Microeconomics of banking* (3rd ed.). MIT Press.

Furlong, F. T., & Keeley, M. C. (1989). Capital regulation and rank risk-taking: A note. *Journal of Banking and Finance*, *13*, 883-891.

Gale, D., & Hellwig, M. (1985). Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies*, *52*, 647-663.

Gangopadhyay, S., & Mukhopadhyay, B. (2002). Multiple bank lending and seniority in claims. *Journal of Economics and Business*, *54*, 7-30.

Geanakoplos, J. (2010). The leverage cycle. *NBER Macroeconomics Annual*, *24*, 1-66. doi: 10.1086/648285

Gennaioli, N., Shleifer, A., & Vishny, R. W. (2013). A model of shadow banking. *Journal of Finance*, *68*(4), 1331-1363. doi: https://doi.org/10.1111/jofi.12031

Gersbach, H., & Papageorgiou, S. (2019, January). *On banking regulation and lobbying*. (Center of Economic Research at ETH Zurich Working Paper 19/308)

Gersbach, H., & Rochet, J.-C. (2017). Capital regulation and credit fluctuations. *Journal of Monetary Economics*, *90*, 113-124. doi: 10.1016/j.jmoneco.2017.05.008

Ghatak, M. (1999). Group lending, local information and peer selection. *Journal of Development Economics*, *60*, 27-50.

Ghatak, M., & Guinnane, T. W. (1999). The economics of lending with joint liability: theory and practice. *Journal of Development Economics*, *60*, 195-228.

Glocker, C. (2021). Reserve requirements and financial stability. *Journal of International Financial Markets, Institutions & Money*, *71*, 101286. doi: 10.1016/j.intfin.2021.101286

Goldstein, I., & Pauzner, A. (2005). Demand deposit contracts and the probability of bank runs. *Journal of Finance*, *60*, 1293-1328.

Gorton, G. (1985). Banks' suspension of convertibility. *Journal of Monetary Economics*, *15*, 177-193.

Gorton, G., & Huang, L. (2006). Bank panics and the endogeneity of central banking. *Journal of Monetary Economics*, *53*(7), 1613-1629. doi: 10.1016/j.jmoneco.2005.05.015

Greenbaum, S. I., Kanatas, G., & Venezia, I. (1989). Equilibrium loan pricing under the bank-client relationship. *Journal of Banking & Finance*, *13*, 221-235.

Greenbaum, S. I., & Thakor, A. V. (1987). Bank funding modes: Securitization vs deposits. *Journal of Banking & Finance*, *11*, 379-401.

Grochulski, B., & Kareken, J. (2004). Allen and gale on risk-taking and competition in banking. *Finance Research Letters*, *1*, 236-240. doi: https://doi.org/10.1016/j.frl.2004.08.001

Górnicka, L. A. (2016). Banks and shadow banks: Competitors or complements? *Journal of Financial Intermediation*, *27*, 118-131. doi: https://doi.org/10.1016/j.jfi.2016.05.002

Hauck, A., Neyer, U., & Vieten, T. (2015). Reestablishing stability and avoiding a credit crunch: Comparing different bad bank schemes. *Quarterly Review of Economics and Finance*, *57*, 116-128. doi: 10.1016/j.qref.2014.10.002

Haufler, A. (2021). Regulatory and bailout decisions in a banking union. *Journal of Banking and Finance*, *133*, 106300. doi: 10.1016/j.jbankfin.2021.106300

Hauswald, R., & Marquez, R. (2006). Competition and strategic information acquisition in credit markets. *Review of Financial Studies*, *19*, 967-1000.

He, Z., Huang, J., & Zhou, J. (2023). Open banking: Credit market competition when borrowers own the data. *Journal of Financial Economics*, *147*, 449-474.

He, Z., & Manela, A. (2016). Information acquisition in rumor-based bank runs. *Journal of Finance*, *71*(3), 1113-1158.

Heddergott, D., & Laitenberger, J. (2017). A simple model of banking competition with bank size heterogeneity and lending spillovers. *Economic Notes by Banca Monte dei Paschi di Siena SpA*, *46*(2), 381-404.

Heider, F., Hoerova, M., & Holthausen, C. (2015). Liquidity hoarding and interbank market rates: The role of counterparty risk. *Journal of Financial Economics*, *118*(2), 336-254.

Hellmann, T. F., Murdock, K. C., & Stiglitz, J. E. (2000). Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American Economic Review*, *90*, 147-165.

Ho, T. S. Y., & Saunders, A. (1981). The determinants of bank interest margins: Theory and empirical evidence. *Journal of Financial and Quantitative Analysis*, *16*, 581-600. doi: 10.2307/2330377

Hoffmann, F., & Vladimirov, V. (2025). Worker runs. *Journal of Finance*, *80*(2), 937-980. doi: 10.1111/jofi.13424

Hotelling, H. (1929). Stability in Competition. *Economic Journal*, *39*(153), 41-57.

Hu, Y., & Varas, F. (2021). A theory of zombie lending. *Journal of Finance*, *76*(4).

Hwang, J., Jiang, N., & Wang, P. (2007). Collusion and overlending. *Economic Inquiry*, *45*(4), 691-707. doi: https://doi.org/10.1111/j.1465-7295.2007.00046.x

Inderst, R., & Ottaviani, M. (2009). Misselling through agents. *American Economic Review*, *99*(3), 883-908. doi: 10.1257/aer.99.3.883

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, *52*(1), 45-67.

Jacklin, C. J. (1987). Demand deposits, trading restrictions, and risk sharing. In E. Prescott & N. Wallace (Eds.), *Contractual arrangements for intertemporal trade.* Minnesotal University Press.

Jappelli, T., Pagano, M., & Bianco, M. (2005). Courts and banks: Effects of judicial enforcement on credit markets. *Journal of Money, Credit and Banking*, *37*(2), 223-244.

Karapetyan, A., & Stacescu, B. (2014). Does information sharing reduce the role of collateral as a screening device? *Journal of Banking & Finance*, *43*, 48-57.

Kashyap, A. K., Rajan, R. G., & Stein, J. C. (2002). Banks as liquidity providers: An explanation for the coexistence of lending and deposit taking. *Journal of Finance*, *57*(1), 33-73.

Keiding, H. (2016a). *The economics of banking*. London: Palgrave Macmillan.

Keiding, H. (2016b). *The economics of banking*. Palgrave Macmillan.

Klein, M. A. (1971). A theory of the banking firm. *Journal of Money, Credit and Banking*, *3*, 205-218.

Krause, A. (2022a). The optimal exclusion length of borrowers after default. *Economics Letters*, *220*, 110881.

Krause, A. (2022b). Strategic default and optimal audit resources with costly state verification. *Research in Economcis*, *76*, 413-421.

Krause, A. (2025a). Bank risk-taking and optimal bailout resources. *Journal of Banking Regulation*, *forthcoming*.

Krause, A. (2025b). *A theoretical model of bank valuation.* (Working Paper, University of Bath)

Leland, H. E., & Pyle, D. H. (1977). Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance*, *32*, 371-387.

Li, M., Milne, F., & Qiu, J. (2016). Uncertainty in an interconnected financial system, contagion, and market freezes. *Journal of Money, Credit and Banking*, *48*(6), 1135-1168.

Ma, C., & Nguyen, X.-H. (2021). Too big to fail and optimal regulation. *International Review of Economics and Finance*, *75*, 747-758. doi: 10.1016/j.iref.2021.04.032

Mailath, G. J., & Mester, L. J. (1994). A positive analysis of bank closure. *Journal of Financial Intermediation*, *3*, 272-299. doi: 10.1006/jfin.1994.1007

Martin, A., Skeie, D., & von Thadden, E.-L. (2014). Repo runs. *Review of Financial Studies*, *27*(4), 957-989.

Martinez-Miera, D., & Repullo, R. (2010). Does competition reduce the risk of bank failure? *Review of Financial Studies*, *23*(10), 3638-3664. doi: https://doi.org/10.1093/rfs/hhq057

Matsuoka, T. (2013). Sunspot bank runs in competitive versus monopolistic banking systems. *Economics Letters*, *118*, 247-249.

Matutes, C., & Padilla, A. J. (1994). Shared atm networks and banking competition. *European Economic Review*, *38*(5), 1113-1138.

Matutes, C., & Vives, X. (1996). Competition for deposits, fragility, and insurance. *Journal of Financial Intermediation*, *5*, 184-216.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, *29*(2), 449-470.

Merton, R. C. (1977). An analytic derivation of the cost of deposit insurance and loan guarantees. *Journal of Banking & Finance*, *1*, 3-11.

Merton, R. C., & Thakor, R. T. (2019). Customers and investors: A framework for understanding the evolution of financial institutions. *Journal of Financial Intermediation*, *39*, 4-18.

Meza, D. D., & Webb, D. C. (1987). Too much investment: A problem of asymmetric information. *Quarterly Journal of Economics*, *102*(2), 281-292.

Milbourne, R. (1983). Optimal money holding under uncertainty. *International Economic Review*, *24*(3), 685-698.

Miller, M. H., & Orr, D. (1966). A model of the demand for money by firms. *Quarterly Journal of Economics*, *80*(3), 413-435.

Mitchell, J. (2001). Bad debts and the cleaning of banks' balance sheets: An application to transition economies. *Journal of Financial Intermediation*, *10*, 1-27. doi: 10.1006/jfin.2000.0304

Monnet, E., & Vari, M. (2023). A dilemma between liquidity regulation and monetary policy: Some history and theory. *Journal of Money, Credit and Banking*, *55*(4), 915-944. doi: 10.1111/jmcb.12930

Morrison, A. D., & White, L. (2005). Crises and capital requirements in banking. *American Economic Review*, *95*, 1548-1572.

Morrison, A. D., & White, L. (2011). Deposit insurance and subsidized recapitalizations. *Journal of Banking & Finance*, *35*, 3400-3416.

Nash, J. (1953). Two-Person Cooperative Games. *Econometrica*, *21*(1), 128–140.

Niinimaki, J.-P. (2007). Evergreening in banking. *Journal of Financial Stability*, *3*, 368-393.

Padilla, A. J., & Pagano, M. (2000). Sharing default information as a borrower discipline device. *European Economic Review*, *44*(10), 1951-1980.

Pandolfi, L. (2021). Bail-in and bailout: Friends or foes? *Management Science*, *forthcoming*. doi: 10.1287/mnsc.2020.3883

Park, H., & Kahn, C. M. (2019). Collateral, rehypothecation, and efficiency. *Journal of Financial Intermediation*, *39*, 34-46.

Petersen, M. A., & Rajan, R. G. (1995). The effect of credit market competition on lending relationships. *Quarterly Journal of Economics*, *110*, 407-443.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, *32*(1/2), 122-136.

Ratnovski, L. (2009). Bank liquidity regulation and the lender of last resort. *Journal of Financial Intermediation*, *18*(4), 541-558. doi: 10.1016/j.jfi.2007.08.001

Repullo, R. (2000). Who should act as a lender of last resort? an incomplete contracts model. *Journal of Money, Credit and Banking*, *32*(3), 580-605.

Rochet, J.-C. (1992). Capital requirements and the behavior of commercial banks. *European Economic Review*, *36*, 1137-1178. doi: 10.1016/0014-2921(92)90051-W

Rochet, J.-C., & Vives, X. (2004). Coordination failures and the lender of last resort: Was bagehot right after all? *Journal of the European Economic Association*, *2*, 1116-1147.

Rosenbaum, J., & Pearl, J. (2009). *Investment banking*. Hoboken, NJ: John Wiley & Sons.

Salop, S. C. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics*, *10*(1), 141-156.

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance*, *49*, 1069-1087.

Shleifer, A., & Vishny, R. W. (1994). Politicians and firms. *Quarterly Journal of Economics*, *109*(4), 995-1025.

Shy, O. (2024). Why do banks require minimum balance to avoid a fee? *Annals of Finance*, *20*, 495-420. doi: 10.1007/s10436-024-00449-2

Shy, O., Stenbacka, R., & Yankov, V. (2016). Limited deposit insurance coverage and bank competition. *Journal of Banking & Finance*, *71*, 95-108.

Smith, B. D., & Stutzer, M. J. (1990). Adverse selection and mutuality: The case of the farm credit system. *Journal Of Financial Intermediation*, *1*, 125-149.

Song, F., & Thakor, A. (2022). Ethics, capital and talent competiton in banking. *Journal of Financial Intermediation*, *52*, 100963.

Song, F., & Thakor, A. V. (2019). Bank culture. *Journal of Financial Intermediation*, *39*, 59-79. doi: https://doi.org/10.1016/j.jfi.2018.06.005

Spulber, D. F. (1999). *Market microstructure: Intermediaries and the theory of the firm*. Cambridge: Cambridge University Press.

Stiglitz, J. E. (1990). Peer monitoring and credit markets. *The World Bank Economic Review*, *4*(3), 351-366.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, *71*, 393-410.

Takáts, E. (2009). A theory of "crying wolf": The economics of money laundering enforcement. *Journal of Law, Economics, & Organization*, *27*(1), 32-78. doi: https://doi.org/10.1093/jleo/ewp018

Tassel, E. V. (1999). Group lending under asymmetric information. *Journal of Development Economics*, *60*, 3-25.

Thakor, A. V. (2021). Politics, credit allocation and bank capital requirements. *Journal of Financial Intermediation*, *45*, 100820. doi: 10.1016/j.jfi.2019.03.005

Thanassoulis, J. (2014). Bank pay caps, bank risk, and macroprudential regulation. *Journal of Banking & Finance*, *48*, 39–151. doi: https://doi.org/10.1016/j.jbankfin.2014.04.004

Thanassoulis, J., & Tanaka, M. (2018). Optimal pay regulation for too-big-to-fail banks. *Journal of Financial Intermediation*, *33*, 83-97. doi: https://doi.org/10.1016/j.jfi.2017.03.001

Thanassoulis, J., & Vadasz, T. (2021, June). *Free banking and credit market competition.* (Working Paper, University of Warwick)

Tomura, H. (2016). Investment horizon and repo in the over-the-counter market. *Journal of Money, Credit and Banking*, *48*(1), 145-164.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, *21*, 265-293.

Vasicek, O. A. (2002). Loan portfolio value. *Risk*(December), 160-162.

Vesala, T. (2007). Switching costs and relationship profits in bank lending. *Journal of Banking & Finance*, *31*, 477-493.

Wagner, W. (2010). Loan market competition and bank risk-taking. *Journal of Financial Services Research*, *37*, 71-81. doi: 10.1007/s10693-009-0073-8

Wallace, N. (1996). *Narrow banking meets the diamond-dybvig model.* (Working Paper, Federal Resrve Bank of Minneapolis Quarterly Review)

Webb, D. C. (1991). Long-term financial contracts can mitigate the adverse selection problem in project financing. *International Economic Review*, *32*(2), 305-320.

Wright, J. (2004). The determinants of optimal interchange fees in payment systems. *Journal of Industrial Economics*, *52*, 1-26.

Xiao, D., & Krause, A. (2022). Balancing liquidity and returns through interbank markets: Endogenous interest rates and network structures. *Journal of Financial Research*, *46*(1), 131‘-149.

Yafeh, Y., & Yosha, A. (2001). Industrial organization of financial systems and strategic use of relationship banking. *European Finance Review*, *5*, 63-78.

Glossary

Bailout A bail-out is a situation where a banks is prevented from failing by government or central bank intervention. Typically depositors do not face a loss from bail-outs.
See also *Bail-in*, *Back office*

Index

Andreas Krause

Theoretical Foundations of Commercial Banking

This book provides readers with a comprehensive and state-of-the-art overview of the theories of banking. It presents theories on lending decisions and any conditions associated with it, as well as deposit-taking. We use a consistent and coherent framework, that allows combining different theories to develop more comprehensive analysis of developments in this important industry. Going beyond the core activities of banks, this book also includes an analysis of some competition between banks, their regulation, and the employment practices and strategies found in banks.