Andreas Krause

Theoretical Foundations of Banking

Volume I: The core activities of commercial banks

Theoretical Foundations of Banking

Theoretical Foundations of Banking

Volume I

The core activities of commercial banks

Andreas Krause

Cover photo: Bernard Spragg, NZ from Christchurch, New Zealand, CC0, via Wikimedia Commons
https://commons.wikimedia.org/wiki/File:Bank_of_ China_ Hong_ Kong._(9532283389).jpg

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

Typeset in Computer Modern 10pt using LaTeX.

Outline

Table of contents

List of Figures

General 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.

Preface to volume I

Banks are mostly associated with the provision of loans and the acceptance of deposits. They provide loans to individuals, ranging from mortgages and home loans, to loans for the purchase of cars, home appliances, and electronic devices, but also to finance expenditures such holidays, weddings, or education. Other loans might be given to clear consolidate debt that often has accumulated from living expenses. To companies they provide loans for investments into real estate, machinery and other equipment, alongside the financing of their current operations. Banks may also provide loans to public bodies, such as government organisations, as well as other banks. Loans are given for time periods ranging from days to years, where the majority of loans has a duration of multiple years.

Similarly, banks accept deposits from individuals and companies, as well as public bodies. Some of these deposits are held temporarily to ensure the depositor can make payments that are due shortly, while other deposits are invested for a longer period of time and can be interpreted as an investment by the depositor. In addition to accepting such temporary deposits, banks also provide services that facilitate making payments between accounts of different depositors, allow the use of payment cards, as well as withdraw cash. While many such deposits are actually maintained for long periods of time, they can mostly be withdrawn without notice. This is in contrast to loans, which are mostly provided for longer periods of time.

In this first volume, we will look in detail not only at the role of banks in the economy, including why deposits can be withdrawn at short notice while loans are given long-term, but also at the provision of loans and acceptance of deposits. We will look at how loan contracts are structured, what the incentives for banks and borrowers are in a loan contract. For deposits we will explore the implications of the ability by depositors to withdraw them instantly and how the services banks provide are affecting bank behaviour.

It is these core activities of banks and their implications that will be out focus and which then will enable us in volume II to explore in more detail the competition between banks, how they operate, and how they are regulated. While some of these aspects are touched upon in the models presented here, they are neither the focus of our analysis nor are they explored thoroughly; we leave such an exploration of these aspects to the second volume.

The aim of this first volume is for readers to gain an understanding of the core activities banks are engaged in and start to appreciate their complexity before we then consider how banks interact with each other, with their employees, and with regulators.

Andreas Krause
Bath, October 2024

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

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

Figure 1: Bank balance sheet

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 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 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 resetrves 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

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

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). 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

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

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

$$\Pi_B = \pi\left(1+r_L\right)L-\left(1+r_D\right)D. \tag{3}$$

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

$$\Pi_D = \pi(1 + r_D)D - (1 + r)D, \quad (4)$$

$$\Pi_D = (1 + r_D)D - (1 + r)D, \quad (5)$$

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

$$\begin{aligned} \Pi_B &= p\left(\pi_H\left(1+r_L\right)L-\left(1+r_D\right)D\right) \\ &\quad +\left(1-p\right)\left(\pi_L\left(1+r_L\right)L-\left(1+r_D\right)D\right) \\ &= \left(p\pi_H+\left(1-p\right)\pi_L\right)\left(1+r_L\right)L-\left(1+r_D\right)D. \end{aligned} \tag{6}$$

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

$$1+r_L = \frac{1+r_D}{p\pi_H+\left(1-p\right)\pi_L}. \tag{7}$$

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

$$\Pi_C^i = \pi_i\left(1+R_i\right)I-\left(1+r_L\right)L \geq 0, \tag{8}$$

where R_i denotes the return of a successful investment for a company of type i. We assume now that $\pi_H\left(1+R_H\right) > 1+r_D > \pi_L\left(1+R_L\right)$, 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 (7) for the loan rate, we can solve equation (8) for the high-quality company as requiring that

$$p \geq p^* = \frac{(1 + r_D) - \pi_H \pi_L (1 + R_H)}{\pi_H - \pi_L}. \tag{9}$$

Hence if the fraction of high-quality companies is too small, there will be demand for loans by these companies. The profits of low-quality companies are given by

$$\begin{aligned} \Pi_C^L &= (\pi_L (1 + R_L) - (1 + r_L)) L \\ &= \frac{(p\pi_H + (1 - p) \pi_L) (1 + R_L) - (1 + r_D)}{p\pi_H + (1 - p) \pi_L} L, \end{aligned} \tag{10}$$

after inserting from equation (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 **?**. 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

$$\Pi_C^i = \pi_i \left((1 + R_i) I - (1 + r_L) L\right) = \pi_i (R_i - r_L) L, \tag{11}$$

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

$$1 + r_L \leq 1 + r_L^* = \frac{\pi_H (1 + R_H) - \pi_L (1 + R_L)}{\pi_H - \pi_L}, \tag{12}$$

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 (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 (9) and (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 1.1 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 1.3 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 2. 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 2.3. 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 3, but it is also one of the causes banks are fragile as we will discover in chapter 13. 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 4 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.

1 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 1.1 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 1.2 considering the case of banks having market power to set loan and deposit rates. The third model in chapter 1.3 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.

1.1 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:

$$C_1 + \hat{L} + D = W, \tag{1.1}$$

$$C_2 = \Pi_C + \Pi_B + \pi\,(1+r_C)\,\hat{L} + \pi\,(1+r_D)\,D, \tag{1.2}$$

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 (1.1) and (1.2). We clearly notice from equations (1.1) and (1.2) that bonds and deposits are perfect substitutes for the consumer and hence in equilibrium, we require that

$$r_D = r_C. \tag{1.3}$$

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

$$\Pi_F = \pi\Big((1+R)\,I - (1+r_C)\,\hat{L} - (1+r_L)\,L\Big) \tag{1.4}$$

and the investment available is

$$I = L + \hat{L}. \tag{1.5}$$

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

$$r_C = r_L. \tag{1.6}$$

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

$$L = D \tag{1.7}$$

and their profits, neglecting operating costs, are given by

$$\Pi_B = \pi \left((1 + r_L) L - (1 + r_D) D \right). \tag{1.8}$$

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 (1.3) and (1.6) which imply

$$r_C = r_L = r_D \tag{1.9}$$

and thus from equation (1.8) we obtain when using $L = D$ that

$$\Pi_B = 0. \tag{1.10}$$

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 2 when introducing transaction costs and chapter 3 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 (2008, ch. 1.7)

1.2 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

$$\begin{aligned} C_i^1 &= W_i - D_i + L_i, \\ C_i^2 &= W_i + (1 + r_D)D_i - (1 + r_L)L_i. \end{aligned} \tag{1.11}$$

The utility function of consumer i is given by

$$U_i\left(C_i^1, C_i^2\right) = u(C_i^1) + u(C_i^2), \tag{1.12}$$

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

$$\begin{aligned} \frac{\partial U_i\left(C_i^1, C_i^2\right)}{\partial D_i} &= -\frac{\partial u(C_i^1)}{\partial C_i^1} + (1 + r_D)\frac{\partial u(C_i^2)}{\partial C_i^2} = 0, \\ \frac{\partial U_i\left(C_i^1, C_i^2\right)}{\partial L_i} &= \frac{\partial u(C_i^1)}{\partial C_i^1} - (1 + r_L)\frac{\partial u(C_i^2)}{\partial C_i^2} = 0. \end{aligned} \tag{1.13}$$

From these conditions we easily get that

$$\begin{aligned} \frac{\partial u(C_i^1)}{\partial u(C_i^2)} &= 1 + r_D, \\ \frac{\partial u(C_i^1)}{\partial u(C_i^2)} &= 1 + r_L, \end{aligned} \tag{1.14}$$

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.

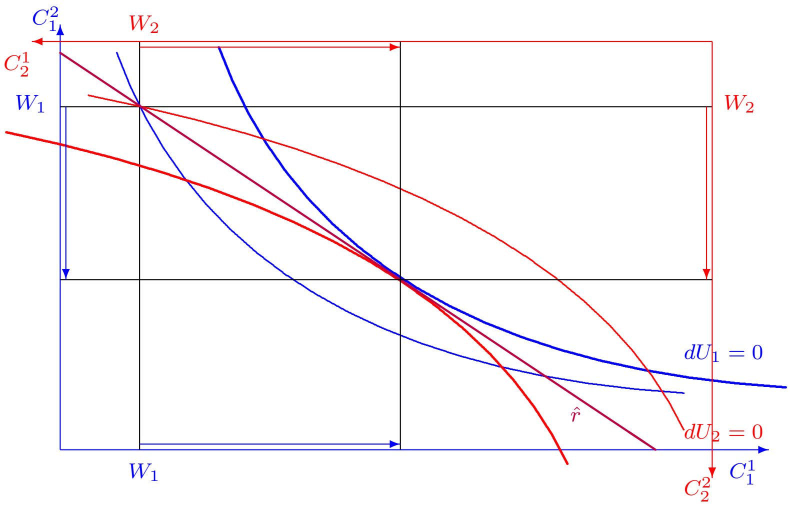
With the usual assumption of concave utility functions $u(\cdot)$, we see from equation (1.14) that consumers with a low income, W_i, are more likely to take out a loan. Their high marginal utility in time period 1 would reduce due to increased consumption and increase in time period 2, reducing the 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 (1.14) 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 (1.14) we require that both, depositors and borrowers, fulfill

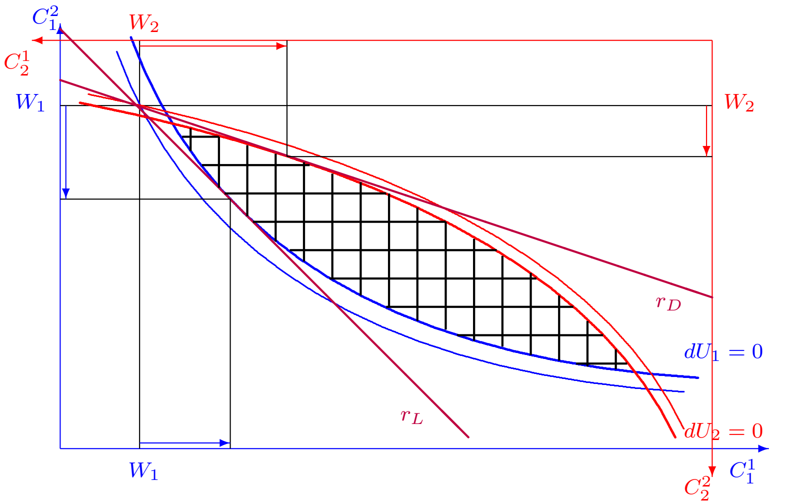
$$\frac{\partial u(C_i^1)}{\partial u(C_i^2)} = 1 + \hat{r}. \tag{1.15}$$

In equilibrium the marginal rates of substitution would be identical for borrowers and lenders. Figure 2a 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 (1.15). 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.

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



(a) Equilibrium with direct lending



(b) Equilibrium with bank lending

Figure 2: Pareto-efficiency with direct and bank lending

interest on loans, which are reduced by paying interest on deposits. Thus bank profits are given by

$$\Pi_B = (1+r)(D-L) + (1+r_L)L - (1+r_D)D. \tag{1.16}$$

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

$$\begin{aligned} \frac{\partial \Pi_B}{\partial r_D} &= (1+r)\frac{\partial D}{\partial r_D} - D - (1+r_D)\frac{\partial D}{\partial r_D} = 0, \\ \frac{\partial \Pi_B}{\partial r_L} &= -(1+r)\frac{\partial L}{\partial r_L} + L + (1+r_L)\frac{\partial L}{\partial r_L} = 0. \end{aligned} \tag{1.17}$$

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

$$\begin{aligned} 1+r_D &= 1+r-\frac{D}{\frac{\partial D}{\partial r_D}} \leq 1+r, \\ 1+r_L &= 1+r-\frac{L}{\frac{\partial L}{\partial r_L}} \geq 1+r, \end{aligned} \tag{1.18}$$

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 (1.14), 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 2b 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)

1.3 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

$$\Pi_D^E = \nu \left(p\pi_H (1+R_H) + (1-p)\pi_L (1+R_L)\right) I - (1+r_D) I. \tag{1.19}$$

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

$$\nu = \frac{1+r_D}{p\pi_H (1+R_H) + (1-p)\pi_L (1+R_L)} \tag{1.20}$$

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

$$\begin{aligned} \Pi_C^E &= (1-\nu)\left(p\pi_H (1+R_H) + (1-p)\pi_L (1+R_L)\right) I \\ &= \left(p\pi_H (1+R_H) + (1-p)\pi_L (1+R_L) - (1+r_D)\right) I, \end{aligned} \tag{1.21}$$

where the last equation has been obtained after inserting for ν from equation (1.20).

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

$$\Pi_D^C = p\pi_H (1 + r_C) L - (1 + r_D) L. \tag{1.22}$$

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

$$1 + r_C = \frac{1 + r_D}{p\pi_H}. \tag{1.23}$$

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

$$\begin{aligned} \Pi_C^C &= p\pi_H ((1 + R_H) L - (1 + r_C) L) + (1 - p) \pi_L (1 + R_L) L \\ &= (p\pi_H (1 + R_H) + (1 - p) \pi_L (1 + R_L) - (1 + r_D)) L \\ &= \Pi_C^E, \end{aligned} \tag{1.24}$$

where the second equality emerges when inserting equation (1.23) 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 (1.21), 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 $\pi_L (1 + R_L) L$, the resources available to the company. The profits of the bank using loan rate r_L are then given by

$$\Pi_B = (p\pi_H (1 + r_L) + (1 - p) \pi_L (1 + R_L)) L - (1 + r_D) D, \tag{1.25}$$

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

$$1 + r_L = \frac{(1 + r_D) - (1 - p) \pi_L (1 + R_L)}{p\pi_H}. \tag{1.26}$$

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 (1.26) for the loan rate and obtain the company profits as

$$\begin{aligned} \Pi_C^B &= p\pi_H ((1 + R_H) L - (1 + r_L) L) \\ &= (p\pi_H (1 + R_H) + (1 - p) \pi_L (1 + R_L) - (1 + r_D)) L \\ &= \Pi_C^E. \end{aligned} \tag{1.27}$$

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 2 or in chapter 3 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 2 to the transformation of short-term deposits into long-term loans in chapter 3. 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 VII.

2

Reducing transaction costs

HAVING CONSIDERED the case of frictionless markets in chapter 1, 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 2.1 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 2.2, 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 2.3 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.

2.1 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

$$\hat{\Pi}_C = \pi\left(\left(1+R\right)L - \left(1+r_C\right)L\right) - C, \tag{2.1}$$

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

$$\hat{\Pi}_D = \pi\left(1+r_C\right)L - L - C. \tag{2.2}$$

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

$$\mathcal{L} = \left(\hat{\Pi}_C + C\right)\left(\hat{\Pi}_D + C\right), \tag{2.3}$$

which gives us

$$\frac{\partial \mathcal{L}}{\partial\left(1+r_C\right)} = \pi L\left(\hat{\Pi}_C + C\right) - \pi L\left(\hat{\Pi}_D + C\right) = 0 \tag{2.4}$$

or $\hat{\Pi}_D = \hat{\Pi}_C$. Solving this relationship after inserting from equations (2.1) and (2.2), we get the expected repayment from the loan as

$$\pi\left(1+r_C\right)L = \frac{1}{2}\left(\pi\left(1+R\right)+1\right)L \tag{2.5}$$

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

$$\hat{\Pi}_C = \hat{\Pi}_D = \frac{1}{2}\left(\pi\left(1+R\right)-1\right)L - C. \tag{2.6}$$

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

$$C \leq C^* = \frac{1}{2}\left(\pi\left(1+R\right)-1\right)L. \tag{2.7}$$

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

$$\begin{aligned} \Pi_C &= \pi\left(\left(1+R\right)L-\left(1+r_L\right)L\right), \\ \Pi_D &= \pi\left(1+r_D\right)L-L, \\ \Pi_B &= \pi\left(\left(1+r_L\right)L-\left(1+r_D\right)L\right). \end{aligned} \tag{2.8}$$

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

$$\frac{\partial\mathcal{L}}{\partial\left(1+r_D\right)} = \pi L\Pi_B - \pi L\Pi_D = 0, \tag{2.9}$$

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

$$\pi\left(1+r_D\right)L = \frac{1}{2}\left(\pi\left(1+r_L\right)+1\right)L. \tag{2.10}$$

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

$$\frac{\partial\mathcal{L}}{\partial\left(1+r_L\right)} = \pi L\Pi_C - \pi L\Pi_B = 0 \tag{2.11}$$

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

$$\pi\left(1+r_D\right)L = 2\pi\left(1+r_L\right)L - \pi\left(1+R\right)L. \tag{2.12}$$

Combining equations (2.10) and (2.12), we solve these two equations for the expected repayments of the loan to the bank and the deposits to the depositors, respectively, to become

$$\begin{aligned} \pi\left(1+r_L\right)L &= \frac{2}{3}\pi\left(1+R\right)+\frac{1}{3}, \\ \pi\left(1+r_D\right)L &= \frac{1}{3}\pi\left(1+R\right)+\frac{2}{3}, \end{aligned} \tag{2.13}$$

from which we easily obtain the expected profits by using equation (2.8) to be

$$\Pi_B = \Pi_C = \Pi_D = \frac{1}{3}\left(\pi\left(1+R\right)-1\right)L. \tag{2.14}$$

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 (2.6). The profits when engaging in bank lending are given by equation (2.8). 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

$$\frac{\partial\mathcal{L}}{\partial\left(1+r_D\right)} = \pi L\Pi_B - \pi L\left(\Pi_D - \hat{\Pi}_D\right) = 0 \tag{2.15}$$

implies $\Pi_B = \Pi_D - \hat{\Pi}_D$. Inserting from equations (2.2) and (2.8), this can be solved for

$$2\pi\left(1+r_D\right)L = \pi\left(1+r_L\right)L + \pi\left(1+r_C\right)L - C. \tag{2.16}$$

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

$$2\pi \left(1 + r_L\right) L = \pi \left(1 + r_D\right) + \pi \left(1 + r_C\right) L + C. \tag{2.17}$$

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 (2.16) and (2.17), we get the expected repayments of the bank and direct loan, respectively, as

$$\begin{aligned} \pi \left(1 + r_L\right) L &= \pi \left(1 + r_D\right) L + \frac{2}{3} C, \\ \pi \left(1 + r_C\right) L &= \pi \left(1 + r_D\right) L + \frac{1}{3} C. \end{aligned} \tag{2.18}$$

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 (2.1), (2.2), and (2.8) that

$$\begin{aligned} \Pi_B &= \frac{2}{3} C > 0, \\ \Pi_D &= \pi \left(1 + r_D\right) L - L, \\ \Pi_C &= \pi \left(1 + R\right) L - \pi \left(1 + r_D\right) L - \frac{2}{3} C, \\ \hat{\Pi}_D &= \pi \left(1 + r_D\right) L - L - \frac{2}{3} C = \Pi_D - \frac{2}{3} C < \Pi_D, \\ \hat{\Pi}_C &= \pi \left(1 + R\right) L - \pi \left(1 + r_D\right) L - \frac{4}{3} C = \Pi_C - \frac{2}{3} C < \Pi_C. \end{aligned} \tag{2.19}$$

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

$$\pi\left(1+r_D\right)L \geq L, \tag{2.20}$$

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

$$C \leq C^{**} = \frac{3}{2}\left(\pi\left(1+R\right)L - \pi\left(1+r_D\right)L\right). \tag{2.21}$$

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 (2.19) 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 (2.6) and bank lending only in equation (2.14) 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

$$C \leq C^{***} = \frac{1}{6}\left(\pi\left(1+R\right)-1\right)L. \tag{2.22}$$

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 (2.6) and (2.19), we find that direct lending is preferred to the co-existence of direct and bank lending if

$$\pi\left(1+r_D\right)L \leq \frac{1}{2}\left(\pi\left(1+R\right)+1\right)L - C. \tag{2.23}$$

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

$$\pi\left(1+r_D\right)L \geq \frac{1}{2}\pi\left(1+R\right)L + \frac{1}{2}L + \frac{1}{3}C. \tag{2.24}$$

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 (2.14) and (2.19) that depositors prefer bank lending if

$$\pi\left(1+r_D\right)L \leq \frac{1}{3}\pi\left(1+R\right)L+\frac{2}{3}L. \tag{2.25}$$

Companies would prefer bank lending if

$$\pi\left(1+r_D\right)L \geq \frac{2}{3}\pi\left(1+R\right)L+\frac{1}{3}L-\frac{2}{3}C. \tag{2.26}$$

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 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, $\pi\left(1+R\right)$, 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 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 (2.19) 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 4, we see that bank lending dominates for high and very low negotiation costs, while direct lending can be observed

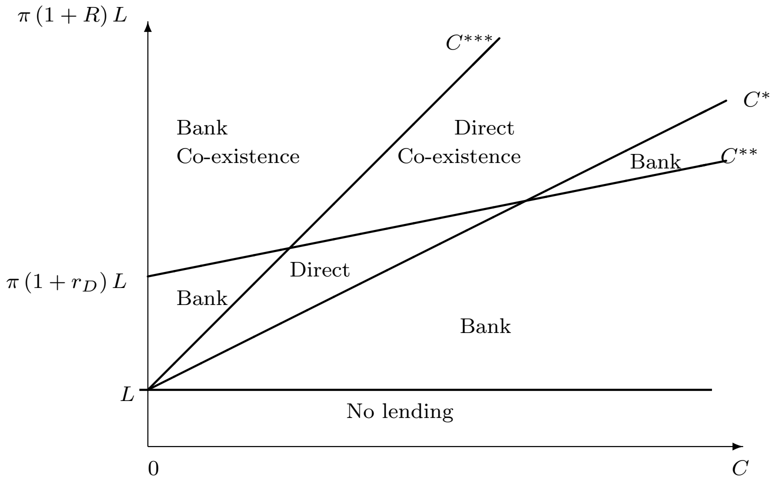


Figure 3: Equilibrium market structures with negotiation costs for direct lending

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 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.

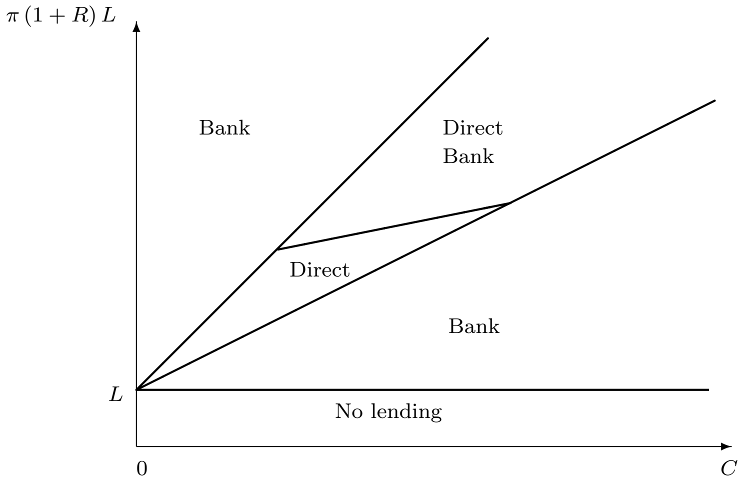


Figure 4: Equilibrium observed lending with negotiation costs for direct lending

Reading Bester (1995)

2.2 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 2.2.1, with the effect of an overall welfare increase. However, as chapter 2.2.2 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 2.2.3 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.

2.2.1 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 $\pi \left(1 + r_C\right) \frac{L}{N}$. 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

$$\pi\left(1+r_C\right)\frac{L}{N}-C\geq\frac{L}{N}. \tag{2.27}$$

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

$$\pi\left(1+r_L\right)ML-NC<\left(1+r_D\right)ML. \tag{2.28}$$

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 (2.27), representing the expected payoff for each of the N loans. This then gives us

$$\begin{aligned}\frac{1}{M}\min\left\{\pi\left(1+r_L\right)ML-NC,\left(1+r_D\right)ML\right\}\\ \geq\pi\left(1+r_C\right)L-NC.\end{aligned} \tag{2.29}$$

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 (2.28) is not fulfilled. In this case equation (2.29) reduces to $\left(1+r_D\right)L\geq\pi\left(1+r_C\right)L-NC$ 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

$$1 + r_D = \pi (1 + r_C) - \frac{NC}{L}. \tag{2.30}$$

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 (2.28), 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 (2.28) 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 (2.28) 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)

2.2.2 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

$$\Pi_B = (p\pi_H + (1-p)\,\pi_L)\,(1+r_L)\,L - (1-r_D)\,D - \frac{1}{2}cLp^2, \tag{2.31}$$

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

$$p^* = \frac{(\pi_H - \pi_L)\,(1+r_L)}{c}. \tag{2.32}$$

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 $\prod_{j=1}^{N}(1-p_j)$, assuming that the success of monitoring is independent across banks. Thus the probability of success in monitoring is given by

$$\hat{p} = 1 - \prod_{j=1}^{N}(1-p_j). \tag{2.33}$$

We can now rewrite equation (2.31) for the profits of a single bank as

$$\hat{\Pi}_B = (\hat{p}\pi_H + (1-\hat{p})\pi_L)(1+r_L)\frac{L}{N} - (1+r_D)\frac{L}{N} - \frac{1}{c}c\frac{L}{N}p_i^2, \tag{2.34}$$

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

$$\frac{\partial \hat{\Pi}_B}{\partial p_i} = \prod_{j=1, j\neq i}^{N}(1-p_j)(\pi_H-\pi_L)(1+r_L)\frac{L}{N} - \frac{cL}{N}p_i = 0, \tag{2.35}$$

where we have used equation (2.33) to replace $\hat{p}$. This expression easily solves for

$$(1-p_i)^{N-1}(\pi_H-\pi_L)(1+r_L) - cp_i = 0. \tag{2.36}$$

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 (2.36) were equal to p^* from equation (2.32), then inserting for cp^* from equation (2.32), we would require that $(\pi_H-\pi_L)(1+r_L)\left((1-p_i^*)^{N-1}-1\right) = 0$. However, the left-hand side is clearly negative unless $p_i^* = p^* = 0$, which due to equation (2.32) 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 (2.36) for $(1-p_i^*)^{N-1}$ and multiplying by $1-p_i^*$, we easily get the aggregate monitoring effort of all banks jointly as

$$\hat{p}^* = 1 - (1-p_i^*)^N = 1 - \frac{cp_i^*(1-p_i^*)}{(\pi_H - \pi_L)(1+r_L)}. \tag{2.37}$$

Comparing the expressions for $\hat{p}^*$ and p^* in equations (2.37) and (2.32), we find that $\hat{p}^* \geq p^*$ if

$$\left((\pi_H - \pi_L)(1+r_L) - \frac{1}{2}c\right)^2 \leq c^2\left(\frac{1}{4} - p_i^*(1-p_i^*)\right). \tag{2.38}$$

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.

2.2.3 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

$$\begin{aligned} \Pi_C^H &= \pi_H \left((1 + R_H) L - (1 + r_C) L \right) \\ &\geq \pi_L \left((1 + R_L) L - (1 + r_C) L \right) = \Pi_C^L, \end{aligned} \tag{2.39}$$

which solves for

$$1 + r_C \leq 1 + r_C^* = \frac{\pi_H (1 + R_H) - \pi_L (1 + R_L)}{\pi_H - \pi_L}. \tag{2.40}$$

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 $\pi_L (1 + R_L) < 1 + r_C$. 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

$$\pi_H (1 + r_C) L \geq (1 + r_D) L \tag{2.41}$$

which we can combine with equation (2.40) to obtain

$$\pi_H (1 + r_C^*) \geq \pi_H (1 + r_C) \geq 1 + r_D. \tag{2.42}$$

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

$$\pi_H \geq \pi_H^* = \frac{1 + r_D}{1 + r_C^*}. \tag{2.43}$$

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

$$\Pi_B^H = \pi_H (1 + r_L) L - (1 + r_D) L - C, \tag{2.44}$$

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

$$1 + r_L \geq 1 + r_L^* = \frac{(1 + r_D) L + C}{\pi_H L} \tag{2.45}$$

with the constraint that $\Pi_C^H \geq 0$, or $1 + r_L \leq 1 + R_H$ from equation (2.39). Using this constraint in equation (2.45), we easily obtain that this implies that

$$\pi_H \geq \pi_H^{**} = \frac{(1 + r_D) L + C}{(1 + R_H) L}. \tag{2.46}$$

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 (2.43) and (2.46), 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 (2.40), this becomes

$$C < \frac{\pi_L (R_H - R_L)(1+r_D)}{\pi_H (1+R_H) - \pi_L (1+R_L)} L. \tag{2.47}$$

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 $\pi \in [\pi_H^*; 1]$ will choose direct lending, while more risky companies with $\pi \in [\pi_H^{**}; \pi_H^*]$ 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 (2016, ch. 1.3.2) or Freixas & Rochet (2008, 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.

2.3 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

$$1 + r_C \geq \frac{1}{\pi} + \frac{z(1-\pi)(1+r_C)^2 L}{2}. \tag{2.48}$$

This solves for a minimum direct loan rate of

$$1 + r_C \geq 1 + r_C^* = \frac{1}{z(1-\pi)L} - \sqrt{\frac{1}{z^2(1-\pi)^2 L^2} - \frac{2}{z\pi(1-\pi)L}}. \tag{2.49}$$

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 $u\left(\pi(1+r_L)NL-(1+r_D)NL-\frac{1}{2N^2}z\pi(1-\pi)(1+r_L)^2N^2L^2\right)$, 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

$$(2.50)\qquad 1+r_L \geq \frac{1+r_D}{\pi}+\frac{z(1-\pi)(1+r_L)^2L}{2N},$$

solving for the minimum bank loan rate

$$(2.51)\quad 1+r_L \geq 1+r_L^* = \frac{N}{z(1-\pi)L}-\sqrt{\frac{N^2}{z^2(1-\pi)^2L^2}-\frac{2N(1+r_D)}{z\pi(1-\pi)L}}.$$

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, $u\left(\pi\left(1+r_C\right)L-\frac{1}{2}z\pi\left(1-\pi\right)\left(1+r_C\right)^2L^2\right)$, we get from comparing coefficients that

$$1+r_D \geq \pi\left(1+r_C\right)-\frac{1}{2}z\pi\left(1-\pi\right)\left(1+r_C\right)^2L. \tag{2.52}$$

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 (2.50), the condition for bank lending to occur can be rewritten as

$$1+r_D \leq \pi\left(1+r_L\right)-\frac{1}{2N}z\pi\left(1-\pi\right)\left(1+r_L\right)^2L. \tag{2.53}$$

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

$$\begin{aligned}\pi\left(1+r_C\right)-\frac{1}{2}z\pi\left(1-\pi\right)\left(1+r_C\right)^2L &\leq 1+r_D \\ &\leq \pi\left(1+r_L\right)-\frac{1}{2N}z\pi\left(1-\pi\right)\left(1+r_L\right)^2L.\end{aligned} \tag{2.54}$$

If we now insert for r_C the value of r_C^* of equation (2.49) and for r_L the value of r_L^* from equation (2.51) 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.

3

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 3.1, with alternative banking specifications explored in chapter 3.2. It is not only that banks provide this liquidity for the benefit of their depositors, but, as we will see in chapter 3.3, 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 3.4.

It is thus that banks increase the welfare of their customers not only by reducing transaction costs as discussed in chapter 2, 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.

3.1 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

$$E\left[U(D)\right] = pu\left(\left(1+r_D^1\right)D\right) + (1-p)\,u\left(\left(1+r_D^2\right)D\right), \tag{3.1}$$

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

$$\begin{aligned} p\left(1+r_D^1\right)D &= D-L, \\ (1-p)\left(1+r_D^2\right)D &= \pi(1+r_L)L, \end{aligned} \tag{3.2}$$

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

$$p\left(1+r_D^1\right) + (1-p)\frac{\left(1+r_D^2\right)}{\pi(1+r_L)} = 1. \tag{3.3}$$

Depositors will maximize their expected utility in equation (3.1), subject to constraint (3.3), 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 (3.1). 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

$$\begin{aligned}\frac{E\left[U\left(D\right)\right]}{\partial\left(1+r_D^1\right)D} &= p\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} \\ &\quad +\left(1-p\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}\left(-\frac{p}{1-p}\pi\left(1+r_L\right)\right) \\ &= 0,\end{aligned} \tag{3.4}$$

which solves for

$$\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} = \pi\left(1+r_L\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}, \tag{3.5}$$

implying that the marginal rate of substitution equals the expected return on the loan. Knowing the utility function $u(\cdot)$, 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

$$\left(1+r_D^1\right)D = D - L + \lambda L = D - \left(1-\lambda\right)L \leq D \tag{3.6}$$

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

$$\begin{aligned}\left(1+r_D^2\right)D &= D - L + \pi\left(1+r_L\right)L \\ &= D - \left(1-\pi\left(1+r_L\right)\right)L \\ &\leq \pi\left(1+r_L\right)D,\end{aligned} \tag{3.7}$$

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 (3.1) remains unchanged. Using the inequalities in equations (3.6) and (3.7), we can now obtain the constraint to our optimization as

$$p\left(1+r_D^1\right)+(1-p)\frac{\left(1+r_D^2\right)}{\pi\left(1+r_L\right)}\leq 1 \tag{3.8}$$

which is more stringent than (3.3) 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

$$\left(1+r_D^1\right)D=D-L+\pi\left(1+r_L\right)LP=D-\left(1-\pi\left(1+r_L\right)P\right)L. \tag{3.9}$$

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

$$\begin{aligned}\left(1+r_D^2\right)D &= \frac{D-L}{P}+\pi\left(1+r_L\right)L \\ &= \frac{1}{P}\left(D-L+P\pi\left(1+r_L\right)L\right) \\ &= \frac{\left(1+r_D^1\right)D}{P}.\end{aligned} \tag{3.10}$$

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 (3.9) 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 (3.10), 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

$$P=\frac{1}{\pi\left(1+r_L\right)}. \tag{3.11}$$

Inserting this into equations (3.9) and (3.10), we obtain

$$\begin{aligned}(1+r_D^1)\,D &= D, \\ (1+r_D^2)\,D &= \pi\,(1+r_L)\,D\end{aligned} \tag{3.12}$$

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.

$$pP\pi\,(1+r_L)\,L = (1-p)\,(D-L)\,, \tag{3.13}$$

which solves, when inserting from equation (3.11), for

$$L = (1-p)\,D < D. \tag{3.14}$$

Using equation (3.12) we obtain the constraint on optimization as

$$p\left(1+r_D^1\right) + (1-p)\,\frac{\left(1+r_D^2\right)}{\pi\,(1+r_L)} = 1, \tag{3.15}$$

identical to the social optimum constraint in equation (3.3). However the first order condition for an optimum, identical to equation (3.5), would only be fulfilled if the utility function is such that

$$\frac{\partial u\,(D)}{\partial\,(1+r_D^1)} = \pi\,(1+r_L)\,\frac{\partial u\,(\pi\,(1+r_L)\,D)}{\partial\,(1+r_D^2)}. \tag{3.16}$$

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 $\pi\,(1+r_L)\,\frac{\partial u(\pi(1+r_L)D)}{\partial D} < \frac{\partial u(D)}{\partial D}$, 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 (3.2) and (3.5) can be achieved. The constraints in equation (3.2) are trivially fulfilled and we can set deposit rates in line with the requirements of equation (3.5). 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 (3.5); 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 (3.5) can never be fulfilled as this would require $\pi\left(1+r_L\right) < 1$, but for a loan to be viable we need that $\pi\left(1+r_L\right) \geq 1$ 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 (3.2) 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 $\frac{D-L}{pP} = \pi\left(1+r_L\right)\frac{D-L}{p}$ in time period 2. The amount $\frac{D-L}{p}$ would buy up all the loans which are worth $P\pi\left(1+r_L\right)L = L$ if using equation (3.11) 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 $\frac{D-L}{pP} = \pi\left(1+r_L\right)\frac{D-L}{p} = \pi\left(1+r_L\right)L$. When retaining the deposit, depositors obtain $\frac{\pi(1+r_L)L}{1-p}$ as derived from equation (3.2), 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 emptyhanded. 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 13.

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 13 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)

3.2 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 3.1 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 13. 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 14) or through regulation and government bailouts (see chapter 16 and part VII), have also to be considered for a complete assessment.

3.2.1 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

$$\left(1+r_D^1\right) D = D - L, \tag{3.17}$$

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

$$\left(1+r_D^2\right) D = \pi \left(1+r_L\right) L + \left(D - L\right), \tag{3.18}$$

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 (3.2) of chapter 3.1, which we reproduce here:

$$\begin{aligned} \left(1+r_D^1\right) D &= (D-L) + \lambda L, \\ \left(1+r_D^2\right) D &= \pi \left(1+r_L\right) L + \left(D - L\right). \end{aligned} \tag{3.19}$$

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)

3.2.2 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

$$p\left(1+r_D^1\right)D = p\left(D-L\right)+pS. \tag{3.20}$$

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 $\pi\left(1+r_L\right)L$ and the shares they bought, each providing them with a payment of $\pi\left(1+r_L\right)L$, giving them a fraction $\frac{(1-p)(D-L)}{S}$ of shares, such that

$$\left(1-p\right)\left(1+r_D^2\right)D = \left(1+\frac{D-L}{S}\right)\left(1-p\right)\pi\left(1+r_L\right)L, \tag{3.21}$$

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 $\left(1-p\right)\left(D-L\right)$. We get from equalling these two expressions that

$$S = \frac{1-p}{p}\left(D-L\right). \tag{3.22}$$

Inserting this into equations (3.20) and (3.21), we get the repayments in time periods 1 and 2, respectively as

$$\begin{aligned} \left(1+r_D^1\right)D &= \frac{D-L}{p}, \\ \left(1+r_D^2\right)D &= \frac{\pi\left(1+r_L\right)}{1-p}L. \end{aligned} \tag{3.23}$$

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 (3.22) 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

$$p\left(1+r_D^1\right)+(1-p)\frac{\left(1+r_D^2\right)}{\pi\left(1+r_L\right)}=1, \tag{3.24}$$

which is identical to the social optimum we obtained in in equation (3.3) of chapter 3.1. Hence equity contracts should be optimal. Furthermore, a bank run where all depositors withdraw their deposit in time period 1 (see chapter 13) 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 14 and 16 as well as part VII.

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).

3.3 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 3.1 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

$$\Pi_D = \pi(1+r_D)D - D, \tag{3.25}$$

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

$$1 + r_D = \frac{1}{\pi}. \tag{3.26}$$

The profits of the bank are then given by

$$\Pi_B = \pi((1+r_L)L - (1+r_D)D) = (\pi(1+r_L)-1)D. \tag{3.27}$$

We assume here that $L = D$ such that banks do not hold any excess cash and the final equality emerges when inserting equation (3.26). 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

$$\hat{\Pi}_D = \left(1+\hat{r}_D^1\right)D - D + \pi\left(1+\hat{r}_D^2\right)D - D, \tag{3.28}$$

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

$$1 + \hat{r}_D^2 = \frac{1 - \hat{r}_D^1}{\pi}. \tag{3.29}$$

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

$$\hat{\Pi}_B = -\hat{r}_D^1 D + \pi \left(r_L - \hat{r}_D^2 \right) D = \left(\pi \left(1 + r_L \right) - 1 \right) D, \tag{3.30}$$

where the final equality arises by inserting from equation (3.29). 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 (3.27) and (3.30), 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 3.1, 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)

3.4 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

$$\Pi_B = E\left[r_L L + \hat{r}_L \hat{L} + \mathbf{1}_{\hat{L}} r_L \hat{L} - r_M M - \frac{1}{2}cM^2 - (1-\mathbf{1}_D)\, r_D D\right], \tag{3.31}$$

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 $E\left[\mathbf{1}_D\right] = \lambda$. 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

$$\begin{aligned} Prob\left(\mathbf{1}_D = \mathbf{1}_{\hat{L}} = 1\right) &= Prob\left(\mathbf{1}_D = \mathbf{1}_{\hat{L}} = 0\right) = \frac{1}{2}\rho, \\ Prob\left(\mathbf{1}_D = 1, \mathbf{1}_{\hat{L}} = 0\right) &= Prob\left(\mathbf{1}_D = 0, \mathbf{1}_{\hat{L}} = 1\right) = \frac{1}{2}\left(1-\rho\right), \end{aligned} \tag{3.32}$$

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

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

$$\begin{aligned} L + C &= D, \\ L + \mathbf{1}_{\hat{L}}\hat{L} &= \left(1-\mathbf{1}_D\right)D + M, \end{aligned} \tag{3.33}$$

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 (3.33) we easily get by eliminating L that

$$M = \mathbf{1}_{\hat{L}}\hat{L} + \mathbf{1}_D D - C, \tag{3.34}$$

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 (3.34) into equation (3.31) 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:

$$\begin{aligned}\frac{\partial \Pi_B}{\partial \hat{L}} &= \hat{r}_L + \gamma\left(r_L - r_M\right) - \frac{1}{2}c\frac{\partial E\left[M^2\right]}{\partial \hat{L}} = 0, \\ \frac{\partial \Pi_B}{\partial C} &= -r_M - \frac{1}{2}c\frac{\partial E\left[M^2\right]}{\partial C} = 0.\end{aligned} \tag{3.35}$$

When using equation (3.32), we obtain that

$$E\left[M^2\right] = \frac{1}{2}\rho\left(\hat{L} + D - C\right)^2 + \frac{1}{2}\left(1-\rho\right)\left(\left(\hat{L} - C\right)^2 + \left(D - C\right)^2\right), \tag{3.36}$$

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 (3.36) we then get

$$\begin{aligned}\frac{\partial E\left[M^2\right]}{\partial \hat{L}} &= \hat{L} - C + \rho D, \\ \frac{\partial E\left[M^2\right]}{\partial C} &= \left(2-\rho\right)C - \hat{L} - D.\end{aligned} \tag{3.37}$$

Inserting the second equation in (3.37) into the final equation in (3.35), we get the optimal cash holdings as

$$C = \frac{\hat{L} + D - \frac{2r_M}{c}}{2-\rho}. \tag{3.38}$$

Using equation (3.38) in the first equation of (3.37), we get from the first equation in (3.35) that the optimal amount of credit lines is given by

$$\hat{L} = \left(1-\rho\right)D + 2\frac{\left(2-\rho\right)\left(\hat{r}_L + \gamma\left(r_L - r_M\right)\right) - r_M}{c\left(1-\rho\right)}. \tag{3.39}$$

We instantly see from equation (3.39) that

$$\begin{aligned}\frac{\partial \hat{L}}{\partial D} &= 1 - \rho < 1, \\ \frac{\partial \hat{L}}{\partial \rho} &= -D + 2\frac{\left(\hat{r}_L - r_M\right) + \gamma\left(r_L - r_M\right)}{c\left(1-\rho\right)^2}.\end{aligned} \tag{3.40}$$

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 (3.40) 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 (3.40) 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.

4 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

$$Prob\,(s = H|H) = Prob\,(s = L|L) = p_j. \tag{4.1}$$

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

$$\begin{aligned} Prob\,(H|s=H) &= p_j^H = \frac{pp_j}{pp_j + (1-p)(1-p_j)}, \\ Prob\,(L|s=L) &= p_j^L = \frac{p(1-p_j)}{p(1-p_j) + (1-p)p_j}, \end{aligned} \tag{4.2}$$

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

$$\begin{aligned} \Pi_D^H &= p_D^H \pi_H \left(1 + r_L^H\right) L + \left(1 - p_D^H\right) \pi_L (1 + R_L) L - L, \\ \Pi_D^L &= p_D^L \pi_L (1 + R_L) L + \left(1 - p_D^L\right) \pi_H \left(1 + r_L^H\right) L - L. \end{aligned} \tag{4.3}$$

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 (4.3) to

$$\begin{aligned} \Pi_D^H &= p_D^H \pi_H \left(1 + r_L^H\right) L - L, \\ \Pi_D^L &= \left(1 - p_D^L\right) \pi_H \left(1 + r_L^H\right) L - L. \end{aligned} \tag{4.4}$$

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

$$
\begin{aligned}
1 + r_L^H &= \frac{1}{p_D^H \pi_H}, \\
1 + r_L^L &= \frac{1}{\left(1 - p_D^L\right) \pi_H}.
\end{aligned}
\tag{4.5}
$$

We see from equations (4.2) and (4.5) that

$$
\begin{aligned}
\frac{\partial p_j^H}{\partial p} &= \frac{p_j\left(1-p_j\right)}{\left(p p_j + (1-p)\left(1-p_j\right)\right)^2}, \\
\frac{\partial p_j^L}{\partial p} &= \frac{p_j\left(1-p_j\right)}{\left(p\left(1-p_j\right) + (1-p) p_j\right)^2}, \\
\frac{\partial\left(1+r_L^H\right)}{\partial p} &= -\frac{\partial p_D^H}{\partial p} \frac{1}{\left(p_D^H\right)^2 \pi_H}, \\
\frac{\partial\left(1+r_L^L\right)}{\partial p} &= \frac{\partial p_D^L}{\partial p} \frac{1}{\left(1-p_D^L\right)^2 \pi_H}.
\end{aligned}
\tag{4.6}
$$

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

$$
\Pi_C = p\left(p_D \pi_H\left(R_H - r_L^H\right) + \left(1 - p_D\right) \pi_H\left(R_H - r_L^L\right)\right) - C,
\tag{4.7}
$$

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

$$
\begin{aligned}
\frac{\partial C}{\partial p} = {} & p_D \pi_H\left(R_H - r_L^H\right) + \left(1 - p_D\right) \pi_H\left(R_H - r_L^H\right) \\
& - p \pi_H p_D \frac{\partial\left(1+r_L^H\right)}{\partial p} - p \pi_H\left(1-p_D\right) \frac{\partial\left(1+r_L^L\right)}{\partial p}.
\end{aligned}
\tag{4.8}
$$

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

$$1 + \hat{r}_L^H = \frac{1}{p_B^H \pi_H}, \tag{4.9}$$

$$1 + \hat{r}_L^L = \frac{1}{\left(1 - p_B^L\right) \pi_H}. \tag{4.10}$$

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

$$\begin{aligned} \hat{\Pi}_C = \; & p \left(p_B \pi_H \left(R_H - \hat{r}_L^H \right) + \left(1 - p_B\right) p_D \pi_H \left(R_H - r_L^H \right) \right. \\ & \left. + \left(1 - p_B\right) \left(1 - p_D\right) \pi_H \left(R_H - r_L^L \right) \right) - \hat{C}. \end{aligned} \tag{4.11}$$

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

$$\text{(4.12)}\quad \begin{aligned}\frac{\partial \hat{C}}{\partial p} &= p_B\pi_H\left(R_H-\hat{r}_L^H\right)+(1-p_B)\,p_D\pi_H\left(R_H-r_L^H\right)\\ &\quad +(1-p_B)\,(1-p_D)\,\pi_H\left(R_H-r_L^L\right)\\ &\quad -pp_B\pi_H\frac{\partial\left(1+\hat{r}_L^H\right)}{\partial p}-p\,(1-p_B)\,p_D\pi_H\frac{\partial\left(1+r_L^H\right)}{\partial p}\\ &\quad -p\,(1-p_B)\,(1-p_D)\,\pi_H\frac{\partial\left(1+r_L^L\right)}{\partial p}.\end{aligned}$$

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 (4.8) and (4.12), we can rewrite the first order condition in equation (4.12) as

$$\text{(4.13)}\qquad \frac{\partial \hat{C}}{\partial p}=\frac{\partial C}{\partial p}\,(1-p_B)+p_B\pi_H\left(R_H-\hat{r}_L^H\right)-pp_B\pi_H\frac{\partial\left(1+\hat{r}_L^H\right)}{\partial p}.$$

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

$$\text{(4.14)}\qquad \begin{aligned}\frac{\partial C}{\partial p} &< \pi_H\left(\left(R_H-\hat{r}_L^H\right)-p\frac{\partial\left(1+\hat{r}_L^H\right)}{\partial p}\right)\\ &= \pi_H\,(1+R_H)+\frac{1-2p_B}{p_B},\end{aligned}$$

where we obtained the second line by inserting for $\hat{r}_L^H$ from equation (4.9) 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 (4.9). 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 (4.14) 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 ?

Review

IF BANKS WERE MERE INTERMEDIARIES between savers and borrowers, their impact would rather minimal. Chapter 1 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 2. 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 3, 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 5. 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 6 and the use of collateral is then explored in more detail in chapter 8 and in chapter 7 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 9 will show. Banks having had interactions with a company before will have an informational advantage over other banks that lack such interactions. In chapter 10 we will explore the consequences of such 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 11, securitization allows the bank to free up resources to grant additional loans. We will see under which conditions such securitization is optimal.

5

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 5.1 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 3. 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 5.3 will explore under which conditions such an extended maturity of loans is optimal. Before this discussion, we will discuss in chapter 5.2 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 10 for a discussion of relationship banking, this is not always the main reason. In chapter 5.4 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.

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.

5.1 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 5.1.3 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 5.1.1 and 5.1.2 will show that if outcomes of investments are common knowledge, different loan contracts are optimal.

5.1.1 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

$$\mathcal{L} = E[U_C(V - R(V))] - \lambda \left(E[U_B(R(V))] - U_B^0 \right). \tag{5.1}$$

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 (5.1).

The first order condition of our maximization yields

$$\frac{\partial \mathcal{L}}{\partial R(V)} = -\frac{\partial E[U_C(V - R(V))]}{\partial R(V)} + \lambda \frac{\partial E[U_B(R(V))]}{\partial R(V)} = 0. \tag{5.2}$$

Using the implicit function theorem to solve equation (5.2) for $R(V)$, we get

$$\begin{aligned} \frac{\partial R(V)}{\partial V} &= -\frac{\frac{\partial^2 \mathcal{L}}{\partial R(V) \partial V}}{\frac{\partial^2 \mathcal{L}}{\partial R(V)^2}} \\ &= \frac{\frac{\partial^2 U_C(V - R(V))}{\partial R(V) \partial V}}{\frac{\partial^2 U_C(V - R(V))}{\partial R(V)^2} + \lambda \frac{\partial^2 U_B(R(V))}{\partial R(V)^2}}. \end{aligned} \tag{5.3}$$

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 (5.2) for λ and insert this expression into (5.3) to obtain

$$\frac{\partial R(V)}{\partial V} = \frac{z_C}{z_C + z_B} > 0. \tag{5.4}$$

Integrating this equation the repayment function becomes

$$R(V) = R_0 + \frac{z_C}{z_C + z_B} V, \tag{5.5}$$

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 (2008, Ch. 4.1), Keiding (2016, Ch. 5.2)

5.1.2 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 $R\left(V\right)$, the company profits are the given by

$$\Pi_C = \pi\left(V - R\left(V\right)\right) - C. \tag{5.6}$$

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 (5.6):

$$\begin{aligned} R(V) &\leq V \\ \pi R(V) &\geq (1+r_D)L \\ \Pi_C^* &\geq \Pi_C. \end{aligned} \tag{5.7}$$

The optimality of the effort level, $\Pi_C^* \geq \Pi_C$, can be replaced by its first order condition, which from equation (5.6) is easily obtained as

$$\frac{\partial \Pi_C}{\partial e} = \frac{\partial \pi}{\partial e}(V - R(V)) - \frac{\partial C}{\partial e} = 0. \tag{5.8}$$

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

$$\begin{aligned} \mathcal{L} &= \pi (V - R(V)) - C + \lambda_1 \left(\frac{\partial \pi}{\partial e}(V - R(V)) - \frac{\partial C}{\partial e} \right) \\ &\quad + \lambda_2 (\pi R(V) - (1+r_D)L) \\ &= \left(V \left(\pi + \lambda_1 \frac{\partial \pi}{\partial e} \right) - C - \lambda_1 \frac{\partial C}{\partial e} \lambda_2 - (1+r_D)L \right) \\ &\quad + R(V) \left((\lambda_2 - 1)\pi - \lambda_1 \frac{\partial \pi}{\partial e} \right), \end{aligned} \tag{5.9}$$

with λ_i denoting the Lagrange multipliers. The first term in the second equality does not depend on $R(V)$ 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 (5.9) that the Lagrangian is increasing in the repayment $R(V)$ and thus it is optimal for the company that the maximal repayment $R(V) = V$ is chosen. If $\frac{\frac{\partial \pi}{\partial e}}{\pi} > \frac{\lambda_2 - 1}{\lambda_1}$, then the Lagrangian is decreasing in the repayment $R(V)$ and we should choose the lowest possible repayment, $R(V) = 0$.

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 $R(V) = V$ and for higher outcomes we require no repayment $R(V) = 0$.

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 $(1 + r_D) L$ 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 $R(V) = 0$ 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)

5.1.3 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 $R(V_i) = \overline{R} = \inf \{R(V_j) | V_j \notin \mathfrak{A}\}$.

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

$$R(V) = \min\{V, \overline{R}\}, \tag{5.10}$$

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, $\text{Prob}\,(V \in \mathfrak{A})\, C$, 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 (5.10), the third equality from $V \in \mathfrak{A}$ and the final inequality from the contract specification in equation (5.10), again. Thus the contract specified in equation (5.10) 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.

5.2 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

$$\pi = \nu\pi_H + (1-\nu)\pi_L. \tag{5.11}$$

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

$$\gamma = \text{Prob}\,(s = H|H) = \text{Prob}\,(s = L|L) > \frac{1}{2}. \tag{5.12}$$

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

$$\begin{aligned} \text{Prob}(s = H) &= \text{Prob}\,(s = H|H)\,\text{Prob}\,(H) \\ &\quad + \text{Prob}\,(s = H|L)\,\text{Prob}\,(L) \\ &= \gamma\nu + (1-\gamma)(1-\nu), \\ \text{Prob}(s = L) &= \text{Prob}\,(s = L|H)\,\text{Prob}\,(H) \\ &\quad + \text{Prob}\,(s = L|L)\,\text{Prob}\,(L) \\ &= \gamma(1-\nu) + (1-\gamma)\nu. \end{aligned} \tag{5.13}$$

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 (5.12) 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

$$\begin{aligned} \text{Prob}\,(H|s = H) &= \frac{\text{Prob}\,(s = H|H)\,\text{Prob}\,(H)}{\text{Prob}\,(s = H)} \\ &= \frac{\gamma\nu}{\gamma\nu + (1-\gamma)(1-\nu)}, \\ \text{Prob}\,(L|s = H) &= 1 - \text{Prob}\,(H|s = H) \\ &= \frac{(1-\gamma)(1-\nu)}{\gamma\nu + (1-\gamma)(1-\nu)}, \\ \text{Prob}\,(H|s = L) &= \frac{\text{Prob}\,(s = L|H)\,\text{Prob}\,(H)}{\text{Prob}\,(s = L)} \\ &= = \frac{(1-\gamma)\nu}{(1-\gamma)\nu + \gamma(1-\nu)}, \\ \text{Prob}\,(L|s = L) &= 1 - \text{Prob}\,(H|s = L) \\ &= \frac{\gamma(1-\nu)}{(1-\gamma)\nu + \gamma(1-\nu)}, \end{aligned} \tag{5.14}$$

where the second equalities arise from using the definition of γ in equation (5.12) 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

$$\begin{aligned} \hat{\pi}_H &= \pi_H \text{Prob}\left(H|s=H\right) + \pi_L \text{Prob}\left(L|s=H\right), \\ \hat{\pi}_L &= \pi_H \text{Prob}\left(H|s=L\right) + \pi_L \text{Prob}\left(L|s=L\right). \end{aligned} \tag{5.15}$$

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 (5.13), (5.14), and (5.15), we can rewrite the probabilities of observing signal H and L, respectively, as

$$\begin{aligned} \text{Prob}\left(s=H\right) &= \frac{\pi - \hat{\pi}_L}{\hat{\pi}_H - \hat{\pi}_L}, \\ \text{Prob}\left(s=L\right) &= \frac{\hat{\pi}_H - \pi}{\hat{\pi}_H - \hat{\pi}_L}. \end{aligned} \tag{5.16}$$

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

$$\begin{aligned} \hat{\Pi}_B^U &= \pi(1+r_L^U)L - (1+r_D)L, \\ \hat{\Pi}_B^H &= \hat{\pi}_H(1+r_L^H)L - (1+r_D)L, \\ \hat{\Pi}_B^L &= \hat{\pi}_L(1+r_L^L)L - (1+r_D)L \end{aligned} \tag{5.17}$$

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

$$
\begin{aligned}
1+r_L^U &\geq 1+\hat{r}_L^U = \frac{1+r_D}{\pi}, \\
1+r_L^H &\geq 1+\hat{r}_L^H = \frac{1+r_D}{\hat{\pi}_H}, \\
1+r_L^L &\geq 1+\hat{r}_L^L = \frac{1+r_D}{\hat{\pi}_L},
\end{aligned}
\tag{5.18}
$$

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

$$\begin{aligned} \lambda_U(r) &= \text{Prob}\left(r_L^U < r\right), \\ \lambda_H(r) &= \text{Prob}\left(r_L^H < r\right), \\ \lambda_L(r) &= \text{Prob}\left(r_L^L < r\right), \end{aligned} \tag{5.19}$$

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 $\text{Prob}(s = H)$ ($\text{Prob}(s = L)$), 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 $\lambda_H\left(r_L^U\right)$ $\left(\lambda_L\left(r_L^U\right)\right)$ and hence the uninformed bank will offer a lower loan rate with probability $1-\lambda_H\left(r_L^U\right)$ $\left(1-\lambda_L\left(r_L^U\right)\right)$.

For informed banks, knowing the signal they have received, their profits will be the value of the loan after having repaid their depositors, $\left(\hat{\pi}_s\left(1+r_L^U\right)-\left(1+r_D\right)\right)L$, 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

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

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 (5.17), 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 (5.17) 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

$$\begin{aligned} (5.21)\ \Pi_B^H &= \left(1 - \lambda_U\left(r_L^H\right)\right)\left(\hat{\pi}_H\left(1 + r_L^H\right) - (1 + r_D)\right)L \\ &= \left(\frac{\hat{\pi}_H}{\pi} - 1\right)(1 + r_D)L, \\ \Pi_B^L &= 0, \\ \Pi_B^U &= \text{Prob}(s = H)\left(1 - \lambda_H\left(r_L^U\right)\right)\left(\hat{\pi}_H\left(1 + r_L^U\right) - (1 + r_D)\right)L \\ &\quad + \text{Prob}(s = L)\left(\hat{\pi}_L\left(1 + r_L^U\right) - (1 + r_D)\right)L \\ &= 0. \end{aligned}$$

Using equation (5.16), these expressions can be solved for the probability distribution of loan rates, which take the form

$$\begin{aligned} (5.22) \qquad \lambda_U\left(r_L^H\right) &= \frac{\hat{\pi}_H}{\pi}\frac{\pi\left(1 + r_L^H\right) - (1 + r_D)}{\hat{\pi}_H\left(1 + r_L^H\right) - (1 + r_D)}, \\ \lambda_H\left(r_L^U\right) &= \frac{\hat{\pi}_H - \hat{\pi}_L}{\pi - \hat{\pi}_L}\frac{\pi\left(1 + r_L^U\right) - (1 + r_D)}{\hat{\pi}_H\left(1 + r_L^U\right) - (1 + r_D)}. \end{aligned}$$

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 (5.15), it is easy to see how the belief on the success rate changes as the precision of information changes. We obtain

$$\begin{aligned} (5.23) \qquad \frac{\partial \hat{\pi}_H}{\partial \gamma} &= \frac{\nu(1 - \nu)(\pi_H - \pi_L)}{(\gamma\nu + (1 - \gamma)(1 - \nu))^2} > 0, \\ \frac{\partial \hat{\pi}_L}{\partial \gamma} &= \frac{\nu(1 - \nu)(\pi_L - \pi_H)}{((1 - \gamma)\nu + \gamma(1 - \nu))^2} < 0, \end{aligned}$$

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

$$\begin{aligned} \frac{\partial \lambda_U\left(r_L^H\right)}{\partial \gamma} &= \frac{\partial \lambda_U\left(r_L^H\right)}{\partial \hat{\pi}_H}\frac{\partial \hat{\pi}_H}{\partial \gamma} < 0, \\ \frac{\partial \lambda_H\left(r_L^U\right)}{\partial \gamma} &= \frac{\partial \lambda_H\left(r_L^U\right)}{\partial \hat{\pi}_L}\frac{\partial \hat{\pi}_L}{\partial \gamma} < 0. \end{aligned} \tag{5.24}$$

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

$$\hat{\hat{r}}_L^H = \min\left\{r_L^H, r_L^U\right\}. \tag{5.25}$$

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

$$1-\hat{\lambda}_H\left(\hat{\hat{r}}_L^H\right) = \left(1-\lambda_H\left(r_L^U\right)\right)\left(1-\lambda_U\left(r_L^H\right)\right), \tag{5.26}$$

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

Using the results from equation (5.24), we easily obtain

$$\begin{aligned} \frac{\partial(1-\hat{\lambda}_H\left(\hat{\hat{r}}_L^H\right))}{\partial\gamma} &= -\frac{\partial\lambda_H\left(r_L^U\right)}{\partial\gamma}\left(1-\lambda_U\left(r_L^H\right)\right) \\ &\quad -\frac{\partial\lambda_U\left(r_L^H\right)}{\partial\gamma}\left(1-\lambda_H\left(r_L^U\right)\right) \\ &> 0. \end{aligned} \tag{5.27}$$

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

$$E[\hat{\hat{r}}_L^H] = \int_0^{+\infty} \hat{\hat{r}}_L^H \left(1-\hat{\lambda}_H\left(\hat{\hat{r}}_L^H\right)\right) d\hat{\hat{r}}_L^H. \tag{5.28}$$

This expression is increasing in the precision of the signal, γ. This is as $1-\hat{\lambda}_H\left(\hat{\hat{r}}_L^H\right)$ is increasing in this precision as shown in equation (5.27) 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)

5.3 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 3. 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

$$\begin{aligned}\Pi_C^i &= \pi_i\left((V-(1+r_L)L)+\pi_i\left(V-(1+r_L)\,L\right)\right)\\ &= \pi_i\left(1+\pi_i\right)\left(V-(1+r_L)\,L\right),\end{aligned} \tag{5.29}$$

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

$$\Pi_B^i = \pi_i\left(1+r_L\right)L-\left(1+r_D\right)D. \tag{5.30}$$

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

$$1+r_L = \frac{1+r_D}{\pi_i}. \tag{5.31}$$

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 (5.30) after inserting all relevant variables.

Inserting equation (5.31) into equation (5.29), company profits are given by

$$\Pi_C^i = \pi_i \left(1 + \pi_i\right) V - \frac{\pi_i \left(1 + \pi_i\right)}{\pi_L} \left(1 + r_D\right) L. \tag{5.32}$$

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 $V - L < \left(1 + \hat{r}_L\right) L$, or $V < \left(2 + \hat{r}_L\right) L$, 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

$$\hat{\Pi}_C^i = \pi_i^2 \left(2V - L - \left(1 + \hat{r}_L\right) L\right). \tag{5.33}$$

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

$$\hat{\Pi}_B^i = \pi_i^2 \left(1 + \hat{r}_L\right) L + \pi_i \left(1 - \pi_i\right) \left(V - L\right) - \left(1 + \hat{r}_D\right) L. \tag{5.34}$$

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

$$1 + \hat{r}_L = \frac{1 + \hat{r}_D}{\pi_i^2} - \frac{1 - \pi_i}{\pi_i} \frac{V - L}{L}. \tag{5.35}$$

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 (5.33), we can obtain the company profits as

$$\hat{\Pi}_C^i = \frac{\pi_i^2}{\pi_H^2} \left(\pi_H \left(1 + \pi_H \right) V - \left(\left(1 + \hat{r}_D \right) + \pi_H \right) L \right). \tag{5.36}$$

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 (5.32) and (5.36), while noting that $1 + \hat{r}_D = (1 + r_D)^2$, this requirement can be written as

$$\pi_L \leq \pi_L^* = \frac{\pi_H \left(1 + \pi_H \right) \left(1 + r_D \right)}{\left(1 + r_D \right)^2 + \pi_H} < \pi_H. \tag{5.37}$$

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 (5.31). 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 (5.32) and (5.36), while noting that $1 + \hat{r}_D = (1 + r_D)^2$, this becomes

$$\pi_L^2 (1 + r_D)^2 + \pi_L^2 \pi_H - \pi_H^2 (1 + \pi_H)(1 + r_D) + \pi_H \pi_L (\pi_H - \pi_L) \frac{V}{L} \geq 0. \tag{5.38}$$

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 (5.37) and (5.38), we require that

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

which admits a viable solution only if

$$\pi_L \geq \pi_L^{**} = (1 + \pi_H)(1 + r_D) \frac{L}{V}. \tag{5.39}$$

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 (5.38) 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)

5.4 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

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

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

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

which solves for

$$1+r_L^i=\frac{\overline{V}-C_i}{L_i}. \tag{5.42}$$

If we assume that banks are competitive, they will make zero profits. Thus with $\Pi_B^i=0$ and inserting equation (5.42) into equation (5.40), we easily get the loan the bank provides to be of size

$$L_i=\frac{\left(\overline{V}-C_i\right)^2}{2\overline{V}\left(1+r_D^i\right)}. \tag{5.43}$$

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

$$\begin{aligned}\Pi_C^i &= \int_{\left(1+r_L^i\right)L_i}^{\overline{V}}\left(V-\left(1+r_L^i\right)L_i\right)dF\left(V\right)-\left(1+r_E\right)\left(I-L_i\right) \\ &= \frac{C_i^2}{2\overline{V}}+\frac{\left(\overline{V}-C_i\right)^2\left(1+r_E\right)}{2\overline{V}\left(1+r_D^i\right)}-\left(1+r_E\right)I,\end{aligned} \tag{5.44}$$

after inserting from equations (5.42) and (5.43) 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

$$\Pi_C^0=\int_0^{\overline{V}}VdF\left(V\right)-\left(1+r_E\right)I=\frac{1}{2}\overline{V}^2-\left(1+r_E\right)I. \tag{5.45}$$

Comparing the profits of taking out a loan in equation (5.44) and financing the investment entirely from equity in equation (5.45), we see that the company takes out a loan only if $\Pi_C^i\geq\Pi_C^0$, which simplifies to become

$$1+r_D^i\leq 1+r_D^{i*}=\frac{\overline{V}-C_i}{\overline{V}+C_i}\left(1+r_E\right), \tag{5.46}$$

which we assume to be fulfilled for both banks. To decide which bank to approach, the company would compare the profits in equation (5.44) for the two banks. Bank 2 gets approached if $\Pi_C^2\geq\Pi_C^1$, which easily becomes

$$\begin{aligned}1+r_D^2 &\leq 1+r_D^{2**} \\ &= \frac{\left(\overline{V}-C_1\right)\left(1+r_E\right)}{\left(\overline{V}-C_1\right)\left(1+r_E\right)+\left(C_1^2-C_2^2\right)\left(1+r_D^1\right)}\left(1+r_D^1\right).\end{aligned} \tag{5.47}$$

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

$$\begin{aligned}\hat{\Pi}_B^i &= \int_0^R \alpha_i V dF(V) + \int_R^{\overline{V}} R_i dF(V) \\ &\quad - \int_0^R C_i dF(V) - \left(1 + r_D^i\right) L_i \\ &= \alpha_i \frac{2\overline{V}R - R^2}{2\overline{V}} - \frac{RC_i}{\overline{V}} - \left(1 + r_D^i\right) L_i. \end{aligned} \tag{5.48}$$

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

$$\frac{\partial \hat{\Pi}_B^i}{\partial R_i} = \frac{\partial \hat{\Pi}_B^i}{\partial R} \frac{\partial R}{\partial R_i} = \alpha_i \frac{\overline{V} - R}{\overline{V}} - \frac{C_i}{\overline{V}} = 0, \tag{5.49}$$

implying

$$R = \frac{\alpha_i \overline{V} - C_i}{\alpha_i}. \tag{5.50}$$

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

$$\alpha_i = \frac{C_i}{C_1 + C_2}. \tag{5.51}$$

Inserting equations (5.50) and (5.51) into equation (5.48), we get the bank profits as

$$\hat{\Pi}_B^i = \frac{C_i}{C_1 + C_2} \frac{\left(\overline{V} - (C_1 + C_2)\right)^2}{2\overline{V}} - \left(1 + r_D^i\right) L_i. \tag{5.52}$$

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

$$L_i = \frac{C_i}{1+r_D^i}\frac{\left(\overline{V}-(C_1+C_2)\right)^2}{2\overline{V}\left(C_1+C_2\right)}. \tag{5.53}$$

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

$$\begin{aligned} \hat{\Pi}_C &= \int_R^{\overline{V}} (V-R)\,dF(V) - (1+r_E)\left(I-(L_1+L_2)\right) \\ &= \frac{(C_1+C_2)^2}{2\overline{V}} - (1+r_E)\,I \\ &\quad + \left(\frac{C_1}{(1+r_D^1)} + \frac{C_2}{(1+r_D^2)}\right)\frac{\left(\overline{V}-(C_1+C_2)\right)^2}{2\overline{V}\left(C_1+C_2\right)}(1+r_E)\,. \end{aligned} \tag{5.54}$$

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 (5.42) and (5.43) 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

$$(5.55)\quad \hat{\hat{\Pi}}_B^j = \int_{R_i}^{R} (V - R_i)\, dF(V) + \int_{R}^{\overline{V}} R_j dF(V) - \int_0^R C_j dF(V) - \left(1 + r_D^j\right) L_j = -\frac{R_j^2}{2\overline{V}} + \frac{C_i - C_j}{\overline{V}} R_j - \frac{\left(\overline{V} - C_i\right) C_j}{\overline{V}} - \left(1 + r_D^j\right) L_j,$$

when inserting from equations (5.42) and (5.43) for the results of the senior bank. The first order condition for a profit maximum is then given by

$$(5.56)\quad \frac{\partial \hat{\hat{\Pi}}_B^j}{\partial \left(1 + r_L^j\right)} = -\frac{\left(1 + r_L^j\right) L_j^2}{\overline{V}} + \frac{C_i - C_j}{\overline{V}} L_j = 0,$$

solving for

$$(5.57)\quad 1 + r_L^j = \frac{C_i - C_j}{L_j},$$

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 (5.57) we obtain from equation (5.55) that

$$(5.58)\quad L_j = \frac{C_i^2 + C_j^2 - 2\overline{V} C_j}{2\overline{V}\left(1 + r_D^j\right)}.$$

From equation (5.57) 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

$$(5.59)\quad \hat{\hat{\Pi}}_C = \int_R^{\overline{V}} \left(V - \left(1+r_L^1\right) L_1 - \left(1+r_L^2\right) L_2\right) dF\left(V\right) - \left(1+r_E\right)\left(I - \left(L_1 + L_2\right)\right)$$
$$= \frac{C_2^2}{2\overline{V}} - \left(1+r_E\right) I + \left(\frac{\left(V - C_1\right)^2}{2\overline{V}\left(1+r_D^1\right)} + \frac{C_1^2 + C_2^2 - 2\overline{V}C_2}{2\overline{V}\left(1+r_D^2\right)}\right)\left(1+r_E\right),$$

where the second equation is obtained by inserting from equations (5.42), (5.43), (5.57), and (5.58), 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 (5.44) and (5.59). 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 (5.44) and (5.59), becomes

$$(5.60)\quad 1 + r_D^2 \geq 1 + r_D^{2****} = \frac{\overline{V} + C_1}{\overline{V} - C_1}\left(1 + r_D^1\right).$$

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

$$(5.61)\quad 1 + r_D^2 \leq 1 + r_D^{2*****} = \frac{C_1^2 + C_2^2 - 2\overline{V}C_2}{C_1^2 - C_2^2}\left(1 + r_E\right).$$

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

$$(5.62)\quad \frac{\overline{V} + C_1}{\overline{V} - C_1}\left(1 + r_D^1\right) \leq 1 + r_D^2 \leq \frac{C_1^2 + C_2^2 - 2\overline{V}C_2}{C_1^2 - C_2^2}\left(1 + r_E\right).$$

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 (5.60). 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 (5.61), 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)

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.

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.

6 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 5.1.3 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 6.1 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 6.2 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 6.3 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.

6.1 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

$$\Pi_C = V - (1 + r_L)\, L. \tag{6.1}$$

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

$$\hat{\Pi}_C = p\,(V - (1 + r_L)\, L - C) + (1 - p)\, V = V - p\,((1 + r_L)\, L + C)\,. \tag{6.2}$$

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

$$p \leq p^* = \frac{(1 + r_L)\, L}{(1 + r_L)\, L + C}. \tag{6.3}$$

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

$$N_D = (1 - \pi)\, N + \kappa\pi N = (1 - \pi\,(1 - \kappa))\, N. \tag{6.4}$$

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

$$p = \frac{W}{CN_D} \tag{6.5}$$

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

$$
\begin{aligned}
(6.6)\,\Pi_B &= (\pi(1-\kappa)N + p\kappa\pi N)(1+r_L)L - p(1-\pi)NC \\
&\quad -(1+r_D)D \\
&= \frac{\pi(\kappa L - (1-\pi(1-\kappa))(1-\kappa)C)(1+r_L) - (1-\pi)C}{(1-\pi(1-\kappa))C}W \\
&\quad -((1+r_D) - \pi(1-\kappa)(1+r_L))D.
\end{aligned}
$$

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 (6.5), the number of defaults N_D with the expression in equation (6.4) 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

$$
(6.7)\quad \frac{\partial \Pi_B}{\partial W} = \frac{\pi(\kappa L - (1-\pi(1-\kappa))(1-\kappa)C)(1+r_L) - (1-\pi)C}{(1-\pi(1-\kappa))C} \lesseqgtr 0.
$$

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 (6.5) that $p = 0$ and hence as equation (6.3) 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 (6.5).

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 (6.3) 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

(6.8)

$$\kappa^{**} = -\frac{L + (1-2\pi)C}{2\pi C} + \sqrt{\left(\frac{L + (1-2\pi)C}{2\pi C}\right)^2 + \frac{1+\pi(1+r_L)}{\pi^2(1+r_L)}(1-\pi)},$$

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 (6.3) must be fulfilled with equality. Inserting for p from equation (6.5) 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

$$W = \begin{cases} W^* = 0 & \text{if} \quad \kappa = \kappa^* = 0 \\ W^{**} & \text{if} \quad \kappa = \kappa^{**} \\ W^{***} & \text{if} \quad \kappa = \kappa^{***} = 1 \end{cases} . \tag{6.9}$$

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:

$$\begin{aligned}
\Pi_C\left(W^*,\kappa^*\right) &< \Pi_C\left(W^*,\kappa^{**}\right)<\Pi_C\left(W^*,\kappa^{***}\right),\\
\Pi_C\left(W^{**},\kappa^*\right) &< \Pi_C\left(W^{**},\kappa^{**}\right)<\Pi_C\left(W^{**},\kappa^{***}\right),\\
\Pi_C\left(W^{***},\kappa^*\right) &> \Pi_C\left(W^{***},\kappa^{**}\right)>\Pi_C\left(W^{***},\kappa^{***}\right).
\end{aligned} \tag{6.10}$$

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:

$$\begin{aligned}
\Pi_B\left(W^*,\kappa^*\right) &> \Pi_B\left(W^{**},\kappa^*\right)>\Pi_B\left(W^{***},\kappa^*\right),\\
\Pi_B\left(W^*,\kappa^{***}\right) &< \Pi_B\left(W^{**},\kappa^{***}\right)<\Pi_B\left(W^{***},\kappa^{***}\right).
\end{aligned} \tag{6.11}$$

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:

$$\begin{aligned}
\Pi_C\left(0,1\right) &> \Pi_C\left(0,0\right)=\Pi_C\left(W^{***},0\right)>\Pi_C\left(W^{***},1\right),\\
\Pi_B\left(0,0\right) &> \Pi_B\left(W^{***},0\right)>\Pi_B\left(W^{***},1\right)>\Pi_B\left(0,1\right).
\end{aligned} \tag{6.12}$$

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

$$\begin{aligned}
&\lambda\Pi_C\left(0,0\right)+\left(1-\lambda\right)\Pi_C\left(W^{***},0\right)\\
&\qquad=\lambda\Pi_C\left(0,1\right)+\left(1-\lambda\right)\Pi_C\left(W^{***},1\right),\\
&\mu\Pi_B\left(0,0\right)+\left(1-\mu\right)\Pi_B\left(0,1\right)\\
&\qquad=\mu\Pi_B\left(W^{***},0\right)+\left(1-\mu\right)\Pi_B\left(W^{***},1\right).
\end{aligned} \tag{6.13}$$

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

$$\begin{aligned} \lambda &= \frac{C}{(1+r_L)\,L+C}, \\ \mu &= \frac{\pi\,(1+r_L)\,L-(1-\pi)\,C}{\pi\,(1+r_L)\,(L+C)-(1-\pi)\,C}. \end{aligned} \tag{6.14}$$

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)

6.2 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

$$\begin{aligned} \Pi_B &= \pi_H ((1+r_L) L + p_H (V_L - (1+r_D) L)) \\ &\quad + \pi_L (V_L + p_L (V_L - (1+r_D) L)) - (1+r_D) L \\ &= \pi_H (1+r_L) L + \pi_L V_L - (1+r_D) L \\ &\quad + (\pi_H p_H + \pi_L p_L)(V_L - (1+r_D) L). \end{aligned} \tag{6.15}$$

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

$$\begin{aligned} \Pi_C &= \pi_H \left(V_H - (1+r_L) L + p_H \left(\overline{V} - V_L\right)\right) \\ &\quad + \pi_L \left(V_L - V_L + p_L \left(\overline{V} - V_L\right)\right) \\ &= \pi_H (V_H - (1+r_L) L) + (\pi_H (p_H - p_L) + p_L) \left(\overline{V} - V_L\right), \end{aligned} \tag{6.16}$$

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

$$\begin{aligned} \hat{\Pi}_C &= \pi_H \left(V_H - V_L + p_L \left(\overline{V} - V_L\right)\right) \\ &\quad + \pi_L \left(V_L - V_L + p_L \left(\overline{V} - V_L\right)\right) \\ &= \pi_H (V_H - V_L) + p_L \left(\overline{V} - V_L\right), \end{aligned} \tag{6.17}$$

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 (6.16) and (6.17), this becomes

$$(p_H - p_L)\left(\overline{V} - V_L\right) \geq (1 + r_L)\, L - V_L \tag{6.18}$$

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 (6.15) after solving for $(1 + r_L)\, L$, we obtain the bank profits as

$$\begin{aligned} \Pi_B &= V_L - (1 + r_D)\, L + \pi_H p_H \left(\overline{V} - (1 + r_D)\, L\right) \\ &\quad - p_L \left(\pi_L (1 + r_D)\, L + \pi_H \overline{V} - V_L\right). \end{aligned} \tag{6.19}$$

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 (6.18) 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 (6.18), the condition that $\Pi_B \geq 0$ becomes

$$(1 + r_D)\, L \leq \frac{V_L + \pi_H \overline{V}}{1 + \pi_H} < \overline{V}, \tag{6.20}$$

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 (6.20) 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 (6.16) 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 (6.18) 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)

6.3 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

$$V_C = \Pi_C + \rho\pi V_C + (1-\pi)\,\rho^T V_C, \tag{6.21}$$

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 (6.21) such that we obtain

$$V_C = \frac{R - r_L}{1 - \pi\rho - (1-\pi)\,\rho^T} L. \tag{6.22}$$

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

$$\hat{V}_C = (1+R)\,L + \rho^T V_C. \tag{6.23}$$

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

$$1 + r_L \leq 1 + r_L^* = \rho\pi \frac{1-\rho^{T-1}}{1-\rho^T}\,(1+R)\,. \tag{6.24}$$

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 (6.22) the bank value is then given by

$$V_B = \Pi_B + \rho\pi V_B + (1-\pi)\,\rho^T V_B, \tag{6.25}$$

which using the definition of Π_B solves for

$$V_B = \frac{\pi\,(1+r_L) - (1+r_D)}{1-\pi\rho-(1-\pi)\,\rho^T} L. \tag{6.26}$$

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

$$1 + r_L^{**} = \frac{1+r_D}{\pi}. \tag{6.27}$$

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 (6.24) and inserting this equation and equalling it with the competitive loan rate from equation (6.27), we easily obtain the optimal exclusion period as

$$T^* = \frac{\ln \frac{\pi^2\rho(1+R)-(1+r_D)}{\pi^2(1+R)-(1+r_D)}}{\ln\rho}. \tag{6.28}$$

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

$$\begin{aligned} \frac{\partial T^*}{\partial \pi} &< 0, \\ \frac{\partial T^*}{\partial (1+R)} &< 0. \end{aligned} \tag{6.29}$$

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.

7 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 7.1 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 7.2 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 7.3 shows. Finally, chapter 7.4 investigates the effect competition has on credit rationing.

7.1 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 is it is successful and no return otherwise, where success is achieved with probability π. This probability as well as the return in the case f 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

$$(7.1) \qquad \begin{aligned} \Pi_C &= \int_{(1+r_L)L}^{+\infty} \pi(1+R)\, L dF(\pi(1+R)L) - E \\ &= L + \int_{(1+r_L)L}^{+\infty} \pi(1+R)\, L dF(\pi(1+R)L) - I. \end{aligned}$$

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 (7.1) 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

$$(7.2) \qquad \begin{aligned} \frac{\partial \Pi_C}{\partial L} &= 1-(1+\overline{r}_L)^2 Lf(\pi(1+\overline{r}_L)L), \\ \frac{\partial \Pi_C}{\partial(1+\overline{r}_L)} &= -(1+\overline{r}_L)L^2 f(\pi(1+\overline{r}+L)L). \end{aligned}$$

Using the implicit function theorem, we easily get that

$$(7.3) \qquad \frac{\partial(1+\overline{r}_L)}{\partial L} = -\frac{\frac{\partial \Pi_C}{\partial L}}{\frac{\partial \Pi_C}{\partial(1+\overline{r}_L)}} = \frac{1-(1+\overline{r}_L)^2 Lf(\pi(1+\overline{r}_L)L)}{(1+\overline{r}_L)L^2 f(\pi(1+\overline{r}+L)L)}.$$

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

$$(7.4) \quad \begin{aligned} \Pi_B &= \int_0^{(1+r_L)L} \pi(1+R)LdF(\pi(1+R)L) \\ &\quad + \int_{(1+r_L)L}^{(1+\overline{r})L} (1+r_L)LdF(\pi(1+R)L) - (1+r_D)L \\ &= \int_0^{(1+r_L)L} \pi(1+R)LdF(\pi(1+R)L) \\ &\quad + (F((1+\overline{r}_L)L) - F((1+r_L)L))(1+r_L)L - (1+r_D)L. \end{aligned}$$

Using the Leibniz integral rule, we easily obtain that

$$(7.5) \qquad \begin{aligned} \frac{\partial \Pi_B}{\partial(1+r_L)} &= (F((1+\overline{r}_L)L) - F((1+r_L)L))L > 0, \\ \frac{\partial \Pi_B}{\partial L} &= (F((1+\overline{r}_L)L) - F((1+r_L)L))(1+r_L) \\ &\quad + \frac{1+r_L}{1+\overline{r}_L} - (1+r_D). \end{aligned}$$

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,

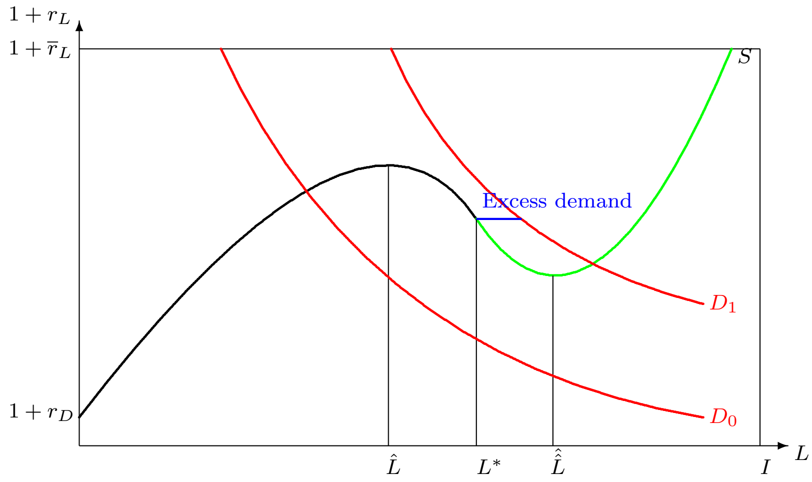


Figure 5: Credit rationing due to uncertain outcomes

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 to get

$$\frac{\partial\left(1+r_L\right)}{\partial L} = -\frac{\frac{\partial \Pi_B}{\partial L}}{\frac{\partial \Pi_B}{\partial(1+r_L)}}, \tag{7.6}$$

which is positive for $L \leq \hat{L}$ and $L > \hat{\hat{L}}$ and negative for $\hat{L} < L \leq \hat{\hat{L}}$. Figure 6 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, $\left(1+r_L\right)L$. The first order condition $\frac{\partial \Pi_B}{\partial(1+r_L)L} = 0$ solves for

$$1 + r_D = \left(F\left(\left(1+\overline{r}_L\right)L\right) - F\left(\left(1+r_L\right)L\right)\right)\left(1+r_L\right), \tag{7.7}$$

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 (7.5), 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 6, 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 6 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)

7.2 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

$$\Pi_C^i = \pi_i \left(\left(1 + R_i\right) L - \left(1 + r_L\right) L \right). \tag{7.8}$$

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

$$1 + r_L \leq 1 + \hat{r}_L = \frac{\pi_H \left(1 + R_H\right) - \pi_L \left(1 + R_L\right)}{\pi_H - \pi_L}. \tag{7.9}$$

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

$$\Pi_B^i = \pi_i \left(1 + r_L\right) L - \left(1 + r_D\right) L, \tag{7.10}$$

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

$$\Pi_B = \begin{cases} \Pi_B^L & \text{if} \quad r_L \leq \hat{r}_L \\ \Pi_B^L & \text{if} \quad r_L > \hat{r}_L \end{cases}. \tag{7.11}$$

The lower right panel in figure 6 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 (7.8) 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 6. 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.

The demand curve of an individual company is given by $\frac{\partial \Pi_C^i}{\partial L} = \pi_i \left(\left(1 + R_i\right) - \left(1 + r_L\right)\right) = 0$ 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 6 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

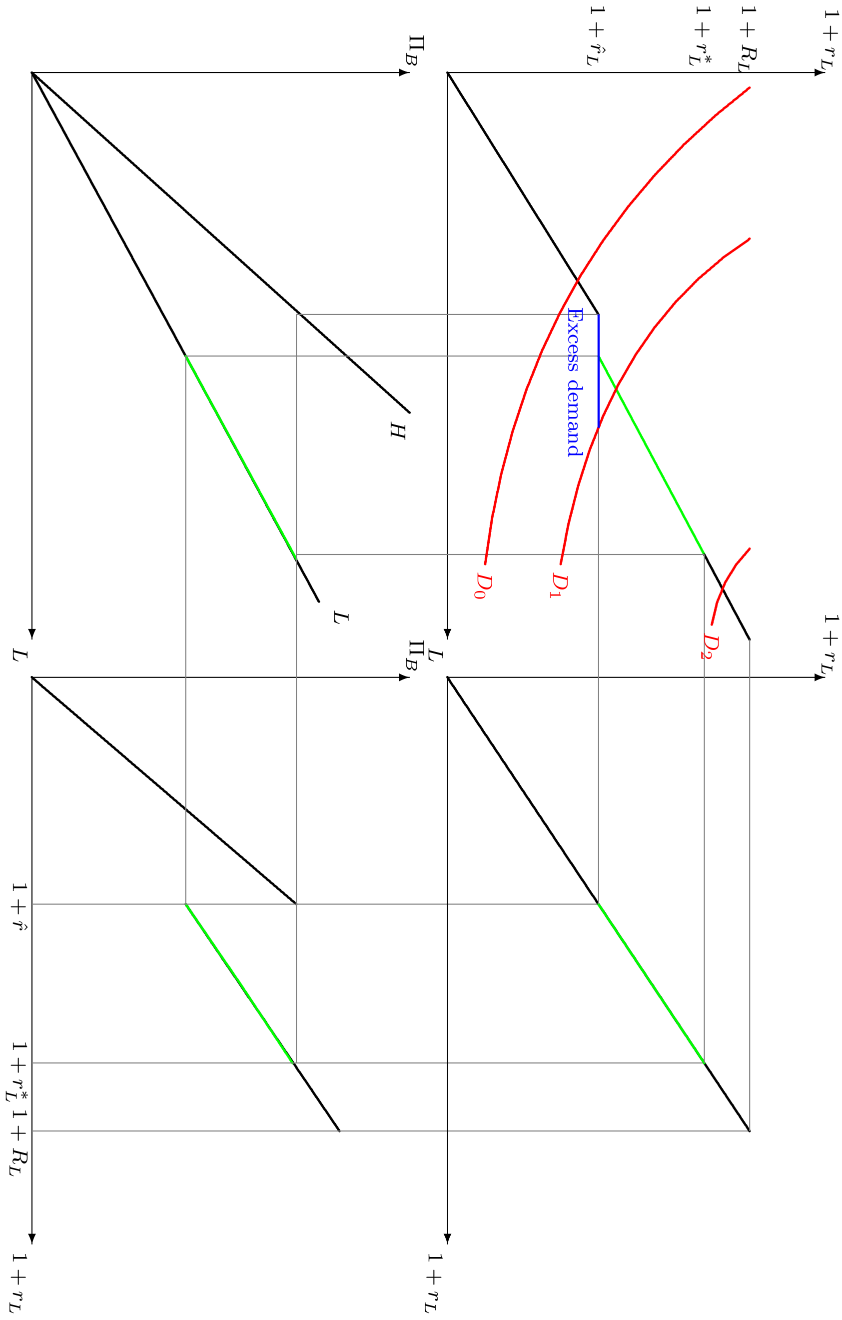


Figure 6: Credit rationing in the presence of moral hazard

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)

7.3 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

$$\Pi_C = \pi \left(V - (1 + r_L)L \right). \tag{7.12}$$

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

$$(7.13) \qquad \hat{\Pi}_C = \sum_{t=0}^{+\infty} \rho^t \Pi_C = \frac{1}{1-\rho}\Pi_C = \frac{\pi\left(V - (1+r_L)L\right)}{1-\rho},$$

with the last equality arising by inserting from equation (7.12). 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

$$(7.14) \qquad \frac{\partial V}{\partial L} = 1 + r_L.$$

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

$$(7.15) \qquad (1 + r_L)L \leq \frac{\rho\pi}{1 - \rho\,(1-\pi)}V.$$

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 (7.15) is fulfilled. With banks maximizing their profits, they would charge the highest possible loan rate such that equation (7.15) is fulfilled with equality. If the condition in equation (7.15) 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 (7.15) to ensure banks are providing loans. Thus with a Lagrangian multiplier ξ, we get as our objective function for the company

$$(7.16) \qquad \mathcal{L} = \frac{1}{1-\rho}\Pi_C + \xi\left((1 + r_L)\,L\,(1 - \rho(1-\pi)) - \rho\pi V\right).$$

The first order condition for a maximum becomes

$$(7.17) \qquad \begin{aligned} \frac{\partial \mathcal{L}}{\partial L} &= \frac{1}{1-\rho}\pi\left(\frac{\partial V}{\partial L} - (1 + r_L)\right) \\ &\quad +\xi\left((1 + r_L)\,(1 - \rho(1-\pi)) - \rho\pi\frac{\partial V}{\partial L}\right) \\ &= 0, \end{aligned}$$

in addition to equation (7.15) being met with equality. We can solve the first order condition (7.17), using equation (7.15) for ξ and obtain

$$\xi = -\frac{1}{\rho(1-\pi)} \frac{\frac{\partial V}{\partial L} - \frac{\rho\pi}{1-\rho(1-\pi)}}{\frac{V}{L} - \frac{\partial V}{\partial L}}, \tag{7.18}$$

which we insert back into equation (7.17) to obtain

$$V = \frac{1-\rho(1-\pi)}{\rho\pi}(1+r_L). \tag{7.19}$$

Using this solution for the successful outcome, we easily get

$$\frac{\partial V}{\partial L} = \frac{1-\rho(1-\pi)}{\rho\pi}(1+r_L) > 1+r_L. \tag{7.20}$$

Therefore the marginal benefits of the investment, $\frac{\partial V}{\partial L}$, will be higher than in the unconstrained optimum as given by equation (7.14). 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 (7.14) violates the constraint in equation (7.15), 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 (7.15).

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 (7.15) 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 (7.15) as

$$\begin{aligned} \Psi &= (1+r_L)L\left(1-\rho(1-\pi)\right) - \rho\pi V \\ &= \frac{\partial V}{\partial L}\left(1-\rho(1-\pi)\right) - \rho\pi V \leq 0, \end{aligned} \tag{7.21}$$

where we used the result on the optimal loan amount for companies from equation (7.14) in the final equality. We then have with $V > (1+r_L)L$ to ensure investments are profitable in case of their success that

$$\begin{aligned}(7.22)\qquad \frac{\partial \Psi}{\partial \pi} &= -\rho\left(V-(1+r_L)L\right) \leq 0 \\ \frac{\partial \Psi}{\partial V} &= \frac{\partial \Psi}{\partial L}\frac{\partial L}{\partial V} = \frac{\frac{\partial^2 V}{\partial L^2} L\left(1-\rho(1-\pi)\right) - \rho\pi\frac{\partial V}{\partial L}}{\frac{\partial V}{\partial L}}, \\ &= \frac{\frac{\partial^2 V}{\partial L^2}}{\frac{\partial V}{\partial L}} L\left(1-\rho(1-\pi)\right) - \rho\pi \leq 0.\end{aligned}$$

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 (7.15), 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)

7.4 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

$$\begin{aligned} \Pi_C &= \pi\left((1+R_H) I - (1+r_L) L\right) \\ &\quad + (1-\pi)\max\left\{(1+R_L) I - (1+r_L) L, 0\right\} - E. \end{aligned} \tag{7.23}$$

Using these profits, we can now derive the optimal loan demand by companies.

Optimal loan demand Using equation (7.23) 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

$$\frac{dr_L}{dL} = \begin{cases} \frac{\pi(R_H-R_L)+(R_L-r_L)}{L} & \text{if} \quad L \leq \frac{1+R_L}{1+r_L} I \\ \frac{R_H-r_L}{L} & \text{if} \quad L > \frac{1+R_L}{1+r_L} I \end{cases}. \tag{7.24}$$

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

$$\begin{aligned} \Pi_B &= \pi(1+r_L) L + (1-\pi)\min\left\{(1+R_L) I, (1+r_L) L\right\} \\ &\quad - (1+r_D) L. \end{aligned} \tag{7.25}$$

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

$$\frac{dr_L}{dL} = \begin{cases} -\frac{r_L-r_D}{L} & \text{if} \quad L \leq \frac{1+R_L}{1+r_L} I \\ -\frac{\pi(1+r_L)+(1-\pi)(1+R_L)-(1+r_D)}{\pi L} & \text{if} \quad L > \frac{1+R_L}{1+r_L} I \end{cases}. \tag{7.26}$$

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 (7.25) this allows us to solve for the loan rate banks will apply, which becomes

$$1 + r_L = \begin{cases} 1 + r_D & \text{if} \quad L \leq \frac{1+R_L}{1+r_L} I \\ \frac{(1+r_D)-(1-\pi)(1+R_L)\frac{L+E}{L}}{\pi} & \text{if} \quad L > \frac{1+R_L}{1+r_L} I \end{cases} . \tag{7.27}$$

Using this result in the slope of the isoprofit curve of banks from equation (7.26), 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 (7.24) 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 (7.24) and (7.26). 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.

8

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 8.1 will discuss the benefits arising from the use of collateral in terms of the cost of loans to companies, before in chapter 8.2 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 8.3.

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 8.4 we discuss the implications of such debt covenants.

8.1 The benefits of collateral

The widespread use of collateral suggests that there are inherent benefits to its use. Chapter 8.1.1 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 8.1.2 shows, collateral can be used to transfer risk from banks, who perceive them to be high, to companies, who perceive them to be low.

8.1.1 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

$$\Pi_C = \pi \left(V - (1 + r_L) L\right) - (1 - \pi) C. \tag{8.1}$$

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

$$\Pi_B = \pi (1 + r_L) L + (1 - \pi) C - (1 + r_D) L. \tag{8.2}$$

In a competitive markets banks make no profits, $\Pi_B = 0$ and the loan rate will be given by

$$1 + r_L = \frac{1 + r_D}{\pi} - \frac{1 - \pi}{\pi} \frac{C}{L}. \tag{8.3}$$

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 (8.3) into equation (8.1), the company profits are

$$\Pi_C = \pi V - (1 + r_D) L, \tag{8.4}$$

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)

8.1.2 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

$$\Pi_B = \pi_C (1 + r_L) L - (1 + r_D) L, \tag{8.5}$$

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

$$1 + r_L = \frac{1 + r_D}{\pi_C}. \tag{8.6}$$

This is the loan rate a company would expect to receive. Its profits are then given by

$$\Pi_C = \pi_C (V - (1 + r_L) L) = \pi_C V - (1 + r_D) L. \tag{8.7}$$

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 (8.6).

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

$$\hat{\Pi}_B = \pi_B (1 + \hat{r}_L) L + (1 - \pi_B) C - (1 + r_D) L. \tag{8.8}$$

With banks being competitive, the requirement that $\hat{\Pi}_B = 0$ gives us a loan rate of

$$1 + \hat{r}_L = \frac{1 + r_D}{\pi_B} - \frac{1 - \pi_B}{\pi_B} \frac{C}{L}. \tag{8.9}$$

The company profits are now reduced by the loss of the collateral if the loan is not repaid, hence

$$\begin{aligned} \hat{\Pi}_C &= \pi_C (V - (1 + \hat{r}_L) L) - (1 - \pi_C) C \\ &= \pi_C V - \frac{\pi_C}{\pi_B} (1 + r_D) L + \frac{\pi_C - \pi_B}{\pi_B} C, \end{aligned} \tag{8.10}$$

where the second equality arises from inserting equation (8.6). 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

$$C = (1 + r_D)\, D. \tag{8.11}$$

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.

8.2 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 8.2.1. 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 8.2.2 how collateral can be used to align the interest of companies the social optimum.

8.2.1 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

$$\Pi_C^i = \pi_i \left((1+R)\,L - \left(1+r_L^i\right) L \right) - (1-\pi_i)\,C_i. \tag{8.12}$$

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

$$\frac{dr_L^i}{dC_i} = -\frac{1-\pi_i}{\pi_i L}. \tag{8.13}$$

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

$$\Pi_B^i = \pi_i \left(1 + r_L^i\right) L + (1-\pi_i)\,\lambda C_i - (1 + r_D)\,L. \tag{8.14}$$

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

$$\frac{dr_L^i}{dC_i} = -\lambda\frac{1-\pi_i}{\pi_i L} \tag{8.15}$$

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 (8.13) and (8.15) 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 7 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

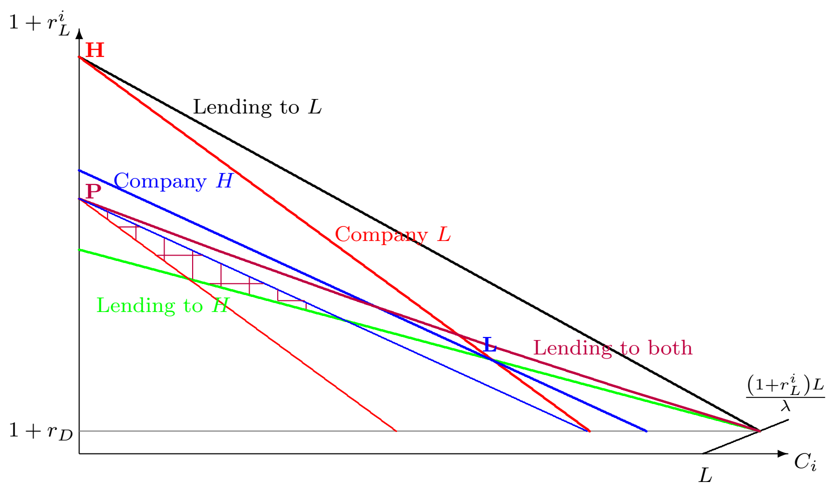


Figure 7: Separating equilibrium with collateral

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.

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 7. 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

$$\Pi_B^P = \pi\,(1+r_L)\,L + (1-\pi)\,\lambda C - (1+r_D)\,L, \tag{8.16}$$

giving us isoprofit curves with slope

$$\frac{dr_L}{dC} = -\lambda\frac{1-\pi}{\pi L}. \tag{8.17}$$

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 7 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.

		Bank 1	
		pooling	separating
Bank 2	pooling	0, 0	Π_B^*, Π_B^{**}
	separating	Π_B^{**}, Π_B^*	0, 0

Figure 8: Strategic choice of loan contracts

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 (8.16) 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

$$\Pi_B^* = \pi_H (1 + r_L) L - (1 + r_D) L = (1 - p) \frac{\pi_H - \pi_L}{\pi} (1 + r_D) L. \tag{8.18}$$

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

$$\Pi_B^{**} = \pi_L (1 + r_L) L - (1 + r_D) L = -p \frac{\pi_H - \pi_L}{\pi} (1 + r_D) L. \tag{8.19}$$

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.

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 8, 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)

8.2.2 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

$$\hat{\pi}_B - \pi_B \geq \hat{\pi}_G - \pi_G, \tag{8.20}$$

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

$$\begin{aligned} \Pi_W^G &= \pi_G V - (1 + r_D) L, \\ \hat{\Pi}_W^G &= \hat{\pi}_G V - (1 + r_D) L - E. \end{aligned} \tag{8.21}$$

It is optimal for the good company to not exert effort if $\Pi_W^G \geq \hat{\Pi}_W^G$, which easily solves for

$$\hat{\pi}_G - \pi_G \leq \frac{E}{V} \tag{8.22}$$

Similarly, for bad companies we have the welfare given by

$$\begin{aligned} \Pi_W^B &= \pi_B V - (1 + r_D) L, \\ \hat{\Pi}_W^B &= \hat{\pi}_B V - (1 + r_D) L - E \end{aligned} \tag{8.23}$$

and the bad company would exert effort if $\Pi_W^B \leq \hat{\Pi}_W^B$, from which we obtain

$$\hat{\pi}_B - \pi_B \geq \frac{E}{V}. \tag{8.24}$$

Combining equations (8.22) and (8.24) 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 (8.20). 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

$$\begin{aligned} \hat{\Pi}_C^G &= \hat{\pi}_G (V - (1 + r_L) L) - E, \\ \Pi_C^G &= \pi_G (V - (1 + r_L) L). \end{aligned} \tag{8.25}$$

If $\hat{\Pi}_C^G \leq \Pi_C^G$, the company will choose to not exert effort. This requires

$$\hat{\pi}_G - \pi_G \leq \frac{E}{V - (1 + r_L) L}. \tag{8.26}$$

Compared to the condition in the social optimum in equation (8.22), 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

$$\begin{aligned} \hat{\Pi}_C^B &= \hat{\pi}_B \left(V - (1 + r_L) L\right) - E, \\ \Pi_C^B &= \pi_B \left(V - (1 + r_L) L\right). \end{aligned} \tag{8.27}$$

If $\hat{\Pi}_C^B \geq \Pi_C^B$, the company will choose to exert effort. This requires

$$\hat{\pi}_B - \pi_B \geq \frac{E}{V - (1 + r_L) L}. \tag{8.28}$$

Compared to the condition in the social optimum in equation (8.24), 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

$$\begin{aligned} \hat{\Pi}_C^G &= \hat{\pi}_G \left(V - (1 + \hat{r}_L) L\right) - (1 - \hat{\pi}_G) C - E, \\ \Pi_C^G &= \pi_G \left(V - (1 + \hat{r}_L) L\right) - (1 - \pi_G) C. \end{aligned} \tag{8.29}$$

If $\hat{\Pi}_C^G \leq \Pi_C^G$, the company will choose to not exert effort. This requires

$$\hat{\pi}_G - \pi_G \leq \frac{E}{V - (1 + \hat{r}_L) L + C}. \tag{8.30}$$

Compared to the condition in the social optimum in equation (8.22), 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

$$\begin{aligned} \hat{\Pi}_C^B &= \hat{\pi}_B (V - (1 + \hat{r}_L) L) - (1 - \hat{\pi}_B) C - E, \\ \Pi_C^B &= \pi_B (V - (1 + \hat{r}_L) L) - (1 - \pi_B) C. \end{aligned} \tag{8.31}$$

If $\hat{\Pi}_C^B \geq \Pi_C^B$, the company will choose to exert effort. This requires

$$\hat{\pi}_B - \pi_B \geq \frac{E}{V - (1 + \hat{r}_L) L + C}. \tag{8.32}$$

Compared to the condition in the social optimum in equation (8.24), 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.

8.3 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

$$\begin{aligned}(8.33)\quad \Pi_C^H &= \pi_H\left(\left(1+R\right)L-\left(1+r_L\right)L\right)-E=\pi_H\left(R-r_L\right)L-E,\\ \Pi_C^L &= \pi_L\left(\left(1+R\right)L-\left(1+r_L\right)L\right)=\pi_L\left(R-r_L\right)L,\end{aligned}$$

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

$$(8.34)\qquad L \geq L^* = \frac{E}{\left(\pi_H-\pi_L\right)\left(R-r_L\right)}.$$

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 (8.34), 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

$$\begin{aligned}(8.35)\qquad \hat{\Pi}_C^H &= \pi_H\left(\left(1+R\right)L-\left(1+r_L\right)L\right)-\left(1-\pi_H\right)C-E\\ &= \pi_H\left(R-r_L\right)L-\left(1-\pi_H\right)C-E,\\ \hat{\Pi}_C^L &= \pi_L\left(\left(1+R\right)L-\left(1+r_L\right)L\right)-\left(1-\pi_L\right)C\\ &= \pi_L\left(R-r_L\right)L-\left(1-\pi_L\right)C.\end{aligned}$$

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

$$(8.36)\qquad L \geq L^{**} = \frac{E-\left(\pi_H-\pi_L\right)C}{\left(\pi_H-\pi_L\right)\left(R-r_L\right)}.$$

In comparison with the constraint in the absence of collateral from equation (8.34), 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

$$\begin{aligned}(8.37)\ \hat{\Pi}_B &= \hat{\pi}\left(\left(1+\hat{R}\right)\hat{L} - \left(1+\hat{r}_L\right)\hat{L} + \pi_H\left(1+r_L\right)L + \left(1-\pi_H\right)C\right) \\ &\quad - \left(1+r_D\right)L.\end{aligned}$$

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

$$(8.38)\qquad \Pi_B = \pi_H\left(1+r_L\right)L + \left(1-\pi_H\right)C - \left(1+r_D\right)L.$$

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

$$\hat{\pi} \geq \hat{\pi}^* = \frac{\pi_H \left(1 + r_L\right) L + \left(1 - \pi_H\right) C}{\left(\hat{R} - \hat{r}_L\right) \hat{L} + \pi_H \left(1 + r_L\right) L + \left(1 - \pi_H\right) C}. \tag{8.39}$$

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 (8.39) 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

$$\begin{aligned} \hat{\hat{\Pi}}_C^H &= \pi_H \left(\left(1 + R\right) L - \hat{\pi} \left(1 + r_L\right) L\right) - \left(1 - \pi_H \hat{\pi}\right) C - E, \\ \hat{\hat{\Pi}}_C^L &= \pi_L \left(\left(1 + R\right) L - \hat{\pi} \left(1 + r_L\right) L\right) - \left(1 - \pi_L \hat{\pi}\right) C. \end{aligned} \tag{8.40}$$

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

$$L \geq L^{***} = \frac{E - \hat{\pi} \left(\pi_H - \pi_L\right) C}{\left(\pi_H - \pi_L\right) \left(\left(1 + R\right) - \hat{\pi} \left(1 + r_L\right)\right)}. \tag{8.41}$$

A smaller loan is viable with rehypothecation if this constraint is less binding than the constraint without rehypthecation, $L^{***} \leq L^{**}$. Using equations (8.36) and (8.41), this solves for

$$E \geq E^* = \left(\pi_H - \pi_L\right) \left(1 + r_L\right) C. \tag{8.42}$$

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 (8.41) 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 (8.42).

In addition, we can easily show that $L^{***} \leq L^{*}$ as a comparison of equations (8.34) and (8.41) 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 (8.40), 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 (8.39), 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 8.1.1 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)

8.4 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

$$\Pi_C = \pi\left(\max\left\{(1+R)L_R + L_S; (1+r_L)L\right\} - (1+r_L)L\right). \tag{8.43}$$

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 $\Pi_C = \pi(RL_R - r_L L)$, 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

$$\Pi_B = \pi \min\{(1+R) L_R + L_S; (1+r_L) L\} + (1-\pi) L_B - (1+r_D) L. \tag{8.44}$$

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)

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.

9

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 9.1 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 9.2, 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 9.3 will explore the impact information disclosure has in this respect.

9.1 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

$$F\left(S^*\right) = \text{Prob}\left(S \leq S^*\right) = \frac{1}{\overline{S}} \int_0^{S^*} dS = \frac{S^*}{\overline{S}}. \tag{9.1}$$

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

$$(9.2)\qquad \begin{aligned} \Pi_C^2 &= \pi\left(\left(1+R\right)L-\left(1+r_L^2\right)L\right), \\ \hat{\Pi}_C^2 &= \pi\left(\left(1+R\right)L-\left(1+\hat{r}_L^2\right)L\right)-S. \end{aligned}$$

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

$$(9.3)\qquad S \leq S^* = \pi\left(\left(1+r_L^2\right)-\left(1+\hat{r}_L^2\right)\right)L.$$

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 (9.1) 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 $1-F\left(S^*\right)$ remaining with their initial bank, the profits the bank will make from their existing companies is given by

$$(9.4)\qquad \begin{aligned} \Pi_B^{2,A} &= \nu\left(\pi\left(1+r_L^2\right)L-\left(1+r_D\right)L\right)\left(1-F\left(S^*\right)\right) \\ &= \frac{\nu}{\overline{S}}\left(\pi\left(1+r_L^2\right)-\left(1+r_D\right)\right)L \\ &\quad \times\left(\overline{S}-\pi\left(\left(1+r_L^2\right)-\left(1+\hat{r}_L^2\right)\right)L\right), \end{aligned}$$

where r_D denotes the interest on deposits that finance the loan and we have used the expression for S^* from equation (9.3), together with the probability distribution in equation (9.1). Maximising these profits over the optimal loan rate to charge their existing companies, we get the first order condition that

$$\begin{aligned} \frac{\partial \Pi_B^{2,A}}{\partial\left(1+r_L^2\right)} &= \frac{\nu\pi L}{\overline{S}}\left(\overline{S}-\pi\left(\left(1+r_L^2\right)-\left(1+\hat{r}_L^2\right)\right)L\right. \\ &\quad \left.-\left(\pi\left(1+r_L^2\right)-\left(1+r_D\right)\right)L\right) \\ &= \;=0, \end{aligned}$$

which easily solves for the loan rate to become

$$(9.5)\qquad 1+r_L^2 = \frac{\pi\left(1+\hat{r}_L^2\right)+\left(1+r_D\right)-\frac{\overline{S}}{L}}{2\pi}.$$

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 (9.3). Hence we have

$$\begin{aligned}(9.6)\ \Pi_B^{2,B} &= \nu\left(\pi\left(1+\hat{r}_L^2\right)L-(1+r_D)L\right)F(S^*)-(1-\nu)(1+r_D)L\\ &= \frac{\nu\pi}{\overline{S}}\left(\pi\left(1+\hat{r}_L^2\right)-(1+r_D)\right)L\\ &\quad\times\left(\left(1+r_L^2\right)L-\left(1+\hat{r}_L^2\right)L\right)-(1-\nu)(1+r_D)L.\end{aligned}$$

Maximizing profits the bank can make from those companies that switch to them, gives rise to the first order condition

$$\begin{aligned}\frac{\partial\Pi_B^{2,B}}{\partial\left(1+\hat{r}_L^2\right)} &= \frac{\nu\pi}{\overline{S}}\left(\pi\left(1+r_L^2\right)-\pi\left(1+\hat{R}_L^2\right)\right.\\ &\quad\left.-\pi\left(1+\hat{r}_L^2\right)+(1+r_D)\right)L^2=0,\end{aligned}$$

and hence

$$1+\hat{r}_L^2=\frac{\pi\left(1+r_L^2\right)+(1+r_D)}{2\pi}. \tag{9.7}$$

Combining equations (9.5) and (9.7) we get the equilibrium loan rates as

$$\begin{aligned}1+r_L^2 &= \frac{1+r_D}{\pi}+\frac{2}{3}\frac{\overline{S}}{\pi L},\\ 1+\hat{r}_L^2 &= \frac{1+r_D}{\pi}+\frac{1}{3}\frac{\overline{S}}{\pi L}.\end{aligned} \tag{9.8}$$

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 (9.8) into equation (9.3), 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

$$\Pi_B^2=\Pi_B^{2,A}+\Pi_B^{2,B}=\frac{5}{9}\nu\overline{S}-(1-\nu)(1+r_D)L \tag{9.9}$$

as we insert the solutions for the loan rate from equations (9.8) into equations (9.4) and (9.6).

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

$$\Pi_B^1 = \nu\pi\left(1+r_L^1\right)L - \left(1+r_D\right)L. \tag{9.10}$$

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

$$\Pi_B = \nu\pi\left(1+r_L^1\right)L - \left(1+r_D\right)L + \frac{5}{9}\nu\overline{S} - (1-\nu)\left(1+r_D\right)L = 0, \tag{9.11}$$

which gives rise to a loan rate in time period 1 of

$$1 + r_L^1 = \frac{2-\nu}{\nu\pi}\left(1+r_D\right) - \frac{5}{9}\frac{\overline{S}}{\pi L}. \tag{9.12}$$

The profits to companies that do not switch are consisting of the profits in time period 1 and time period 2, giving us

$$\begin{aligned} \Pi_C &= \pi\left(\left(1+R\right)L - \left(1+r_L^1\right)L\right) \\ &\quad +\pi\left(\left(1+R\right)L - \left(1+r_L^2\right)L\right) \\ &= 2\pi\left(1+R\right)L - 2\frac{1-\nu}{\nu}\left(1+r_D\right)L - \frac{1}{9}\overline{S}, \end{aligned} \tag{9.13}$$

inserting for the loan rates from equations (9.8) and (9.12). Similarly, for those companies that do switch banks, their profits are given by

$$\begin{aligned} \hat{\Pi}_C &= \pi\left(\left(1+R\right)L - \left(1+r_L^1\right)L\right) \\ &\quad +\pi\left(\left(1+R\right)L - \left(1+\hat{r}_L^2\right)L\right) - S \\ &= 2\pi\left(1+R\right)L - 2\frac{1-\nu}{\nu}\left(1+r_D\right)L + \frac{2}{9}\overline{S} - S. \end{aligned} \tag{9.14}$$

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 $F\left(S^*\right)$, such their expected profits are given by

$$\begin{aligned} \overline{\Pi}_C &= \Pi_C\left(1-F\left(S^*\right)\right) + \hat{\Pi}_C F\left(S^*\right) \\ &= 2\pi\left(1+R\right)L - 2\frac{1-\nu}{\nu}\left(1+r_D\right)L - \frac{1}{18}\overline{S}, \end{aligned} \tag{9.15}$$

noting that the last term in equation (9.14) 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

$$\nu \geq \frac{36\,(1+r_D)}{36\,(1+r_D) + 36\pi\,(1+R) - \frac{\overline{S}}{L}}. \tag{9.16}$$

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

$$\hat{\nu}_D = \frac{\nu\,(1-\pi)}{\nu\,(1-\pi) + (1-\nu)}. \tag{9.17}$$

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

$$\begin{aligned} (9.18) \qquad \Pi_B^{2,A} &= \nu\pi \left(\pi \left(1+r_L^2\right) L - (1+r_D) L\right) (1 - F(S^*)) \\ &\quad + \nu (1-\pi) (\pi (1+R) L - (1+r_D) L) \\ &= \frac{\nu\pi}{\overline{S}} \left(\pi \left(1+r_L^2\right) - (1+r_D)\right) L \\ &\quad \times \left(\overline{S} - \pi \left(\left(1+r_L^2\right) - \left(1+\hat{r}_L^2\right)\right) L\right) \\ &\quad + \nu (1-\pi) (\pi (1+R) L - (1+r_D) L), \end{aligned}$$

where S^* is defined as in equation (9.3); 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

$$\begin{aligned} \frac{\partial \Pi_B^{2,A}}{\partial \left(1+r_L^2\right)} &= \frac{\nu\pi^2 L^2}{\overline{S}} \left(\overline{S} - \pi \left(1+r_L^2\right) L + \pi \left(1+\hat{r}_L^2\right) L \right. \\ &\quad \left. - \left(\pi \left(1+r_L^2\right) - (1+r_D)\right)\right) = 0, \end{aligned}$$

which easily solves for

$$(9.19) \qquad 1 + r_L^2 = \frac{\pi \left(1+\hat{r}_L^2\right) + (1+r_D) - \frac{\overline{S}}{L}}{2\pi}.$$

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

$$\begin{aligned}(9.20)\Pi_B^{2,B} &= \nu\pi\left(\pi\left(1+\hat{r}_L^2\right)L-(1+r_D)L\right)F(S^*)\\ &= \frac{\nu\pi^2}{\overline{S}}\left(\pi\left(1+\hat{r}_L^2\right)L-(1+r_D)L\right)\left(\left(1+r_L^2\right)-\left(1+\hat{r}_L^2\right)\right)L.\end{aligned}$$

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

$$1+\hat{r}_L^2 = \frac{\pi\left(1+r_L^2\right)+(1+r_D)}{2\pi} \tag{9.21}$$

From equations (9.19) and (9.21) we get the same loan rates as in the absence of information sharing. Thus as in equation (9.8), we have the loan rates given by

$$\begin{aligned}1+r_L^2 &= \frac{1+r_D}{\pi}+\frac{2}{3}\frac{\overline{S}}{\pi}\\ 1+\hat{r}_L^2 &= \frac{1+r_D}{\pi}+\frac{1}{3}\frac{\overline{S}}{\pi}.\end{aligned} \tag{9.22}$$

Using these loan rates, we get the profits of banks in time period 2 from both existing and switching companies, given by

$$\Pi_B^2 = \Pi_B^{2,A}+\Pi_B^{2,B} = \frac{5}{9}\nu\pi\overline{S}+\nu(1-\pi)\left(\pi(1+R)-(1+r_D)\right)L. \tag{9.23}$$

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 (9.10) 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

$$1+r_L^1 = \frac{1+\nu(1-\pi)}{\pi\nu}(1+r_D)-\frac{5}{9}\overline{s}-(1-\pi)(1+R). \tag{9.24}$$

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

$$\begin{aligned}E\left[1+r_L^2\right] &= \pi\left(1+r_L^2\right)+(1-\pi)(1+R)\\ &= (1+r_D)+\frac{2}{3}\frac{\overline{S}}{L}+(1-\pi)(1+R).\end{aligned}$$

Thus the profits of companies not switching and switching, respectively, are given by

$$
\begin{aligned}
(9.25) \qquad \Pi_C &= 2\pi(1+R)L - \pi\left(1+r_L^1\right)L - \pi E\left[1+r_L^2\right]L \\
&= 2\pi(1+R)L - \frac{1}{3}\pi\overline{S} - \frac{1+\nu}{\nu}(1+r_D)L \\
\hat{\Pi}_C &= 2\pi(1+R)L - \pi\left(1+r_L^1\right)L - \pi\left(1+\hat{r}_L^2\right)L \\
&= \pi(3-\pi)(1+R)L - \frac{2-5\pi}{9}\overline{S} \\
&\quad + \frac{1+\nu(2-\pi)}{\nu}(1+r_D)L - S.
\end{aligned}
$$

The average profits are then given as

$$
\begin{aligned}
(9.26) \qquad \overline{\Pi}_C &= \frac{2}{3}\Pi_C + \frac{1}{3}\left(\hat{\Pi}_C + S\right) - \frac{1}{18}\overline{S} \\
&= \frac{\pi(7-\pi)}{3}(1+R)L - \frac{9+2\pi}{54}\overline{S} \\
&\quad - \frac{3+\nu(4-\pi)}{3\nu}(1+r_D)L,
\end{aligned}
$$

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 (9.15), 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 (9.3) for companies not defaulting and defaulting, respectively, as

$$
\begin{aligned}
S &\leq S^{*}=\pi\left(\left(1+r_{L}^{2}\right)-\left(1+\hat{r}_{L}^{2}\right)\right)L, \\
S &\leq S^{**}=\pi\left(\left(1+r_{L}^{2,D}\right)-\left(1+\hat{r}_{L}^{2,D}\right)\right)L,
\end{aligned} \tag{9.27}
$$

we obtain the profits banks make from their own companies and those switching towards them as

$$
\begin{aligned}
\Pi_{B}^{2,A} &= \nu\pi\left(\pi\left(1+r_{L}^{2}\right)L-\left(1+r_{D}\right)L\right)\left(1-F\left(S^{*}\right)\right) \\
&\quad +\nu\left(1-\pi\right)\left(\pi\left(1+r_{L}^{2,D}\right)L-\left(1+r_{D}\right)L\right)\left(1-F\left(S^{**}\right)\right), \\
\Pi_{B}^{2,B} &= \nu\pi\left(\pi\left(1+\hat{r}_{L}^{2}\right)L-\left(1+r_{D}\right)L\right)F\left(S^{*}\right) \\
&\quad +\nu\left(1-\pi\right)\left(\pi\left(1+\hat{r}_{L}^{2,D}\right)L-\left(1+r_{D}\right)L\right)F\left(S^{**}\right).
\end{aligned} \tag{9.28}
$$

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 (9.4) and (9.6) shows, hence as in equation (9.8) we will get the loan rates in the second time period as

$$
\begin{aligned}
1+r_{L}^{2} &= 1+r_{L}^{2,D}=\frac{1+r_{D}}{\pi}+\frac{2}{3}\frac{\overline{S}}{\pi L}, \\
1+\hat{r}_{L}^{2} &= 1+\hat{r}_{L}^{2,D}=\frac{1+r_{D}}{\pi}+\frac{1}{3}\frac{\overline{S}}{\pi L}.
\end{aligned} \tag{9.29}
$$

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

$$
1+r_{L}^{1}=\frac{1+r_{D}}{\nu\pi}-\frac{5}{9}\frac{\overline{S}}{\pi L}. \tag{9.30}
$$

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

$$
(9.31) \qquad \begin{aligned} \Pi_C &= 2\pi\left(1+R\right)L - \frac{1+\nu}{\nu}\left(1+r_D\right)L - \frac{1}{9}\overline{S}, \\ \hat{\Pi}_C &= 2\pi\left(1+R\right)L - \frac{1+\nu}{\nu}\left(1+r_D\right)L + \frac{2}{9}\overline{S} - S. \end{aligned}
$$

We see that all but the second terms are identical to the profits in the case of no information disclosure in equations (9.13) and (9.14). 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)

9.2 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

$$\begin{aligned} \Pi_B^S &= \pi \min\left\{\left(1+r_L^S\right)L; V_H^S + E\right\} \\ &\quad + (1-\pi)\min\left\{\left(1+r_L^S\right)L; V_L^S + E\right\} - (1+r_D)L \\ &= \left(1+r_L^S\right)L - (1+r_D)L, \end{aligned} \tag{9.32}$$

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

$$\begin{aligned} \hat{\Pi}_B^1 &= \pi \min\left\{(1+r_L)L; (1-\phi)V_H^L + E\right\} \\ &\quad + (1-\pi)\min\left\{(1+r_L)L; (1-\phi)V_L^L + E\right\} \\ &\quad - (1+r_D)L \\ &= (1+r_L)L - (1+r_D)L, \end{aligned} \tag{9.33}$$

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

$$\begin{aligned} \hat{\Pi}_B^2 &= \pi \min\left\{(1+r_L)L; (1-\phi)V_H^L + E - (1+r_L)L\right\} \\ &\quad + (1-\pi)\min\left\{(1+r_L)L; (1-\phi)V_L^L + E - (1+r_L)L\right\} \\ &\quad - (1+r_D)L \\ &= \pi \min\left\{(1+r_L)L; (1-\phi)V_H^L + E - (1+r_L)L\right\} \\ &\quad - (1+r_D)L, \end{aligned} \tag{9.34}$$

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 (9.34) as

$$\hat{\Pi}_B^2 = \begin{cases} \pi(1+r_L)L - (1+r_D)L & \text{if} \quad \phi \leq \phi_0 \\ \pi\left((1-\phi)V_H^L + E - (1+r_L)L\right) & \text{if} \quad \phi > \phi_0 \\ -(1+r_D)L & \end{cases} . \tag{9.35}$$

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

$$1 + r_L = \begin{cases} 2\frac{1+r_D}{1+\pi} & \text{if} \quad \phi \leq \phi_0 \\ \frac{2(1+r_D)L - \pi(1-\phi)V_H^L - \pi E}{(1-\pi)L} & \text{if} \quad \phi > \phi_0 \end{cases} , \tag{9.36}$$

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

$$\begin{aligned} \Pi_C^1 &= \pi \max\left\{V_H^S + E - \left(1 + r_L^S\right)L; 0\right\} \\ &\quad + (1-\pi)\pi \max\left\{V_L^S + E - \left(1 + r_L^S\right)L; 0\right\} - E \\ &= \overline{V}_S - \left(1 + r_L^S\right)L, \end{aligned} \tag{9.37}$$

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

$$\begin{cases} \pi \geq -\frac{1}{2}\frac{V_H^S}{V_H^S - V_L^S} + \sqrt{\frac{2(1+r_D)L - V_L^S}{V_H^S - V_L^S} - \frac{1}{4}\left(\frac{V_H^S}{V_H^S - V_L^S}\right)^2} & \text{if} \quad \phi \leq \phi_0 \\ \phi \leq \frac{(1-\pi)+\pi\left(V)H^L + E\right) - 2(1+r_D)L}{V_H^L} & \text{if} \quad \phi > \phi_0 \end{cases} \tag{9.38}$$

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

$$(9.39)\quad \begin{aligned}\Pi_C^2 &= \pi\left(\phi V_H^L + \max\left\{(1-\phi)V_H^L + E - 2(1+r_L)L;0\right\}\right) \\ &\quad + (1-\pi)\left(\phi V_L^L + \max\left\{(1-\phi)V_L^L + E - 2(1+r_L)L;0\right\}\right) \\ &\quad -E \\ &= \phi\overline{V}_L + \pi\max\left\{(1-\phi)V_H^L + E - 2(1+r_L)L;0\right\} - E,\end{aligned}$$

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

$$(9.40)\quad \Pi_C^2 = \begin{cases} \phi(1-\pi)V_L^L - (1-\pi)E - 2(1+r_L)L & \text{if} \quad \phi \leq \phi_0 \\ \phi\overline{V}_L - E & \text{if} \quad \phi > \phi_0 \end{cases}.$$

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 (9.36) into the respective profits of equations (9.37) and (9.40), this requirement solves for

$$(9.41)\quad \phi \geq \phi^* = \begin{cases} \frac{(1+\pi)\overline{V}_S + \left(1-\pi^2\right)E + 2(1+r_D)L}{(1-\pi^2)V_L^L} & \text{if} \quad \phi \leq \phi_0 \\ \frac{(1-\pi)\overline{V}_S + \pi V_H^L + E - 2(1+r_D)L}{(1-\pi)V_L + \pi V_H^L} & \text{if} \quad \phi > \phi_0 \end{cases}.$$

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 (9.39) that the second term cannot be negative, easily becomes

$$(9.42) \qquad \phi \geq \phi^{**} = \begin{cases} \frac{\left(1-\pi^2\right)E+4(1+r_D)L}{(1-\pi^2)V_L^L} & \text{if} \quad \phi \leq \phi_0 \\ \frac{E}{\overline{V}_L} & \text{if} \quad \phi > \phi_0 \end{cases} .$$

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 (9.32) and (9.33) 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 (9.37), we get the profits of the company for this small investment given by

$$(9.43) \qquad \Pi_C^S = \overline{V}_S - (1 + r_D) L > 0$$

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 (9.35) 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

$$(9.44) \qquad 1 + r_L^2 = \begin{cases} \frac{1+r_D}{\pi} & \text{if} \quad \phi \leq \phi_0 \\ \frac{1+r_D}{\pi} - \frac{(1-\phi)V_H^L+E}{L} & \text{if} \quad \phi > \phi_0 \end{cases} ,$$

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 (9.39) by

$$\text{(9.45)} \qquad \Pi_C^L = \phi \overline{V}_L + \pi \max\left\{(1-\phi) V_H^L + E - \left(1+r_L^1\right) L - \left(1+r_L^2\right) L; 0\right\} - E.$$

Inserting the loan rate $r_L^1 = r_D$ and for r_L^2 from equation (9.44), we get these profits as

$$\text{(9.46)} \qquad \Pi_C^L = \begin{cases} \phi \overline{V}_L - E & \text{if} \quad \phi \leq \hat{\phi}_0 \\ \phi\left(V_L^L - \pi V_H^L\right) + 2\pi V_H^L + (2\pi - 1) E & \text{if} \quad \phi > \hat{\phi}_0 \\ -(1+\pi)(1+r_D) L & \end{cases},$$

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 (9.43) and (9.46) gives us

$$\text{(9.47)} \qquad \phi \geq \hat{\phi}^* = \begin{cases} \frac{E}{\overline{V}_L} & \text{if} \quad \phi \leq \hat{\phi}_0 \\ \frac{(1+\pi)(1+r_D)L - 2\pi V_H^L - (2\pi-1)E}{V_L^L - \pi V_H^L} & \text{if} \quad \phi > \hat{\phi}_0 \end{cases}.$$

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

$$\text{(9.48)} \qquad \phi \geq \hat{\phi}^{**} = \begin{cases} \frac{V_S + E - (1+r_D)L}{\overline{V}_L} & \text{if} \quad \phi \leq \hat{\phi}_0 \\ \frac{((1+\pi)(1+r_D)L - 2\pi V_H^L - (2\pi-1)E}{V_L^L - \pi V_H^L} & \text{if} \quad \phi > \hat{\phi}_0 \end{cases}.$$

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.

We can now compare the result with and without disclosure of the fact that a loan has been granted or applied for. In figure 9 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.

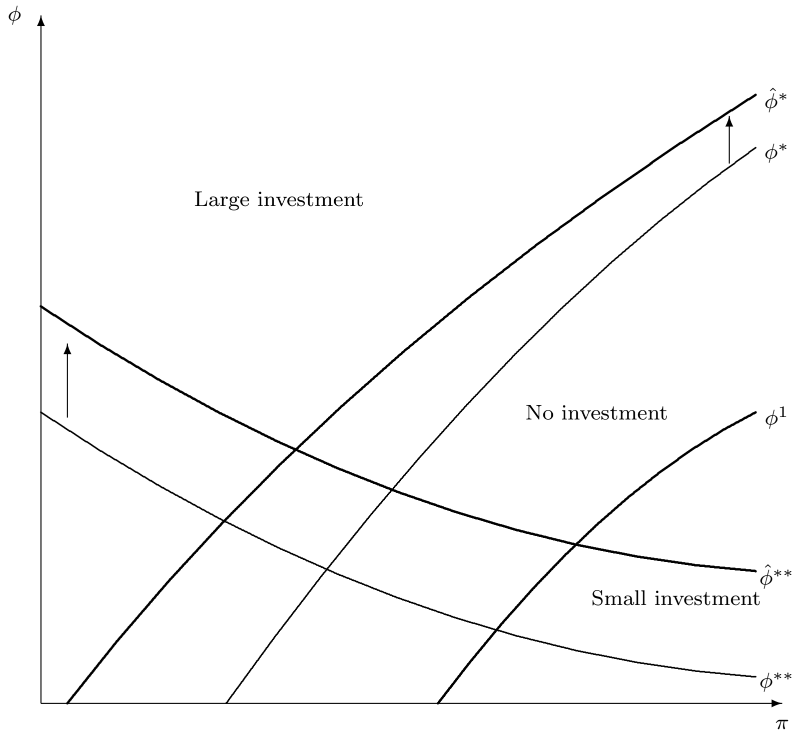


Figure 9: Investment choice with and without information disclosure

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)

9.3 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

$$\begin{aligned} \Pi_B^1 &= \nu\pi\left(1+r_L^1\right)L-\left(1+r_D\right)L, \\ \Pi_B^2 &= \nu\left(\pi\left(1+r_L^2\right)L-\left(1+r_D\right)L\right). \end{aligned} \tag{9.49}$$

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

$$\hat{\Pi}_B^2 = \hat{\nu}\pi\left(1+\hat{r}_L^2\right)L-\left(1+r_D\right)L, \tag{9.50}$$

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

$$\Pi_C^t = \pi \left((1 + R) L - \left(1 + r_L^t\right) L \right) \tag{9.51}$$

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 (9.49) and (9.50), this easily becomes

$$\begin{aligned} 1 + r_L^2 &\geq \frac{1 + r_D}{\pi}, \\ 1 + \hat{r}_L^2 &\geq \frac{1 + r_D}{\nu\pi}, \end{aligned} \tag{9.52}$$

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

$$\frac{1 + r_D}{\nu\pi} \geq 1 + r_L^2 \geq \frac{1 + r_D}{\pi}. \tag{9.53}$$

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 (9.51), 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 (9.53) and noting that the bank would charge the highest possible loan rate, we have

$$(9.54) \qquad 1 + r_L^2 = \min\left\{\frac{1 + r_D}{\nu\pi}; 1 + R\right\}.$$

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

$$(9.55) \qquad \begin{aligned} 1 + r_L^1 &= \frac{1 + \nu}{\nu\pi}\left(1 + r_D\right) - \left(1 + R\right), \\ 1 + r_L^2 &= 1 + R, \end{aligned}$$

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 $\Pi_C = 2\pi\left(1 + R\right)L - \pi\left(1 + r_L^1\right)L - \frac{1+r_D}{\nu}L$. In order for the company to accept a loan, we need to ensure that $\Pi_C \geq 0$, with solves for $1 + r_L^1 \leq 2\left(1 + R\right) - \frac{1+r_D}{\nu\pi}$.

The bank makes profits of $\Pi_B^2 = \left(1 - \nu\right)\left(1 + r_D\right)L$ and this gives us aggregate profits of $\Pi_B = \nu\pi\left(1 + r_L^1\right)L - \nu\left(1 + r_D\right)L$. 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 $2\left(1 + R\right) - \frac{1+r_D}{\nu\pi} \geq 1 + r_L^1 \geq \frac{1+r_D}{\pi}$. This provides a feasible solution if $2\left(1 + R\right) - \frac{1+r_D}{\nu\pi} \geq \frac{1+r_D}{\pi}$, 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

$$
\begin{aligned}
1+r_L^1 &= \frac{1+r_D}{\pi}, \\
1+r_L^2 &= \frac{1+r_D}{\nu\pi}
\end{aligned}
\tag{9.56}
$$

if $\pi > \frac{1}{\nu}\frac{1+r_D}{1+R}$.

Based on the results in equations (9.55) and (9.56) 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 illustrates the loan rates for different success rates of the company investments. We see that for $\pi < \frac{1}{\nu} 1 + r_D 1 + R$ 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 (9.55) and (9.56), we easily obtain that

$$
\Pi_C = \begin{cases} 2\pi\left(1+R\right)L - \frac{1+\nu}{\nu}\left(1+r_D\right) & \text{if} \quad \pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \\ 0 & \text{if} \quad \pi < \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \end{cases}.
\tag{9.57}
$$

Having established the equilibrium without the disclosure of information, we can now proceed to evaluate the situation if the bank lending in time

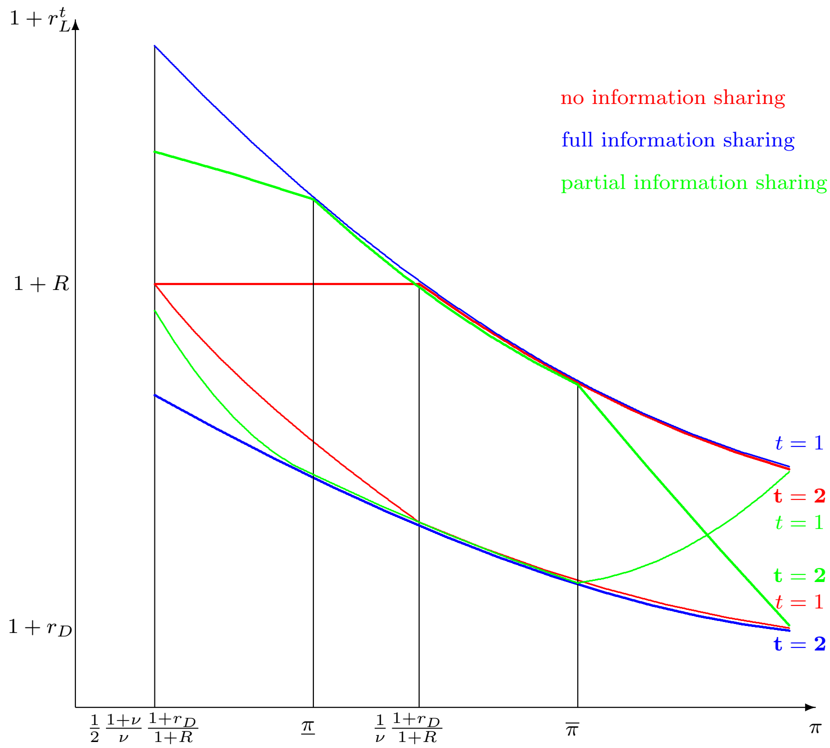


Figure 10: Equilibrium loan rates with different levels of information disclosure

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

$$1 + r_L^2 = \frac{1 + r_D}{\pi}. \tag{9.58}$$

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

$$1 + r_L^1 = \frac{1 + r_D}{\nu\pi}, \tag{9.59}$$

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 (9.58) and (9.58) becomes

$$\pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R}. \tag{9.60}$$

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 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 (9.58) and (9.59), such that we obtain

$$\Pi_C = \begin{cases} 2\pi\left(1+R\right)L - \frac{1+\nu}{\nu}\left(1+r_D\right) & \text{if} \quad \pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \\ 0 & \text{if} \quad \pi < \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \end{cases}. \tag{9.61}$$

Comparing this result with the case of banks not sharing their information about the company from equation (9.56), 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

$$\hat{\nu}_D = \frac{\nu\left(1-\pi\right)}{\nu\left(1-\pi\right)+\left(1-\nu\right)}. \tag{9.62}$$

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 (9.49) and (9.50), this easily becomes

$$
\begin{aligned}
1+r_L^2 &\geq \frac{1+r_D}{\pi}, \\
1+\hat{r}_L^2 &\geq \frac{1+r_D}{\hat{\nu}_D\pi}, \\
1+\hat{r}_L^2 &\geq \frac{1+r_D}{\hat{\nu}_N\pi},
\end{aligned}
\tag{9.63}
$$

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

$$
\frac{1+r_D}{\hat{\nu}_i\pi} \geq 1+r_L^2 \geq \frac{1+r_D}{\pi}. \tag{9.64}
$$

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 (9.51), 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 (9.64) and noting that the bank would charge the highest possible loan rate, we have

$$
1+r_L^2 = \begin{cases} \min\left\{\frac{1+r_D}{\pi}; 1+R\right\} & \text{if no default in t=1} \\ \min\left\{\frac{1+r_D}{\hat{\nu}_D\pi}; 1+R\right\} & \text{if default in t=1} \end{cases}. \tag{9.65}
$$

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 (9.62) 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 (9.65). Thus the expected loan rate is given by

$$(9.66)\qquad E\left[1+r_L^2\right] = \pi\frac{1+r_D}{\pi}+(1-\pi)\frac{1-\nu\pi}{\nu\pi(1-\pi)}(1+r_D) = \frac{1+r_D}{\nu\pi}.$$

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 $\pi\in[\underline{\pi};\overline{\pi}]$, 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 (9.65) we see that $1+r_L^2=1+R$. This now gives us an expected loan rate of

$$(9.67)\qquad E\left[1+r_L^2\right]=\pi\frac{1+r_D}{\pi}+(1-\pi)(1+R)=(1+r_D)+(1-\pi)(1+R).$$

The condition $\frac{1+r_D}{\hat{\nu}_D\pi}>1+R$ implied that this applies to $\pi\notin[\underline{\pi};\overline{\pi}]$.

The loan rate in the second time period therefore is given by

$$(9.68)\qquad E\left[1+r_L^2\right]=\begin{cases}\frac{1+r_D}{\nu\pi} & \text{if}\quad \pi\in[\underline{\pi};\overline{\pi}]\\ (1+r_D)+(1-\pi)(1+R) & \text{if}\quad \pi\notin[\underline{\pi};\overline{\pi}]\end{cases}.$$

The expected profits of the bank in time period 2 are from equation (9.68) given by $\Pi_B^2=\nu\left(\pi E\left[1+r_L^2\right]-(1+r_D)\right)L$, which after inserting from equation (9.68) becomes

$$(9.69)\qquad \Pi_B^2=\begin{cases}(1-\nu)(1+r_D)L & \text{if}\quad \pi\in[\underline{\pi};\overline{\pi}]\\ \nu(1-\pi)(\pi(1+R)-(1+r_D))L & \text{if}\quad \pi\notin[\underline{\pi};\overline{\pi}]\end{cases}.$$

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 (9.49) and (9.69) becomes

$$(9.70)\qquad \Pi_B=\begin{cases}\nu\pi\left(1+r_L^1\right)L-\nu(1+r_D)L & \text{if}\quad \pi\in[\underline{\pi};\overline{\pi}]\\ \nu\pi\left(\left(1+r_L^1\right)(1-\pi)(1+R)\right)L & \text{if}\quad \pi\notin[\underline{\pi};\overline{\pi}]\\ \quad-(1+\nu(1-\pi))(1+r_D)L & \end{cases}.$$

If banks are competitive, we require that $\Pi_B=0$, which solves for

$$(9.71)\quad 1+r_L^1=\begin{cases}\frac{1+r_D}{\pi} & \text{if}\quad \pi\in[\underline{\pi};\overline{\pi}]\\ \frac{1+\nu(1-\pi)}{\nu\pi}(1+r_D)-(1-\pi)(1+R) & \text{if}\quad \pi\notin[\underline{\pi};\overline{\pi}]\end{cases}.$$

The profits of companies are given by equation (9.51) and after inserting from equations (9.65) and (9.71), we easily obtain that $\Pi_C=2\pi(1+R)L-$

$\frac{1+\nu}{\nu}(1+r_D)L$. 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 (9.65) 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

$$\Pi_C = \begin{cases} 2\pi(1+R)L - \frac{1+\nu}{\nu}(1+r_D) & \text{if} \quad \pi \geq \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \\ 0 & \text{if} \quad \pi < \frac{1}{2}\frac{1+\nu}{\nu}\frac{1+r_D}{1+R} \end{cases}, \tag{9.72}$$

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 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.

10

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 10.1. 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 10.2, where we will also see why companies might seek to have relationships with more than just a single bank. Finally, chapter 10.3 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.

10.1 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 10.1.1 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 10.1.2 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 10.1.3 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 10.1.4. Thus banks can anticipate informational advantages and charge more consistent loan rates to companies.

10.1.1 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

$$\Pi_B^2 = \pi \left(1 + r_L^2\right) L - (1 + r_D)(L + C), \tag{10.1}$$

If banks are competitive such that their profits are eliminated, $\Pi_B^2 = 0$, the loan rate in time period 2 is given by

$$1 + r_L^2 = \frac{1 + r_D}{\pi} \frac{L + C}{L}. \tag{10.2}$$

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

$$\begin{aligned} \Pi_B^1 &= \pi \left(\left(1 + r_L^1\right) L + \left(\pi \left(1 + r_L^2\right) L - (1 + r_D) L\right)\right) \\ &\quad - (1 + r_D)(L + C). \end{aligned} \tag{10.3}$$

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 (10.2), we easily get

$$1 + r_L^1 = \frac{1 + r_D}{\pi} \frac{(1 - \pi)(L + C) + \pi L}{L} < 1 + r_L^2. \tag{10.4}$$

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 (10.2), where when staying with the same bank in time period 2 we would have $C = 0$. The total loan costs are therefore given by

$$\begin{aligned} \left(1 + \hat{r}_L^1\right) L + \left(1 + \hat{r}_L^2\right) L &= \frac{1 + r_D}{\pi} (L + C) + \frac{1 + r_D}{\pi} L \\ &= \frac{1 + r_D}{\pi} (2L + C). \end{aligned} \tag{10.5}$$

The loan costs here, however, are given as

$$\begin{aligned} \left(1 + r_L^1\right) L + \left(1 + r_L^2\right) L &= \frac{1 + r_D}{\pi} ((1 - \pi)(L + C) + \pi L) \\ &\quad + \frac{1 + r_D}{\pi} (L + C) \\ &= \frac{1 + r_D}{\pi} ((2 - \pi)(L + C) + \pi L), \end{aligned} \tag{10.6}$$

where we insert from equations (10.2) and (10.4) for the loan rates. We easily see that the expression in equation (10.6) is higher than the expression in equation (10.5) 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 (2008, Ch. 3.6.1)

10.1.2 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

$$\hat{\Pi}_B^2 = \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)\left(\hat{\pi}_1\left(1+\hat{r}_L^2\right)L - \left(1+r_D\right)L\right), \tag{10.7}$$

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

$$\begin{aligned} 1+\hat{r}_L^2 &= \frac{1+r_D}{\hat{\pi}_1} - \frac{\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+\hat{r}_L^2\right)}} \\ &= \frac{1+r_D}{\hat{\pi}_1} + \frac{\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}, \end{aligned} \tag{10.8}$$

where for the last equality we used that $\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+\hat{r}_L^2\right)} = -\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)} < 0$.

We can insert this result into equation (10.7) to obtain the profits of the new bank in time period 2 as

$$\hat{\Pi}_B^2 = \frac{\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)^2}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}} L. \tag{10.9}$$

Similarly, for the initial bank we obtain their profits in time period 2 as

$$\Pi_B^2 = \left(1 - \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)\right)\left(\pi_2\left(1+r_L^2\right)L - \left(1+r_D\right)L\right) \tag{10.10}$$

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

$$1+r_L^2 = \frac{1+r_D}{\pi_2} + \frac{1-\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}. \tag{10.11}$$

The profits of the initial bank are then given after inserting this result into equation (10.10) and we obtain

$$\Pi_B^2 = \frac{\left(1-\text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)\right)^2}{\frac{\partial \text{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}} L. \tag{10.12}$$

As from statistics we now that, as an approximation, for a random variable x we have $\text{E}\left[\frac{1}{x}\right] \approx \frac{1}{\text{E}[x]} + \frac{\text{Var}[x]}{\text{E}[x]^3}$, we easily get the expected loan rates of the new and existing bank given by

$$\text{(10.13)} \quad \begin{aligned} \mathrm{E}\left[1+\hat{r}_L^2\right] &= \frac{1+r_D}{\pi}\left(1+\frac{\sigma_1^2}{\pi^2}\right)+\frac{\mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}, \\ \mathrm{E}\left[1+r_L^2\right] &= \frac{1+r_D}{\pi}\left(1+\frac{\sigma_2^2}{\pi^2}\right)+\frac{1-\mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}. \end{aligned}$$

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 (10.13) 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 $\mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right) < \frac{1}{2}$, 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

$$\text{(10.14)} \quad \begin{aligned} \Pi_B^1 &= \left(1-\mathrm{Prob}\left(\hat{r}_L^1 < r_L^1\right)\right)\left(\pi_1\left(1+r_L^1\right)L-\left(1+r_D\right)L+\pi_1\Pi_B^2\right) \\ &\quad +\mathrm{Prob}\left(\hat{r}_L^1 < r_L^1\right)\pi_1\hat{\Pi}_B^2. \end{aligned}$$

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 (10.9) and (10.12) 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 $\mathrm{Prob}\left(\hat{r}_L^1 < r_L^1\right) = \frac{1}{2}$. This then solves for the loan rate in time period 1 to be given by

$$\text{(10.15)} \quad 1+r_L^1 = \frac{1+r_D}{\pi_1} - \frac{1-2\mathrm{Prob}\left(\hat{r}_L^2 < r_L^2\right)\left(1-\mathrm{Prob}\left(\hat{r}_L^2 < r_L^2\right)\right)}{2\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}.$$

The expected loan rate in time period 1 is then obtained as

$$\text{(10.16)} \quad \begin{aligned} \mathrm{E}\left[1+r_L^1\right] &= \left(\frac{1}{\pi}+\frac{\sigma_1^2}{\pi^3}\right)\left(1+r_D\right) \\ &\quad -\frac{1-2\mathrm{Prob}\left(\hat{r}_L^2 < r_L^2\right)\left(1-\mathrm{Prob}\left(\hat{r}_L^2 < r_L^2\right)\right)}{2\frac{\partial \mathrm{Prob}\left(\hat{r}_L^2 \leq r_L^2\right)}{\partial\left(1+r_L^2\right)}}. \end{aligned}$$

Comparing the first terms of the expected loan rates in time periods 1 and 2 from equations (10.13) and (10.16), 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 $\mathrm{E}\left[1+r_L^1\right] < \mathrm{E}\left[1+r_L^2\right]$. 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 $\mathrm{Prob}\left(\hat{r}_L^2 < r_L^2\right)$ 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)

10.1.3 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

$$\hat{\nu} = \rho p + (1 - \rho)(1 - p). \tag{10.17}$$

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

$$\begin{aligned} \hat{\pi}_H &= \frac{\rho p}{\hat{\nu}}, \\ \hat{\pi}_L &= \frac{\rho(1-p)}{1-\hat{\nu}}. \end{aligned} \tag{10.18}$$

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

$$\begin{aligned} (\hat{\pi}_H \pi_H + (1-\hat{\pi}_H)\pi_L) V &> (1+r_D) L, \\ (\hat{\pi}_L \pi_L + (1-\hat{\pi}_L)\pi_H) V &< (1+r_D) L. \end{aligned} \tag{10.19}$$

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

$$\begin{aligned} \Pi_C^i &= \pi_i (V - (1+r_L) L), \\ \hat{\Pi}_C^i &= \pi_i (V - (1+\hat{r}_L) L) - S. \end{aligned} \tag{10.20}$$

Companies will take the loan offer of their existing bank if $\Pi_C^i \geq \hat{\Pi}_C^i$, or

$$1 + r_L \leq 1 + \hat{r}_L + \frac{S}{\pi_i L}, \tag{10.21}$$

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 (10.21) will hold with equality. We thus obtain the relationship between the loan rates of the initial and any new banks as

$$1 + r_L = 1 + \hat{r}_L + \frac{S}{\pi_H L}. \tag{10.22}$$

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

$$\hat{p} = \frac{\lambda_H (1 - \rho) p}{\lambda_H (1 - \rho) p + \lambda_L \rho (1 - p)}. \tag{10.23}$$

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

$$\hat{\Pi}_B = \hat{\pi}(1 + \hat{r}_L)L - (1 + r_D) L. \tag{10.24}$$

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

$$(10.25) \qquad 1 + \hat{r}_L = \frac{1 + r_D}{\hat{\pi}}.$$

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 (10.23) that $\hat{p} = 1 - \hat{\pi}_L$. The company profits are then given by

$$(10.26) \qquad \begin{aligned} \hat{\Pi}_C^1 &= \pi_i \left(V - (1 + \hat{r}_L) L\right) - S \\ &= \pi_i \left(V - \frac{1 + r_D}{(1 - \hat{\pi}_L)\, \pi_H + \hat{\pi}_L \pi_L} L\right) - S < 0, \end{aligned}$$

where we inserted for the loan rate from equation (10.25) and for $\hat{\pi}$ from equation (10.23) with $\lambda_H = \lambda_L = 1$. Comparing this result with our assumption in equation (10.19), 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 (10.21), we obtain that the loan rate of the new bank is given by

$$(10.27) \qquad 1 + \hat{r}_L = \frac{\pi_L V - S}{\pi_L L}.$$

Comparing equations (10.25) and (10.27), we easily get

$$(10.28) \qquad \hat{p} = \pi_L \frac{(1 + r_D)\, L - \pi_L V + S}{(\pi_H - \pi_L)\, (\pi_L V - S)}.$$

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 (10.23) can be solved for the probability that companies with low success rates are switching banks:

$$(10.29) \qquad \lambda_L = \frac{1 - \hat{p}}{\hat{p}} \frac{(1 - \rho)p}{\rho(1 - p)}.$$

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 (10.28), this condition requires that

$$S < S^* = \pi_L \left(V - \frac{1+r_D}{\pi_H} L \right). \tag{10.30}$$

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 (10.28) into the definition of $\hat{\pi}$, we obtain that $\hat{\pi} = \pi_L \frac{(1+r_D)L}{\pi_L V - S}$. Using equations (10.22) and (10.27), we easily get the loan rates of the initial and new bank as

$$\begin{aligned} 1 + r_L &= \frac{V}{L} - \frac{S}{L}\left(\frac{1}{\pi_L} - \frac{1}{\pi_H}\right), \\ 1 + \hat{r}_L &= \frac{V}{L} - \frac{S}{\pi_L L}. \end{aligned} \tag{10.31}$$

As the switching costs S increase, the loan rate falls. We easily obtain that

$$\frac{\partial(1+\hat{r}_L)}{\partial S} = -\frac{1}{\pi_L L} < -\frac{1}{L}\left(\frac{1}{\pi_L} - \frac{1}{\pi_H}\right) = \frac{\partial(1+r_L)}{\partial S} < 0. \tag{10.32}$$

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

$$\begin{aligned} 1 + r_L &= \frac{(1+r_D)\,L + S}{\pi_H L}, \\ 1 + \hat{r}_L &= \frac{1+r_D}{\pi_H}. \end{aligned} \tag{10.33}$$

The loan rates of the initial banks are increasing while those of new banks remain constant,

$$0 = \frac{\partial(1+\hat{r}_L)}{\partial S} < \frac{1}{\pi_H} = \frac{\partial(1+r_L)}{\partial S}. \tag{10.34}$$

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

$$\hat{\Pi}_C^H = \pi_H \left(V - (1 + r_L) L\right) - S = \pi_H V - (1 + r_D) L - S \tag{10.35}$$

and requiring company profits to be positive, $\hat{\Pi}_C^H \geq 0$, is only achieved if

$$S \leq S^{**} = \pi_H V - (1 + r_D) L. \tag{10.36}$$

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)

10.1.4 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

$$\Pi_B^2 = \pi\left(1+\hat{r}_L^2\right)L - (1+r_D)\,L. \tag{10.37}$$

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

$$1+\hat{r}_L^2 = 1+\hat{r}_L^1 = \frac{1+r_D}{\pi}. \tag{10.38}$$

The initial bank will have learned the type of company they are lending to and hence make profits of

$$\hat{\Pi}_B^{2,i} = \pi_i\left(1+\hat{r}_L^{2,i}\right)L - (1+r_D)\,L, \tag{10.39}$$

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

$$1+\hat{r}_L^{2,i} = \frac{1+r_D}{\pi_i}, \tag{10.40}$$

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 (10.40) and use the definition of π, these profits become

$$\hat{\Pi}_B^{2,H} = p\frac{(1-p)\,(\pi_H-\pi_L)}{\pi}\,(1+r_D)\,L_H, \tag{10.41}$$

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 (10.40) and will make no profits, such that the total profits of a bank is given by equation (10.41).

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

$$\begin{aligned}(10.42)\qquad \Pi_B &= \left(\pi\left(1+r_L^1\right)-(1+r_D)\right)L \\ &\quad +\pi\left(p\left(\pi_H\left(1+r_L^{2H}\right)-(1+r_D)\right)L_H\right. \\ &\quad \left.+(1-p)\left(\pi_L\left(1+r_L^{2L}\right)-(1+r_D)\right)L_L\right).\end{aligned}$$

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

$$\begin{aligned}(10.43)\qquad \Pi_C &= \pi\left((1+R)L-\left(1+r_L^1\right)L\right) \\ &\quad +\pi p\pi_H\left((1+R)L_H-\left(1+r_L^{2,H}\right)L_H\right) \\ &\quad +\pi(1-p)\pi_L\left((1+R)L_L-\left(1+r_L^{2,L}\right)L_L\right).\end{aligned}$$

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 (10.43), 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

$$\begin{aligned} L &= -\frac{R - r_L^1}{\frac{\partial R}{\partial L}}, \\ L_H &= -\frac{R - r_L^{2,H}}{\frac{\partial R}{\partial L_H}}, \\ L_L &= -\frac{R - r_L^{2,L}}{\frac{\partial R}{\partial L_L}}, \end{aligned} \tag{10.44}$$

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

$$\mathcal{L} = \Pi_C + \zeta \left(\Pi_B - \hat{\Pi}_B^{2,H} \right), \tag{10.45}$$

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

$$\begin{aligned} 1 + r_L^1 &= \frac{1 + r_D}{\pi} + \frac{1 - \zeta}{\zeta} \frac{L}{\frac{\partial L}{\partial \left(1 + r_L^1\right)}} \\ &= 1 + \hat{r}_L^1 + \frac{1 - \zeta}{\zeta} \frac{L}{\frac{\partial L}{\partial \left(1 + r_L^1\right)}}, \\ 1 + r_L^{2,H} &= \frac{1 + r_D}{\pi_H} + \frac{1 - \zeta}{\zeta} \frac{L_H}{\frac{\partial L_H}{\partial \left(1 + r_L^{2,H}\right)}} \\ &= 1 + \hat{r}_L^{2,H} + \frac{1 - \zeta}{\zeta} \frac{L_H}{\frac{\partial L_H}{\partial \left(1 + r_L^{2,H}\right)}}, \\ 1 + r_L^{2,L} &= \frac{1 + r_D}{\pi_L} + \frac{1 - \zeta}{\zeta} \frac{L_L}{\frac{\partial L_L}{\partial \left(1 + r_L^{2,L}\right)}} \\ &= 1 + \hat{r}_L^{2,L} + \frac{1 - \zeta}{\zeta} \frac{L_L}{\frac{\partial L_L}{\partial \left(1 + r_L^{2,L}\right)}}. \end{aligned} \tag{10.46}$$

Inserting these loan rates into the constraint $\hat{\Pi}_B^{2,H} = \Pi_B$, we obtain that

$$\frac{1-\zeta}{\zeta} = \frac{p(1-p)(\pi_H - \pi_L)(1+r_D)L_H}{\pi \frac{L^2}{\frac{\partial L}{\partial\left(1+r_L^1\right)}} + \pi^2 p \frac{L_H^2}{\frac{\partial L_H}{\partial\left(1+r_L^{2,H}\right)}} + \pi^2(1-p)\frac{L_L^2}{\frac{\partial L_L}{\partial\left(1+r_L^{2,L}\right)}}} < 0, \tag{10.47}$$

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 $\frac{\partial \Pi_C}{\partial\left(1+r_L^{2,L}\right)} = -\pi(1-p)\pi_L L_L$. 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 (10.46) 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.

10.2 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 10.2.1. 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 10.2.2, 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 10.2.3 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 10.2.4 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.

10.2.1 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

$$\Pi_B^2 = p\left(1+r_L^1\right)\left(1+r_L^2\right)L - \left(1+r_L^1\right)\left(1+r_D\right)L, \tag{10.48}$$

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

$$1 + r_L^2 = \frac{1 + r_D}{p}. \tag{10.49}$$

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 (10.49). 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

$$\Pi_B^1 = \begin{cases} \left(1 + r_L^1\right) - (1 + r_D) L & \text{if} \quad p \geq p^* \\ \pi \left(1 + r_L^1\right) L - (1 + r_D) L & \text{if} \quad p < p^* \end{cases}. \tag{10.50}$$

If we assume again that banks are competitive and $\Pi_B^1 = 0$, the loan rates are then given by

$$1 + r_L^1 = \begin{cases} 1 + r_D & \text{if} \quad p \geq p^* \\ \frac{1+r_D}{\pi} & \text{if} \quad p < p^* \end{cases}. \tag{10.51}$$

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

$$\begin{aligned} \Pi_C &= \begin{cases} \pi \left(V - \left(1 + r_L^1\right) L\right) & \\ \quad + (1 - \pi) p \left(V - \left(1 + r_L^1\right) \left(1 + r_L^2\right) L\right) & \text{if} \quad p \geq p^* \\ \pi \left(V - \left(1 + r_L^1\right) L\right) & \text{if} \quad p < p^* \end{cases} \\ &= \begin{cases} (\pi + (1 - \pi) p) V & \\ \quad - \left(\pi (1 + r_D) + (1 - \pi) (1 + r_D)^2\right) L & \text{if} \quad p \geq p^* \\ \pi V - (1 + r_D) L & \text{if} \quad p < p^* \end{cases}, \end{aligned} \tag{10.52}$$

where the second equality uses the loan rates from equations (10.49) and (10.51).

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

$$(10.53) \qquad 1 + \hat{r}_L^2 = \begin{cases} \frac{1+r_D}{p} & \text{if} \quad p \geq p^* \\ \frac{V}{L} & \text{if} \quad p < p^* \end{cases} .$$

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

$$(10.54) \quad \begin{aligned} \hat{\Pi}_B = & \pi \left(\left(1 + \hat{r}_L^1\right) L - (1 + r_D) L \right) - (1 - \pi)(1 - p)(1 + r_D) L \\ & + (1 - \pi) p \left(\left(1 + \hat{r}_L^1\right) \left(1 + \hat{r}_L^2\right) L - \left(1 + \hat{r}_L^1\right)(1 + r_D) L \right) \\ & - C. \end{aligned}$$

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 (10.53). the loan rate in time period 1 as

$$(10.55) \qquad 1 + \hat{r}_L^1 = \begin{cases} \frac{(1-p(1-\pi))(1+r_D)+\frac{C}{L}}{\pi+(1-\pi)(1-p)(1+r_D)} & \text{if} \quad p \geq p^* \\ \frac{(1-p(1-\pi))(1+r_D)+\frac{C}{L}}{\pi+(1-\pi)p\frac{V}{L}-(1-\pi)p(1+r_D)} & \text{if} \quad p < p^* \end{cases} .$$

The profits of the company are now given by

$$(10.56) \quad \hat{\Pi}_C = \pi \left(V - \left(1 + \hat{r}_L^1\right) L \right) + (1 - \pi) p \left(V - \left(1 + \hat{r}_L^1\right) \left(1 + \hat{r}_L^2\right) L \right),$$

where the relationship bank will extend the loan if the company is highly skilled. Inserting the loan rates from equations (10.53) and (10.55) 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 (10.52), and relationship banking, equation (10.56) 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

$$\frac{C}{L} \geq (1-\pi)(1-p)(1+r_D)\,r_D. \tag{10.57}$$

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

$$\begin{aligned}\frac{C}{L} \quad < \quad & \frac{\left((1-\pi)\,p\frac{V}{L} + (1+r_D)\right)\left(\pi + (1-\pi)\,p\left(\frac{V}{L} - (1+r_D)\right)\right)}{\pi + (1-\pi)\,p\frac{V}{L}} \\ & -(1-p(1-\pi))(1+r_D)\,.\end{aligned} \tag{10.58}$$

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)

10.2.2 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

$$\Pi_B = \pi\,(1 + r_L)\,L - (1 + r_D)\,L, \tag{10.59}$$

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

$$1 + r_L = \frac{1 + r_D}{\pi}. \tag{10.60}$$

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

$$\begin{aligned} 1 - \hat{p} &= \text{Prob}\,(L|\text{switch}) \\ &= \frac{\text{Prob}\,(\text{switch}|L)\,\text{Prob}\,(L)}{\text{Prob}\,(\text{switch})} \\ &= \frac{1 - p}{\text{Prob}\,(\text{switch})}, \end{aligned} \tag{10.61}$$

where $\text{Prob}\,(\text{switch}|L) = 1$ as we had indicated above that any company with a low success rate would seek to switch banks. We furthermore have

$$\begin{aligned} \text{Prob}\,(\text{switch}) &= \text{Prob}\,(\text{switch}|L)\,\text{Prob}\,(L) \\ &\quad + \text{Prob}\,(\text{switch}|H)\,\text{Prob}\,(H) \\ &= (1 - p) + \text{Prob}\,(r_L > \hat{r}_L)\,p \\ &= 1 - p\,(1 - \text{Prob}\,(\hat{p} > p)) \\ &= 1 - p\text{Prob}\,(\hat{p} \leq p)\,. \end{aligned} \tag{10.62}$$

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 (10.61), we get

$$p\,(1 - Prob\,(\hat{p} \leq p)) - \hat{p}\,(1 - pProb\,(\hat{p} \leq p)) = 0. \tag{10.63}$$

As p is known, we can determine $\text{Prob}\,(\hat{p} \leq p)$ as being either 0 or 1. If $\hat{p} > p$, then $\text{Prob}\,(\hat{p} \leq p) = 0$ and equation (10.63) solves for $p = \hat{p}$, violating the assumption that $\hat{p} > p$. If $\hat{p} \leq p$ and therefore $\text{Prob}\,(\hat{p} \leq p) = 1$, equation (10.63) 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 $\hat{\Pi}_B = \pi_L\,(1 + \hat{r}_L)\,L - (1 + r_D)\,L = 0$, 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

$$\Pi_B^* = p\pi_H\left(1+r_L^*\right)L+(1-p)\,L-(1+r_D)\,L. \tag{10.64}$$

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

$$1+r_L^*=\frac{(1+r_D)-(1-p)}{p\pi_H}. \tag{10.65}$$

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)

10.2.3 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

$$\Pi_C = \pi \left((1+R) L - (1+r_L) \left(L + \hat{L} \right) \right). \tag{10.66}$$

Companies will take this loan if its expected profits are positive, $\Pi_C \geq 0$, which requires a loan rate of no more than

$$1 + r_L \leq 1 + r_L^* = (1 + R) \frac{L}{L + \hat{L}}. \tag{10.67}$$

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

$$\Pi_B = \pi (1 + r_L) \left(L + \hat{L} \right) - (1 + r_D) \left(L + \hat{L} \right), \tag{10.68}$$

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

$$1 + r_L \geq 1 + r_L^{**} = \frac{(1 + r_D) \left(L + \hat{L} \right) - \hat{L}}{\pi \left(L + \hat{L} \right)}. \tag{10.69}$$

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

$$\Pi_B = \pi (1 + R) L - (1 + r_D) \left(L + \hat{L} \right) \tag{10.70}$$

and the condition to provide a loan, $\Pi_B \geq -\hat{L}$ yields

$$\hat{L} \leq \hat{L}^* = \frac{\pi (1 + R) - (1 + r_D)}{r_D} L. \tag{10.71}$$

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

$$\hat{\Pi}_B = \pi (1 + r_L) L - (1 + r_D) L \tag{10.72}$$

and the loan is given as long as $\hat{\Pi}_B \geq 0$, or

$$1 + r_L \geq 1 + \hat{r}_L^* = \frac{1 + r_D}{\pi}. \tag{10.73}$$

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

$$\hat{\Pi}_B = \pi (1 + R) \frac{L}{L + \hat{L}} L - (1 + r_D) L \tag{10.74}$$

and the loan is given as long as $\hat{\Pi}_B \geq 0$, or

$$\hat{L} \leq \hat{L}^{**} = \frac{\pi (1 + R) - (1 + r_D)}{1 + r_D} L. \tag{10.75}$$

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)

10.2.4 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

$$\hat{\lambda} = p\lambda^N + (1-p)\,. \tag{10.76}$$

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

$$\hat{\pi} = \frac{p\pi_H\lambda^N + (1-p)\,\pi_L}{p\lambda^N + (1-p)}. \tag{10.77}$$

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 (10.76).

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

$$\hat{\Pi}_B^{2,i} = \pi_i\left(1+r_L^{2,i}\right)\left(1+r_L^1\right)\frac{L}{\hat{N}} - \left(1+r_D\right)\left(1+r_L^1\right)\frac{L}{\hat{N}} - C, \tag{10.78}$$

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

$$1+r_L^{2,i} = \frac{1+r_D}{\pi_i} + \frac{\hat{N}}{1+r_L^1}\frac{C}{L}. \tag{10.79}$$

$\hat{N}$ 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 $\binom{N-1}{i}$ permutations of banks for this scenario. The expected number banks rolling over the loan are thus given by

$$\begin{aligned}\hat{N} &= 1+\sum_{i=0}^{N-1} i\binom{N-1}{i}\lambda^{N-1-i}(1-\lambda)^i \\ &= 1+(N-1)(1-\lambda),\end{aligned} \tag{10.80}$$

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 (10.77) and thus its profits are given by

$$\hat{\Pi}_B^2 = \hat{\pi}(1+\hat{r}_L^2)\left(1+r_L^1\right)L - \left(1+r_D\right)\left(1+r_L^1\right)L. \tag{10.81}$$

If banks are competitive such that $\hat{Pi}_B^2=0$ the loan rate these banks require is given by

$$1+\hat{r}_L^2 = \frac{1+r_D}{\hat{\pi}}. \tag{10.82}$$

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 (10.79) and (10.82) for companies with high success rates, we see that they seek to remain with their bank if

$$C \leq \frac{(1+r_D)\left(1+r_L^1\right)}{\hat{N}}\left(\frac{1}{\hat{\pi}}-\frac{1}{\pi_H}\right)L. \tag{10.83}$$

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

$$\hat{\hat{\Pi}}_B = \pi\left(1+\hat{r}_L^2\right)\left(1+r_L^1\right)\frac{L}{\hat{N}} - (1+r_D)\left(1+r_L^1\right)\frac{L}{\hat{N}} < 0, \tag{10.84}$$

which after inserting from equation (10.79) for the loan rate as charged for companies with high success rates, becomes

$$C < \frac{1}{\hat{N}}\left(\frac{1}{\pi}-\frac{1}{\pi_H}\right)(1+r_D)\left(1+r_L^1\right)L. \tag{10.85}$$

As $\pi > \hat{\pi}$, the constraint in equation (10.85) is more restrictive than the constraint in equation (10.83), 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

$$\Pi_B^1 = \left(1-\lambda^N\right) p \left(1+r_L^1\right) \frac{L}{N} - \left(1+r_D\right) \frac{L}{N} \tag{10.86}$$

If banks are competitive such that $\Pi_B^1 = 0$, the loan rate in time period 1 is given by

$$1+r_L^1 = \frac{1+r_D}{p\left(1-\lambda^N\right)}. \tag{10.87}$$

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

$$\Pi_C = p\pi_H \left(1-\lambda^N\right) \left(V - \left(1+r_L^1\right)\left(1+\hat{r}_L^{2,H}\right) L\right). \tag{10.88}$$

After inserting for the loan rates $1+r_L^1$ and $1+\hat{r}_L^{2,H}$ from equations (10.79) and (10.87), this becomes

$$\Pi_C = p\pi_H \left(1-\lambda^N\right) \left(V - \hat{N}C\right) - \left(1+r_D\right)^2 L. \tag{10.89}$$

Noting the expression for $\hat{N}$ from equation (10.80), we obtain the optimal number of relationship banks to be given by the first order condition

$$\begin{aligned} \frac{\partial \Pi_C}{\partial N} &= p\pi_H \left(-\lambda^N \ln\lambda \left(V - C\left((N-1)(1-\lambda)+1\right)\right)\right. \\ &\quad \left. - \left(1-\lambda^N\right)(1-\lambda) C\right) = 0. \end{aligned} \tag{10.90}$$

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

$$
\begin{aligned}
(10.91)\quad \frac{\partial^2 \Pi_C}{\partial N^2} &= -p\pi_H \left(\lambda^N (\ln \lambda)^2 \left(V - C\hat{N}\right) - \lambda^N \ln \lambda C (1-\lambda)\right. \\
&\quad \left. -\lambda^N (1-\lambda) \ln \lambda C\right) < 0, \\
\frac{\partial^2 \Pi_C}{\partial N \partial C} &= p\pi_H \left(\lambda^N \ln \lambda \hat{N} L - \left(1-\lambda^N\right)(1-\lambda) L\right) < 0, \\
\frac{\partial^2 \Pi_C}{\partial N \partial V} &= -\lambda^N \ln \lambda > 0, \\
\frac{\partial^2 \Pi_C}{\partial N \partial \lambda} &= -\pi_H N p \lambda^{N-1} \left(C (1-\lambda)\left(\lambda^N - 1\right)\right. \\
&\quad -C\left((N-1)(1-\lambda)+1\right)+V) \ln \lambda \\
&\quad -\pi_H p \lambda^N \left(-C\left(\lambda^N - 1\right) + CN(1-\lambda)\lambda^{N-1}\right. \\
&\quad -C(1-N)) \ln \lambda \\
&\quad -\pi_H p \lambda^{N-1} \left(C(1-\lambda)\left(\lambda^N - 1\right)\right. \\
&\quad -C\left((N-1)(1-\lambda)+1\right)+V) > 0.
\end{aligned}
$$

Using the implicit function theorem, we then have the relationships between the optimal number of relationship banks and various parameters given as

$$
\begin{aligned}
(10.92)\quad \frac{\partial N}{\partial C} &= -\frac{\frac{\partial^2 \Pi_C}{\partial N \partial C}}{\frac{\partial^2 \Pi_C}{\partial N^2}} < 0, \\
\frac{\partial N}{\partial \lambda} &= -\frac{\frac{\partial^2 \Pi_C}{\partial N \partial \lambda}}{\frac{\partial^2 \Pi_C}{\partial N^2}} > 0, \\
\frac{\partial N}{\partial V} &= -\frac{\frac{\partial^2 \Pi_C}{\partial N \partial V}}{\frac{\partial^2 \Pi_C}{\partial N^2}} > 0.
\end{aligned}
$$

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.

10.3 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 10.3.1 how the presence of adverse selection affects the ability to sustain relationship banking and chapter 10.3.2 looks at the optimal investment banks should make into relationship banking as competition increases.

10.3.1 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

$$\Pi_B = \pi\,(1+r_L)\,L - (1+r_D)\,L, \tag{10.93}$$

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

$$1 + r_L = \frac{1+r_D}{\pi} + \frac{\Pi_B}{\pi L}. \tag{10.94}$$

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

$$\hat{\Pi}_B^i = \pi_i\left(1+\hat{r}_L^i\right)L - (1+r_D)\,L - C, \tag{10.95}$$

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

$$1 + \hat{r}_L^i = \frac{1+r_D}{\pi_i} + \frac{\hat{\Pi}_B^i + C}{\pi_i L}. \tag{10.96}$$

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 (10.94) 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

$$\hat{\Pi}_B^i = \left(\frac{\pi_i}{\pi}\Pi_B - C\right) + \pi_i \left(\frac{1}{\pi} - \frac{1}{\pi_i}\right)(1+r_D)\,L. \tag{10.97}$$

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 (10.97) 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

$$\hat{\hat{\Pi}}_B^i = (1-\theta)\left(\frac{\pi_i}{\pi}\Pi_B - C\right) + \pi_i \left(\frac{1}{\pi} - \frac{1}{\pi_i}\right)(1+r_D)\,L. \tag{10.98}$$

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

$$\hat{\hat{\Pi}}_B^H \geq \Pi_B = 0, \tag{10.99}$$

which we can solve for

$$\frac{\pi_L}{\pi_H} \leq \xi^* = \frac{(1+r_D)(1-p) - (1-\theta)\,p\frac{C}{L}}{(1-p)\left(1+r_D+(1-\theta)\frac{C}{L}\right)}. \tag{10.100}$$

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

$$
\begin{aligned}
(10.101)\quad \frac{\partial \xi^*}{\partial \theta} &= \frac{C}{L}\frac{1+r_D-(1-\theta)(1-p)\frac{C}{L}}{(1-p)\left(1+r_D+(1-\theta)\frac{C}{L}\right)^2} > 0, \\
\frac{\partial \xi^*}{\partial C} &= -\frac{(1-\theta)(1-p)p\left(1+r_D+(1-\theta)\frac{C}{L}\right)+(1-\theta)}{(1-p)\left(1+r_D+(1-\theta)\frac{C}{L}\right)^2} < 0, \\
\frac{\partial \xi^*}{\partial p} &= -\frac{r_D+(1-\theta)p}{(1-p)(1+r_D+(1-\theta)p)} < 0.
\end{aligned}
$$

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 (10.100) 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)

10.3.2 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

$$\Pi_B^{i,R} = \pi_i \left(1 + \hat{r}_L^i\right) L - (1 + r_D)\, L, \tag{10.102}$$

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

$$1 + \hat{r}_L^i = \frac{1 + r_D}{\pi_i}, \tag{10.103}$$

giving rise to company profits of

$$\Pi_C^{i,R} = \pi_i \left(V - \left(1 + \hat{r}_L^i\right) L\right) = \pi_i V - (1 + r_D)\, L. \tag{10.104}$$

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 (10.104) and we can interpret θ as the level of competition in the market. The actual profits that companies will obtain are thus given by

$$\hat{\Pi}_C^{i,R} = \theta \Pi_C^{i,R}. \tag{10.105}$$

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

$$1 + \hat{r}_L^i = \theta \frac{1 + r_D}{\pi_i} + (1 - \theta) \frac{V}{L}. \tag{10.106}$$

Using the bank profits as defined in equation (10.102) and inserting the loan rate from equation (10.106), we get the profits of relationship banks as

$$\hat{\Pi}_B^{i,R} = (1 - \theta) \left(\pi_i V - (1 + r_D) L\right). \tag{10.107}$$

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

$$\Pi_B^T = \pi (1 + r_L) L - (1 + r_D) L, \tag{10.108}$$

where banks charge a loan rate r_L. The profits of a company receiving such a loan is given by

$$\Pi_C^{i,T} = \pi_i \left(V - (1 + r_L) L\right). \tag{10.109}$$

If banks are competitive, $\Pi_B = 0$, we have from equation (10.108) the loan rate given by

$$1 + r_L = \frac{1 + r_D}{\pi} \tag{10.110}$$

and hence company profits are

$$\Pi_C^{i,T} = \pi_i V - \frac{\pi_i}{\pi} (1 + r_D) L. \tag{10.111}$$

If we again assume that competition will be imperfect and companies can retain only a fraction θ of their profits, we obtain

$$\hat{\Pi}_C^{i,T} = \theta \Pi_C^{i,T} \tag{10.112}$$

and hence the loan rate applied by transaction banks is given as

$$1 + r_L = \theta \frac{1 + r_D}{\pi} + (1 - \theta) \frac{V}{L}. \tag{10.113}$$

Bank profits from providing such a loan are then obtained as

$$\hat{\Pi}_B^T = (1-\theta)\left(\pi V - (1+r_D)\, L\right). \tag{10.114}$$

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

$$\Pi_B^i = \rho\left(\hat{\Pi}_B^{i,R} - C\right) + (1-\rho)\hat{\Pi}_B^T. \tag{10.115}$$

Thus the optimal investment into relationship banking is given from the first order condition

$$\frac{\partial \Pi_B^i}{\partial C} = \frac{\partial \rho}{\partial C}\left(\hat{\hat{\Pi}}_B^i - \hat{\Pi}_B\right) - \frac{\partial \rho}{\partial C} C - \rho = 0. \tag{10.116}$$

From inserting equations (10.107) and (10.114) we know that $\hat{\hat{\Pi}}_B^i - \hat{\Pi}_B = (1-\theta)\,(\pi_i - \pi)V$. Hence differentiating the expression in expression (10.116) for θ we get

$$\frac{\partial^2 \rho}{\partial C^2}\frac{\partial C}{\partial \theta}\left((1-\theta)\left(\pi_i - \pi\right) V - C\right) - \frac{\partial \rho}{\partial C}(\pi_i - \pi)V - \frac{\partial \rho}{\partial C}\frac{\partial C}{\partial \theta} = 0, \tag{10.117}$$

which solves for

$$\frac{\partial C}{\partial \theta} = \frac{\frac{\partial \rho}{\partial C}\left(\pi_i - \pi\right) V}{\frac{\partial^2 \rho}{\partial C^2}\left((1-\theta)\left(\pi_i - \pi\right) V - C\right) - \frac{\partial \rho}{\partial C}} < 0, \tag{10.118}$$

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.

11

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

$$\Pi_D = \pi\,(1+r_D)\,L + (1-\pi)\,E - C - L. \tag{11.1}$$

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

$$\Pi_B = \pi\,((1+r_L)\,L - (1+r_D)\,L + E) - E. \tag{11.2}$$

With banks being competitive, their profits will vanish and $\Pi_B = 0$, such that

$$1 + r_L = (1+r)D) + (1-\pi)\,\frac{E}{L}. \tag{11.3}$$

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

$$
\begin{aligned}
(11.4) \qquad \Pi_S &= \pi\,(1+r_S)\,L + (1-\pi)\,\theta\,(1+r_S)\,L - L \\
&= (\pi + (1-\pi)\,\theta)\,(1+r_S)\,L - L.
\end{aligned}
$$

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

$$
\begin{aligned}
(11.5) \qquad \hat{\Pi}_B &= \pi\,((1+\hat{r}_L)\,L - (1+r_S)\,L + E) \\
&\quad + (1-\pi)\,(E - \theta\,(1+r_S)\,L) - E \\
&= \pi\,(1+\hat{r}_L)\,L - (\pi + (1-\pi)\,\theta)\,(1+r_S)\,L.
\end{aligned}
$$

If banks are competitive again such that $\hat{\Pi}_B = 0$, we have

$$
(11.6) \qquad 1+\hat{r}_L = \frac{\pi + (1-\pi)\,\theta}{\pi}\,(1+r_S)\,.
$$

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 (11.3) and (11.6), this easily solves for

$$
(11.7) \qquad 1+r_S \leq \frac{\pi}{\pi + (1-\pi)\,\theta}\,(1+r_D) + \frac{1-\pi}{\pi + (1-\pi)\,\theta}\,\frac{E}{L}.
$$

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 (11.1) and (11.4) becomes

$$
(11.8) \quad 1+r_S \geq \frac{\pi}{\pi + (1-\pi)\,\theta}\,(1+r_D) + \frac{1-\pi}{\pi + (1-\pi)\,\theta}\,\frac{E}{L} - \frac{1}{\pi + (1-\pi)\,\theta}\,\frac{C}{L}.
$$

Combing these last two equations, we get

$$\begin{aligned}(11.9)\ \frac{\pi}{\pi+(1-\pi)\theta}(1+r_D)+\frac{1-\pi}{\pi+(1-\pi)\theta}\frac{E}{L} &\geq 1+r_S \\ &\geq \frac{\pi}{\pi+(1-\pi)\theta}(1+r_D)+\frac{1-\pi}{\pi+(1-\pi)\theta}\frac{E}{L}-\frac{1}{\pi+(1-\pi)\theta}\frac{C}{L}.\end{aligned}$$

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

$$1+r_S=\frac{\pi}{\pi+(1-\pi)\theta}(1+r_D)+\frac{1-\pi}{\pi+(1-\pi)\theta}\frac{E}{L}. \tag{11.10}$$

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

$$\begin{aligned}(11.11)\ \frac{\partial(1+r_S)}{\partial\theta} &= -\frac{1-\pi}{(\pi+(1-\pi)\theta)^2}\left(\pi(1+r_D)+(1-\pi)\frac{E}{L}\right)<0, \\ \frac{\partial(1+r_S)}{\partial\pi} &= \frac{\theta(1+r_D)-\frac{E}{L}}{(\pi+(1-\pi)\theta)^2}<0.\end{aligned}$$

We can now see when using equation (11.10) that $r_S \geq r_D$ if $E \geq \theta(1+r_D)L$ and the final inequality in equation (11.11) has the same requirement. As we had assumed that for the guarantee of the bank to be credible we require that $E \geq \theta(1+r_S)L$ and $r_S \geq r_D$, this condition is fulfilled.

Using the implicit function theorem we easily obtain that

$$\frac{\partial\theta}{\partial\pi}=-\frac{\frac{\partial(1+r_S)}{\partial\pi}}{\frac{\partial(1+r_S)}{\partial\theta}}<0 \tag{11.12}$$

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 (11.9) are fulfilled. With company profits given by

$$\Pi_C=\pi\left((1+R)L-(1+\hat{r}_L)L\right), \tag{11.13}$$

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 (11.6), (11.6) and (11.10), we easily get that $1 + r_L = 1 + \hat{r}_L = (1 + r_D) + \frac{1-\pi}{\pi}\frac{E}{L}$. 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 (11.9), 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 (11.9) 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 (11.12) 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 12 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 13 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 14 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 15 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 16 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 17 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.

12

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 12.1 and chapter 12.2 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 12.3 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 12.4 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.

12.1 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

$$\Pi_D = \pi (1 + r_D) D + (1 - \pi) \lambda (1 + r_D) D - (1 + r) D. \tag{12.1}$$

If we assume that depositors are competitive, then $\Pi_D = 0$ and the deposit rate is therefore given by

$$1 + r_D = \frac{1 + r}{\pi (1 - \lambda) + \lambda}. \tag{12.2}$$

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

$$r_D - r \approx (1 - \lambda)(1 - \pi). \tag{12.3}$$

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)

12.2 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

$$\text{(12.4)} \quad \begin{aligned} \Pi_D^L &= (p(\pi_L + \alpha(1-\lambda))\pi_H + (1-p)\pi_L)(1+r_L)L - D \\ &= (\pi + \alpha p(1-\lambda)\pi_H)(1+r_L)L - D. \end{aligned}$$

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

$$\text{(12.5)} \quad \begin{aligned} \Pi_D^H &= (p(\pi_L + \alpha(\pi_H - \pi_L)) + (1-p)\pi_L)(1+r_L)L - D \\ &= (\pi_L + \alpha p(\pi_H - \pi_L))(1+r_L)L - D. \end{aligned}$$

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

$$\text{(12.6)} \quad \begin{aligned} \hat{\Pi}_D &= (p\pi_H + (1-p)(1+\alpha(1-\lambda))\pi_L)(1+r_L)L - D \\ &= (\pi + \alpha(1-p)(1-\lambda)\pi_L)(1+r_L)L - D. \end{aligned}$$

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

$$\text{(12.7)} \quad \frac{\pi_H}{\pi_L} \geq \frac{1-p}{p}.$$

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

$$\frac{\pi_H}{\pi_L} \leq \frac{1 - \alpha p - (1-p)\left(1 + \alpha\left(1-\lambda\right)\right)}{p\left(1-\alpha\right)}. \tag{12.8}$$

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)

12.3 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

$$\hat{\lambda} = \lambda \frac{1 + r_L}{1 + r_D}. \tag{12.9}$$

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

$$\lambda \geq \lambda^* = \frac{1 + r_D}{1 + r_L}. \tag{12.10}$$

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

$$\Pi_D = \begin{cases} \pi\left(\left(1+r_D\right)D+B\right) & \\ \quad +\left(1-\pi\right)\left(\left(1+r_D\right)D+B\right)-D & \text{if} \quad \lambda \geq \lambda^* \\ \pi\left(\left(1+r_D\right)D+B\right) & \\ \quad +\left(1-\pi\right)\hat{\lambda}\left(1+r_D\right)D-D & \text{if} \quad \lambda < \lambda^* \end{cases}. \tag{12.11}$$

Let us now assume that deposit markets are competitive such that $\Pi_D = 0$. Inserting for $\hat{\lambda}$ from equation (12.9), this solves for

$$1+r_D = \begin{cases} 1-\frac{B}{D} & \text{if} \quad \lambda \geq \lambda^* \\ \frac{1}{\pi}-\frac{B}{D}-\frac{1-\pi}{\pi}\lambda\left(1+r_L\right) & \text{if} \quad \lambda < \lambda^* \end{cases}. \tag{12.12}$$

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

$$P = \left(1-\pi\right)\lambda\left(1+r_L\right)D. \tag{12.13}$$

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

$$\Pi_B = \begin{cases} \pi\left(\left(1+r_L\right)D-\left(1+r_D\right)D\right) & \\ \quad +\left(1-\pi\right)\left(\lambda\left(1+r_L\right)D-\left(1+r_D\right)D\right) & \\ \quad -P & \text{if} \quad \lambda \geq \lambda^* \\ \pi\left(\left(1+r_L\right)D-\left(1+r_D\right)D\right)-P & \text{if} \quad \lambda < \lambda^* \end{cases}. \tag{12.14}$$

Inserting equations (12.12) and (12.13) into equation (12.14) we obtain the profits of banks buying insurance cover as

$$\Pi_B = \begin{cases} \pi\left(1+r_L\right)D + B - D & \text{if} \quad \lambda \geq \lambda^* \\ \pi\left(1+r_L\right)D + \pi B - D & \text{if} \quad \lambda < \lambda^* \end{cases} . \tag{12.15}$$

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 (12.14), 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 (12.13) implies that $\lambda = 0$ and no coverage of deposits is available. Without any deposit insurance, we easily see from equation (12.14) that

$$\Pi_B = \pi\left(\left(1+r_L\right)D - \left(1+r_D\right)D\right). \tag{12.16}$$

We can now compare the profits of a bank purchasing full insurance, $\lambda \geq \lambda^*$, from equation (12.15) with the profits of a bank purchasing no insurance, $\lambda = 0$, from equation (12.16) and we easily see that the former is giving the bank higher profits is

$$B \geq \left(1 - \pi\left(1+r_D\right)\right)D. \tag{12.17}$$

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)

12.4 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

$$\Pi_D^i = \pi_i \left(1 + r_D^i\right) D - c d_i D, \tag{12.18}$$

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

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\pi_i \left(1 + r_D^i\right) - \pi_j \left(1 + r_D^j\right)}{2c}. \tag{12.19}$$

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

$$\Pi_B^i = \pi_i \left(1 + r_L\right) D_i - \left(1 + r_D^i\right) D_i, \tag{12.20}$$

where we assume that loans are fully financed by deposits. Inserting for $D_i = d_i^* D$ and for d_i^* from equation (12.19) we get the first order condition for the optimal deposit rate as

$$\frac{\partial \Pi_B^i}{\partial \left(1 + r_D^i\right)} = -D_i + \left(\pi_i \left(1 + r_L\right) - \left(1 + r_D^i\right)\right) \frac{\pi_i}{2c} D = 0, \tag{12.21}$$

which, after inserting for D_i and D solves for

$$1 + r_D^i = \frac{\pi_i^2 \left(1 + r_L\right) + \pi_j \left(1 + r_D^j\right) - c}{2\pi_i}. \tag{12.22}$$

We can easily obtain that

$$\begin{aligned} \frac{\partial \left(1 + r_D^i\right)}{\partial \pi_j} &= \frac{1 + r_D^j}{2\pi_i} > 0, \\ \frac{\partial \left(1 + r_D^i\right)}{\partial \pi_i} &= \frac{1 + r_L}{2} - \frac{\pi_j \left(1 + r_D\right) + c}{2\pi_i^2} \lesseqqgtr 0. \end{aligned} \tag{12.23}$$

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 (12.18), we obtain that a bank is attracting deposits only if $\pi_i \left(1 + r_D\right) \geq 1 + r$.

A bank will only accept deposits if this is profitable for them, thus we require $\Pi_B^i \geq 0$, which using equation (12.20), 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

$$\pi_i^2 \geq \frac{1 + r}{1 + r_L}. \tag{12.24}$$

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 (12.18) requires that

$$d_i \leq d_i^{**} = \frac{\pi_i \left(1 + r_D^i\right) - (1 + r)}{c}. \tag{12.25}$$

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 (12.20) and maximize for the optimal deposit rate set by banks maximizing their profits. This gives us the deposit rate as

$$1 + r_D^i = \frac{\pi_i^2 (1 + r_L) + (1 + r)}{2\pi_i} \tag{12.26}$$

and hence after inserting this result into the equation (12.25), the market share of bank i becomes

$$d_i^{**} = \frac{\pi_i^2 (1 + r_L) - (1 + r)}{2c}. \tag{12.27}$$

This maximization is only relevant if $d_i^{**} \leq 1$ and not all depositors will use banks, which implies that

$$\pi_i^2 \leq \frac{2c + (1 + r)}{1 + r_L}. \tag{12.28}$$

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

$$1 + r_D^i = \frac{c + (1 + r)}{\pi_i}. \tag{12.29}$$

We illustrate these results graphically in figure 11. If the condition in equation (12.24) 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 (12.26) and the respective market shares by equation (12.27). When inserting for the market shares from equation (12.25), the condition that $d_i^{**} + d_j^{**} < 1$ solves for

$$\pi_i^1 + \pi_j^2 < 2\frac{c + (1+r)}{1+r_L}. \tag{12.30}$$

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 (12.30) 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 (12.22) and market shares by equation (12.19). Solving equation (12.22) for $1 + r_D^i$ by inserting from $1 + r_D^j$, we easily get

$$1 + r_D^i = \frac{2}{3}\frac{\pi_i^2 + \frac{1}{2}\pi_j^2}{\pi_i}(1+r_L) - \frac{c}{\pi_i} \tag{12.31}$$

and from inserting this expression into equation (12.19), we get the market share of bank i as

$$d_i^* = \frac{\pi_i^2 + \pi_j^2}{2c}(1+r_L) - \frac{1}{2}. \tag{12.32}$$

For both banks to be accepting deposits, we require that no bank obtains the entire market, thus $d_i < 1$, or

$$\pi_i^2 + \pi_j^2 < \frac{c}{1+r_L}. \tag{12.33}$$

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 (12.20). Inserting this deposit rate into equation (12.19) 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

$$1 + r_D^i = \frac{\pi_j^2 (1 + r_L) + c}{\pi_i}. \tag{12.34}$$

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.

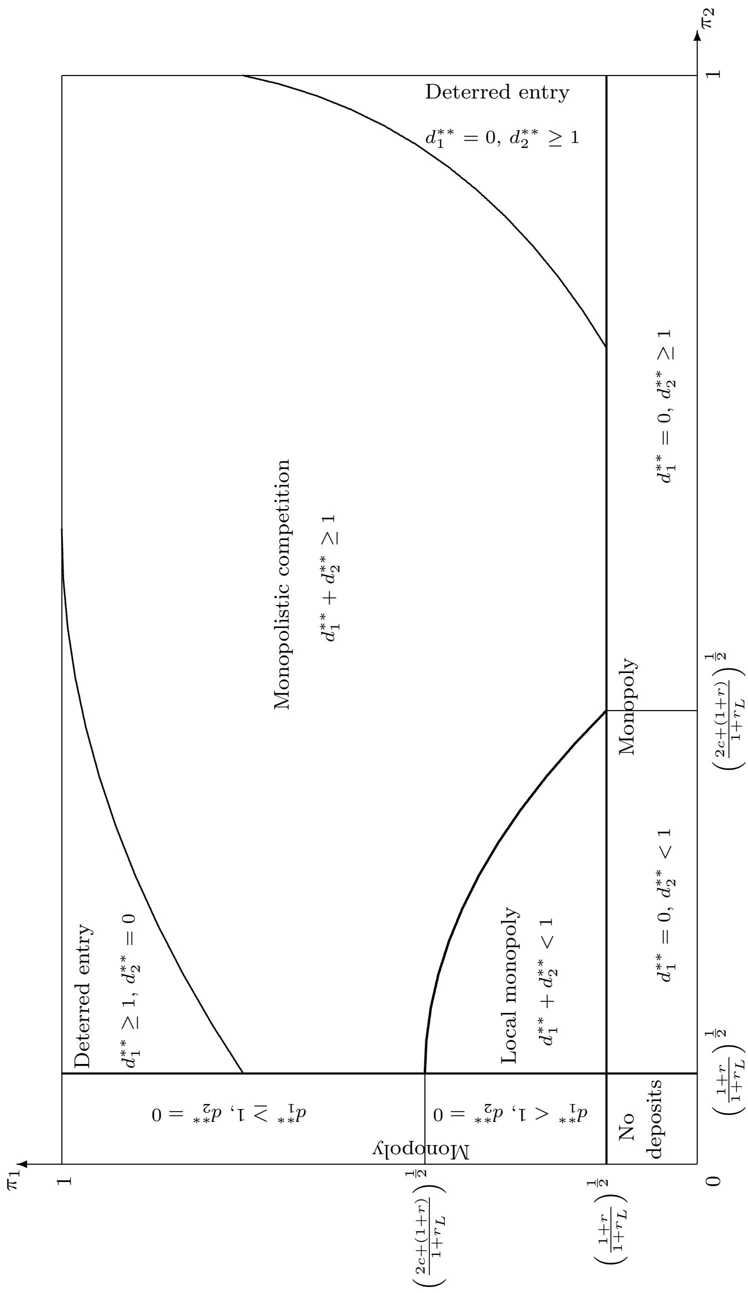


Figure 11: Equilibria characteristics in the deposit market

13

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 13.1 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 13.2. 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 13.3 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 13.4.

13.1 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 13.1.1 we see how the expectations about the behaviour of other depositors can trigger a bank run and chapter 13.1.2 shows how a lack of coordination in the behaviour of depositors can lead to a bank run that is detrimental to all depositors.

13.1.1 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

$$\begin{aligned} 1 + \hat{r}_D &= \pi\,(1 + r_L)\,\frac{L - \hat{L}}{D - \gamma L} \\ &= \pi\,(1 + r_L)\,\frac{\lambda - \gamma}{\lambda\,(1 - \gamma)}. \end{aligned} \tag{13.1}$$

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 $\pi\,(1 + r_L)\,\frac{\lambda-\gamma}{\lambda(1-\gamma)} \geq 1$, 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 $D - L + \lambda L = D - (1 - \lambda)\,L$, as all loans are sold at a discount λ and $D - L$ denotes the initial cash holding; the amount of deposits outstanding is $D - (D - L) - \gamma L = (1 - \gamma)\,L$, 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

$$1 + r_D = \frac{D - (1 - \lambda) L}{(1 - \gamma) L} > 0. \tag{13.2}$$

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

$$\gamma > \gamma^* = \lambda \frac{\pi (1 + r_L) - 1}{\pi (1 + r_L) - \lambda}. \tag{13.3}$$

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)

13.1.2 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

$$R + \lambda\pi (1 + r_L) \hat{L} = \gamma D, \tag{13.4}$$

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

$$\hat{L} = \frac{\gamma D - R}{\lambda\pi (1 + r_L)}. \tag{13.5}$$

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

$$\pi\left(1+r_L\right)\left(L-\hat{L}\right)+R-\gamma D<\left(1-\gamma\right)\left(1+r_D\right)D. \tag{13.6}$$

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 (13.6) solves for

$$\pi<\pi^{*}=\frac{\left(1+\left(1-\gamma\right)r_D\right)D-R}{\left(1+r_L\right)\left(D-R\right)}, \tag{13.7}$$

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 (13.4) that a bank would fail to repay all withdrawn deposits if $R+\lambda\pi\left(1+r_L\right)L<\gamma D$. Hence a bank would fail instantly due to their inability to repay withdrawn deposits if

$$\pi<\pi^{**}=\frac{\gamma D-R}{\lambda\left(1+r_L\right)\left(D-R\right)}. \tag{13.8}$$

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 (13.5) for the amount of loans that need to be sold into condition (13.6) for the failure of a company. A bank would fail due to not being able to repay the remaining depositors if

$$\pi<\pi^{***}=\frac{\left(\lambda\left(1+\left(1-\gamma\right)r_D\right)+\gamma\right)D-\left(1+\lambda\right)R}{\lambda\left(1+r_L\right)\left(D-R\right)}. \tag{13.9}$$

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 (13.6) to represent the situation that the resources available to the bank are not sufficient to repay the remaining depositors their initial deposits: $\pi\left(1+r_L\right)\left(L-\hat{L}\right)+R-\gamma D<\left(1-\gamma\right)D$ and obtain

$$\pi<\hat{\pi}^{***}=\frac{\left(\lambda+\gamma\right)D-\left(1+\lambda\right)R}{\lambda\left(1+r_L\right)\left(D-R\right)}. \tag{13.10}$$

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 12. 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 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

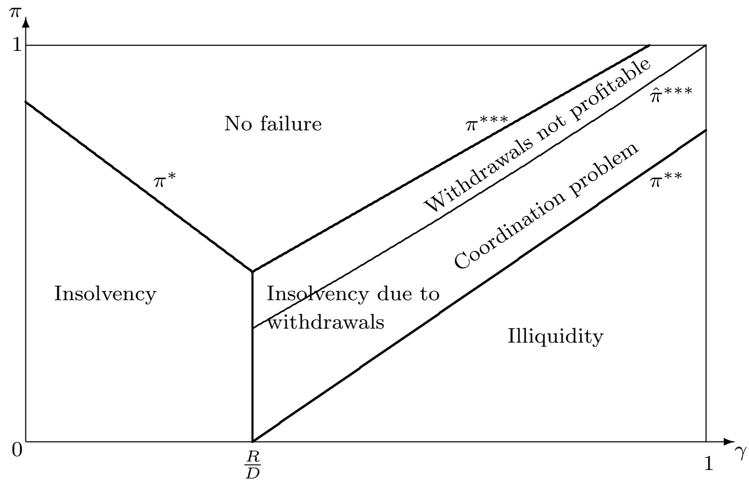


Figure 12: Bank failures due to deposit withdrawals

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)

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.

13.2 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 13.2.1 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 13.2.2 will show. Meanwhile, chapter 13.2.3 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 13.2.4 we will therefore explore how such loan guarantees affect the emergence of bank runs.

13.2.1 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

$$\begin{aligned} \theta_t^\gamma &= \frac{F_\gamma(t + \Delta t^*) - F_\gamma(t)}{1 - F_\gamma(t)} \\ &= \frac{(1 - F_\gamma(t)) - (1 - F_\gamma(t + \Delta t^*))}{1 - F_\gamma(t)} \\ &= \gamma \Delta t^*. \end{aligned} \tag{13.11}$$

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

$$\frac{d(1 - F_\gamma(t))}{1 - F_\gamma(t)} = \gamma dt,$$

which is a differential equation that can be solved for

$$F_\gamma(t) = 1 - e^{-\gamma t}. \tag{13.12}$$

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

$$F_p(t_i | t^*) = 1 - e^{-p(t_i - t^*)}. \tag{13.13}$$

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

$$
\begin{aligned}
(13.14) \qquad f_p(t_i|t^*) f_\gamma(t^*) &= \gamma p e^{-(\gamma-p)t^*} e^{-pt_i}, \\
\int_{t_i-\Delta t^*}^{t_i} f_p(t_i|s) f_\gamma(s)\, ds &= \gamma p e^{-pt_i} \int_{t_i-\Delta t^*}^{t_i} e^{-(\gamma-p)s} ds \\
&= \frac{\gamma p e^{-pt_i}}{\gamma - p} e^{-(\gamma-p)t_i} \left(e^{(\gamma-p)\Delta t^*} - 1 \right).
\end{aligned}
$$

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

$$
\begin{aligned}
(13.15) \qquad f(t^*|t_i) &= \frac{f_p(t_i|t^*) f_\gamma(t^*)}{\int_{t_i-\Delta t^*}^{t_i} f_p(t_i|s) f_\gamma(s)\, ds} \\
&= \frac{\gamma - p}{e^{(\gamma-p)\Delta t^*} - 1} e^{(\gamma-p)(t_i-t^*)},
\end{aligned}
$$

where the second equality uses the results of equation (13.14). The probability that the liquidity happened before time t^*, given the information was received at t_i, is then given by

$$
\begin{aligned}
(13.16) \qquad F(t^*|t_i) &= \int_{t_i-\Delta t^*}^{t_i} f(s|t_i)\, ds \\
&= \frac{(\gamma-p) e^{(\gamma-p)t_i}}{e^{(\gamma-p)\Delta t^*} - 1} \int_{t_i-\Delta t^*}^{t_i} e^{-(\gamma-p)sds} \\
&= \frac{e^{(\gamma-p)\Delta t^*} - e^{(\gamma-p)(t_i-t^*)}}{e^{(\gamma-p)\Delta t^*} - 1}.
\end{aligned}
$$

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

$$h_i(\tau) = \frac{f(t_i + \tau - T|t_i)}{1 - F(t_i + \tau - T|t_i)} = \frac{(\gamma - p)\, e^{(\gamma-p)(T-\tau)}}{e^{(\gamma-p)(T-\tau)} - 1}, \tag{13.17}$$

using equations (13.15) and (13.16) 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

$$\frac{\partial h_i(\tau)}{\partial \tau} = \frac{(\gamma - p)^2 e^{(\gamma-p)(T-\tau)}}{\left(e^{(\gamma-p)(T-\tau)} - 1\right)^2} > 0. \tag{13.18}$$

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

$$h_i(\tau)(\lambda D - V(\tau)) + \frac{\partial V(\tau)}{\partial \tau} = 0. \tag{13.19}$$

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 (13.19), we obtain after inserting for the probability of a bank failing from equation (13.17), that the value of deposits is given by

$$V(\tau) = \lambda D \tag{13.20}$$

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

$$d\Pi_D^i(\tau) = r_D V(\tau)\,dt - h_i(\tau)(1-\lambda)\,Ddt, \tag{13.21}$$

where $V(\tau)$ and $h(\tau)$ are given by equations (13.20) and (13.17), 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 (13.17) becomes

$$\tau \leq \tau^* = T - \frac{1}{\gamma - p}\ln\frac{\frac{\lambda}{1-\lambda}r_D}{\frac{\lambda}{1-\lambda}r_D - (\gamma - p)}, \tag{13.22}$$

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 (13.13), we know that with each deposit D we have at the time the bank fails due to running out of cash reserves

$$\begin{aligned} D\int_{t^*}^{t^*+(T-\tau)} f_p(s|t^*)\,ds &= D\int_{t^*}^{t^*+(T-\tau)} pe^{-p(s-t^*)}ds \\ &= \left(1-e^{-p(T-\tau)}\right)D \\ &\leq R \end{aligned} \tag{13.23}$$

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

$$\tau \geq \tau^{**} = T + \frac{\ln\left(1-\frac{R}{D}\right)}{p}. \tag{13.24}$$

If we now combine the two conditions in equations (13.22) and (13.24), we get that depositors do not withdraw and banks do not fail if $\tau^{**} \leq \tau \leq$

τ^*, which has a viable solution if

$$\frac{R}{D} \geq 1 - \left(\frac{\frac{\lambda}{1-\lambda} r_D}{\frac{\lambda}{1-\lambda} r_D - (\gamma - p)} \right)^{\frac{p}{\gamma - p}}. \tag{13.25}$$

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 (13.26) 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 (13.24) 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

$$\lambda > \lambda^* = \frac{(\gamma - p)\, e^{(\gamma - p)T}}{(\gamma - p)\, e^{(\gamma - p)T} + r_D \left(e^{(\gamma - p)T} - 1 \right)}. \tag{13.26}$$

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)

13.2.2 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

$$D_i = \frac{D - p_i(1+r_D)D}{1-p_i}\pi_i(1+r_L). \tag{13.27}$$

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

$$\text{(13.28)} \qquad \Pi_D^i = \int_{-\infty}^{+\infty} \frac{D - p_i (1 + r_D) D}{1 - p_i} \pi_i (1 + r_L) \, dF(\varepsilon) - (1 + r_D) D.$$

If $\pi_i = 0$, we see immediately that the first term in equation (13.28) 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 (13.28) 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

$$\text{(13.29)} \qquad \frac{\partial \lambda_i}{\partial \pi} = \frac{\partial \lambda_i}{\partial p_i} \frac{\partial p_i}{\partial \Pi_D^i} \frac{\partial \Pi_D^i}{\partial \pi} < 0.$$

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)

13.2.3 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

$$\pi_1 \geq \pi_1^* = \frac{1 + r_D}{1 + r_L}. \tag{13.30}$$

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

$$\lambda_1 = \begin{cases} \pi_1 \frac{1+r_L}{1+r_D} & \text{if} \quad \pi_1 < \pi_1^* \\ 1 & \text{if} \quad \pi_1 \geq \pi_1^* \end{cases}. \tag{13.31}$$

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

$$\pi_2 \geq \pi_2^* = \left(\frac{1 + r_D}{1 + r_L} \right)^2 = \pi_1^{*\,2}. \tag{13.32}$$

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

$$\lambda_2 = \begin{cases} \pi_2 \left(\frac{1+r_L}{1+r_D}\right)^2 & \text{if} \quad \pi_2 < \pi_2^* \\ 1 & \text{if} \quad \pi_2 \geq \pi_2^* \end{cases} . \tag{13.33}$$

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

$$\pi_2 - \overline{\pi} = \theta (\pi_1 - \overline{\pi}) + \varepsilon_2. \tag{13.34}$$

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 (13.34) into equation (13.33) that

$$\begin{aligned} \mathrm{E}[\lambda_2|\pi_1] &= \int_{-\infty}^{\varepsilon_2^*} (\theta\pi_1 + (1-\theta)\overline{\pi} + \varepsilon_2) \left(\frac{1+r_L}{1+r_D}\right)^2 dF(\varepsilon_2) \\ &\quad + \int_{\varepsilon_2^*}^{+\infty} dF(\varepsilon_2) \\ &= 1 - F(\varepsilon_2^*) + \frac{\theta\pi_1 + (1-\theta)\overline{\pi}}{\pi_2^*} F(\varepsilon_2^*) \\ &\quad + \frac{1}{\pi_2^*} \int_{-\infty}^{\varepsilon_2^*} \varepsilon_2 dF(\varepsilon_2). \end{aligned} \tag{13.35}$$

From this we obtain using the Leibniz integral rule that

$$\frac{\partial \mathrm{E}[\lambda_2|\pi_1]}{\partial \pi_1} = \frac{\theta}{\pi_2^*} F(\varepsilon_2^*). \tag{13.36}$$

If depositors withdraw after time period 1, they will obtain a payment of $\lambda_1 (1 + r_D) D$ and if they retain their deposits, they expect to obtain after time period 2 the amount of $\mathrm{E}[\lambda_2|\pi_1](1 + r_D)^2 D$. Depositors will withdraw if $\lambda_1 (1 + r_D) D \geq \mathrm{E}[\lambda_2|\pi_1](1 + r_D)^2 D$, which solves for

$$1 + r_D \leq 1 + r_D^* = \frac{\lambda_1}{\mathrm{E}[\lambda_2|\pi_1]}. \quad (13.37)$$

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 (13.31) to obtain the fraction of deposits repaid, λ_1. If $\pi_1 \geq \pi_1^*$, we know that $\lambda_1 = 1$ and using equation (13.36), we see that the critical deposit rate for a withdrawal is decreasing in the repayment rate of the bank as we have

$$\frac{\partial(1 + r_D^*)}{\partial \pi_1} = -\frac{1}{\mathrm{E}[\lambda_2|\pi_1]^2} \frac{\theta}{\pi_2^*} F(\varepsilon_2^*) < 0. \quad (13.38)$$

In the case that $\pi_1 < \pi_1^*$, this becomes

$$\begin{aligned} \frac{\partial(1 + r_D^*)}{\partial \pi_1} &= \frac{1}{\pi_1^*} \frac{1 - F(\varepsilon_2^*) + (1 - \theta)\frac{\overline{\pi}}{\pi_2^*} F(\varepsilon_2^*) + \frac{1}{\pi_2^*}\int_{-\infty}^{\varepsilon_2^*} \varepsilon_2 dF(\varepsilon_2)}{\mathrm{E}[\lambda_2|\pi_1]^2} \\ &> 0. \end{aligned} \quad (13.39)$$

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

$$\begin{aligned} \text{(13.40)} \qquad \mathrm{E}[\lambda_1] &= \int_{-\infty}^{\varepsilon_1^*} \pi_1 \frac{1+r_L}{1+r_D} dF(\varepsilon_1) + \int_{\varepsilon_1^*}^{+\infty} dF(\varepsilon_1) \\ &= 1 - F(\varepsilon_1^*) + \frac{\overline{\pi}}{\pi_1^*} F(\varepsilon_1^*) + \frac{1}{\pi_1^*} \int_{-\infty}^{\varepsilon_1^*} \varepsilon_1 dF(\varepsilon_1), \end{aligned}$$

where we made use the definition of π_1^* in equation (13.30). Using from the definition of ε_1^* that $\frac{\partial \varepsilon_1^*}{\partial \overline{\pi}} = -1$, we get using the Leibniz integration rule that

$$\text{(13.41)} \qquad \frac{\partial \mathrm{E}[\lambda_1]}{\partial \overline{\pi}} = \frac{F(\varepsilon_1^*)}{\pi_1^*}.$$

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 (13.35). It is thus

$$\begin{aligned} \text{(13.42)} \qquad \mathrm{E}[\lambda_2] &= \int_{-\infty}^{+\infty} \mathrm{E}[\lambda_2|\pi_1] \, dF(\varepsilon_1) \\ &= 1 - F(\varepsilon_2^*) + \frac{\overline{\pi}}{\pi_2^*} F(\varepsilon_2^*) + \frac{1}{\pi_2^*} \int_{-\infty}^{\varepsilon_2^*} \varepsilon_2 dF(\varepsilon_2), \end{aligned}$$

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

$$\text{(13.43)} \qquad \frac{\partial \mathrm{E}[\lambda_2]}{\partial \overline{\pi}} = \frac{F(\varepsilon_2^*)}{\pi_2^*}.$$

Depositors will withdraw if their expected payments from withdrawing exceed the expected payments from remaining with the bank, $\mathrm{E}[\lambda_1](1+r_D)D \geq \mathrm{E}[\lambda_2](1+r_D)^2 D$, or

$$\text{(13.44)} \qquad 1 + r_D \leq 1 + r_D^{**} = \frac{\mathrm{E}[\lambda_1]}{\mathrm{E}[\lambda_2]}.$$

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 (13.40) and (13.43), as well as nothing that $\pi_2^* = \pi_1^{*\ 2}$, we obtain

$$\frac{\partial\left(1+r_D^{**}\right)}{\partial \overline{\pi}} = \frac{\pi_1^* F\left(\varepsilon_1^*\right) - F\left(\varepsilon_2^*\right)}{\pi_2^*} + \frac{\pi_1^* - 1}{\pi_2^*} F\left(\varepsilon_1^*\right) F\left(\varepsilon_2^*\right) + \frac{F\left(\varepsilon_1^*\right)\int_{-\infty}^{\varepsilon_2^*}\varepsilon_2 dF\left(\varepsilon_2\right) - F\left(\varepsilon_2^*\right)\int_{-\infty}^{\varepsilon_1^*}\varepsilon_1 dF\left(\varepsilon_1\right)}{\pi_1^*\pi_2^*} < 0. \tag{13.45}$$

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 13 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 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

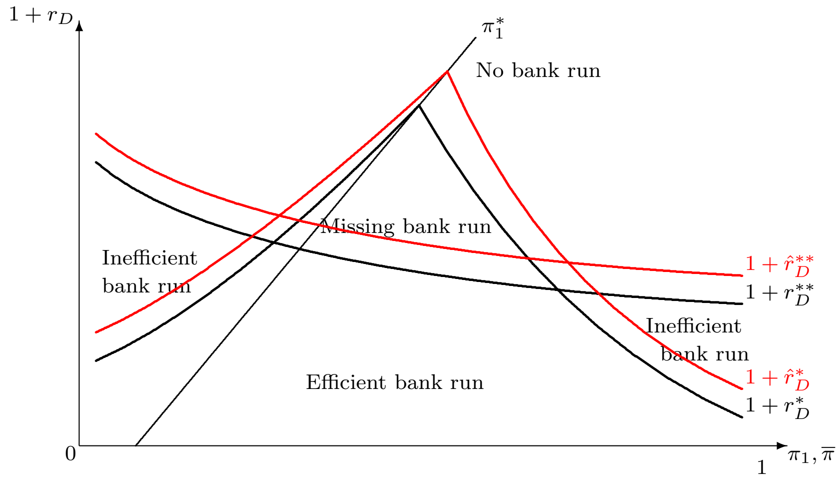


Figure 13: Efficient and inefficient bank runs

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 13 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)

13.2.4 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

$$\begin{aligned} \Pi_B &= p\left(\pi_H + \left(1-\pi_H\right)\lambda\right)\left(1+r_L\right)L \\ &\quad + \left(1-p\right)\left(\alpha\pi_H + \left(1-\alpha\pi_H\right)\lambda\right)\left(1+r_L\right)L \\ &\quad - \left(1+r_D\right)L - C. \end{aligned} \tag{13.46}$$

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

$$\pi_H^* = \frac{p + \left(1-p\right)\alpha}{c}\left(1+r_L\right)\left(1-\lambda\right). \tag{13.47}$$

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 $C = \frac{1}{2}\frac{(p+(1-p)\alpha)^2}{c}\left(1+r_L\right)^2\left(1-\lambda\right)^2 L$ from inserting for π_H from (13.46).

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 $\left(1+r_D\right)L$ 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 $\left(\pi_H + \left(1-\pi_H\right)\lambda\right)\left(1+r_L\right)L -$

$C \geq (1 + r_D) L$, which, when inserting for π_H from equation (13.47) and for C from the expression above, solves for

$$(13.48)\quad (p + (1-p)\alpha)\left(1 - \frac{1}{2}(p + (1-p)\alpha)\right) \geq c\frac{(1 + r_D) - \lambda(1 + r_L)}{(1 + r_L)^2 (1-\lambda)^2}.$$

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

$$(13.49)\quad (p + (1-p)\alpha)\left(\alpha - \frac{1}{2}(p + (1-p)\alpha)\right) \geq c\frac{(1 + r_D) - \lambda(1 + r_L)}{(1 + r_L)^2 (1-\lambda)^2}.$$

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 (13.48) and (13.49) 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

$$(13.50)\qquad \lambda > \lambda^* = 2\frac{1 + r_D}{1 + r_L} - 1.$$

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 (13.48) and (13.49) 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 (13.48) and (13.49) 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

$$(13.51)\qquad \lambda > \lambda^{**} = 2 - \frac{1 + r_L}{1 + r_D}.$$

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.

13.3 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

$$\Pi_D = \gamma u\left(\left(1+r_D^1\right)D\right) + (1-\gamma)\, u\left(\left(1+r_D^2\right)D\right). \tag{13.52}$$

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

$$\gamma\left(1+r_D^1\right)D \leq R. \tag{13.53}$$

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, $\pi\,(1+r_L)\,L$. Hence we require furthermore that

$$(1-\gamma)\left(1+r_D^2\right)D \leq R-\gamma\left(1+r_D^1\right)D + \pi\,(1+r_L)\,L. \tag{13.54}$$

Denoting by ξ_1, and ξ_2 the Lagrange coefficients for the constraints in equations (13.53) and (13.54), respectively, the first order conditions for maximizing the expected utility of depositors, equation (13.52), are given by

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\left(1+r_D^1\right)D} &= \gamma\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} - \xi_1\gamma - \xi_2\gamma = 0, \\ \frac{\partial\mathcal{L}}{\partial\left(1+r_D^2\right)D} &= (1-\gamma)\,\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D} - \xi_2\,(1-\gamma) = 0, \end{aligned} \tag{13.55}$$

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

$$\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} - \xi_1 = \frac{\partial\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}. \tag{13.56}$$

As $\xi_2 > 0$, the constraint in equation (13.54) is binding. If $\xi_1 > 0$, then the condition in equation (13.53) is also binding, implying from solving these conditions as equalities that

$$\begin{aligned} (13.57) \qquad \left(1+r_D^1\right)D &= \frac{R}{\gamma}, \\ \left(1+r_D^2\right)D &= \frac{\pi\left(1+r_L\right)L}{1-\gamma}, \end{aligned}$$

after inserting the solution for $\left(1+r_D^1\right)D$ into equation (13.54) to obtain the second expression.

If $\xi_1 = 0$, then equation (13.53) 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 (13.54), we get the expression for $\left(1+r_D^1\right)D$ as

$$(13.58) \qquad \left(1+r_D^1\right)D = \left(1+r_D^2\right)D = R + \pi\left(1+r_L\right)L.$$

The constraint on the cash reserves held by the bank, $\gamma\left(1+r_D^1\right)D \leq R$, becomes binding if

$$(13.59) \qquad \pi > \pi^* = \frac{1-\gamma}{\gamma}\frac{R}{\left(1+r_L\right)L}.$$

If the constraint is binding, we have both constraints, equations (13.53) and (13.54), being fulfilled with equality and the repayments to depositors are given by equation (13.57), and if it is not binding the deposit rates are given by equation (13.58).

As we can see from the depiction of our result in figure 14, 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 (13.53) and the total resources available to repay depositors in time period 2, equation(13.54), become

$$\begin{aligned} (\gamma + (1-\gamma)\hat{\gamma})\left(1+\hat{r}_D^1\right) D &\leq R \\ (1-\gamma)(1-\hat{\gamma})\left(1+\hat{r}_D^2\right) D &\leq \left(R - \left(1+\hat{r}_D^1\right) D\right) \\ &\quad +\pi(1+r_L) L, \end{aligned} \tag{13.60}$$

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 (13.55) that the second constraint is binding. Let us furthermore assume that banks offer a deposit contract that promises to pay depositors $(1+\hat{r}_D^*) D$ 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 $\left(1+\hat{r}_D^1\right) D = \min\left\{(1+\hat{r}_D^*) D; \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}\right\}$, where the second expression has been obtained from the first constraint in equation (13.60) being fulfilled with equality due to it being binding. Solving the binding second constraint in equation ((13.60)), the repayment for depositors remaining with the bank are given by

$$\left(1+\hat{r}_D^2\right) D = \frac{(R - (\gamma + (1-\gamma)\hat{\gamma})(1+\hat{r}_D^*) D) + \pi(1+r_L) L}{(1-\gamma)(1-\hat{\gamma})} \tag{13.61}$$

if $(1+\hat{r}_D^*) D \leq \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}$ 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 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, $(1+\hat{r}_D^*) D > \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}$, the deposits rate for those remaining with the bank is given by

$$\left(1+\hat{r}_D^2\right) D = \frac{\pi(1+r_L) L}{(1-\gamma)(1-\hat{\gamma})} \tag{13.62}$$

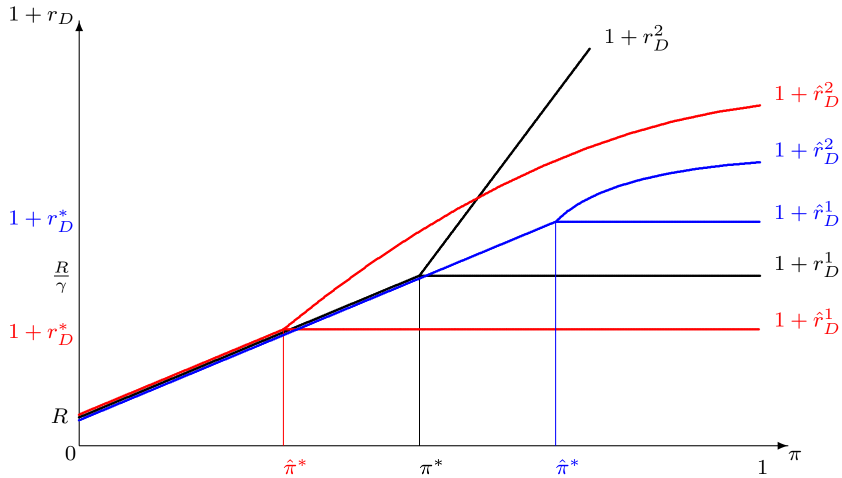


Figure 14: Deposit rates with early withdrawals

If we have a shortage of cash reserves, $(1+\hat{r}_D^*)\,D \geq \frac{R}{\gamma+(1-\gamma)\hat{\gamma}}$, 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 (13.60) and using the expression in equation (13.61), we obtain the equilibrium early withdrawal rate as

$$\hat{\gamma} = \frac{(1-\gamma)\,R - \gamma\pi\,(1+r_L)\,L}{(1-\gamma)\,(R+\pi\,(1+r_L)\,L)}. \tag{13.63}$$

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

$$\begin{aligned} \pi < \hat{\pi}^* &= \frac{(1-\gamma)\,(1-\hat{\gamma}^*)}{1-(1-\gamma)\,(1-\hat{\gamma}^*)}\frac{R}{(1+r_L)\,L} \\ &= \frac{(1+\hat{r}_D^*)\,D - R}{(1+r_L)\,L}. \end{aligned} \tag{13.64}$$

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 (13.61) and (13.61), we find the deposit rates of those remaining with the bank is given by
(13.65)

$$\left(1+\hat{r}_D^2\right) D = \begin{cases} R + \pi\left(1+r_L\right) L + \frac{\pi(1+r_L)L-((1+\hat{r}_D^*)D-R)}{\pi(1+r_L)L} R & \text{if} \quad \pi > \hat{\pi}^* \\ R + \pi\left(1+r_L\right) L & \text{if} \quad \pi < \hat{\pi}^* \end{cases}.$$

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 14 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)

13.4 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

$$\gamma\left(1+r_D^1\right)D = D - L. \tag{13.66}$$

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

$$(1-\gamma)\left(1+r_D^2\right)D = \pi\left(1+r_L\right)L. \tag{13.67}$$

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 (13.66) and (13.67). We obtain the utility of the depositors, who do not know whether they want to withdraw early or late, as

$$\Pi_D = \gamma u\left(\left(1+r_D^1\right)D\right) + (1-\gamma)\,u\left(\left(1+r_D^2\right)D\right), \tag{13.68}$$

which after inserting for $\left(1+r_D^i\right)D$ from equations (13.66) and (13.67) gives us the optimal level of investment into loans L for the bank as

$$\frac{\partial u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D} = \pi\left(1+r_L\right)\frac{\partial u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D}. \tag{13.69}$$

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

$$\gamma \leq \gamma^* = \frac{D-(1-\lambda)L^*}{\left(1+r_D^{1,*}\right)D}. \tag{13.70}$$

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 (13.66) 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

$$\Pi_B = \pi\left(1+r_L\right)\hat{L} - (1-\gamma)\left(1+\hat{r}_D^2\right)D. \tag{13.71}$$

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

$$\gamma u\left(\left(1+\hat{r}_D^1\right)D\right) + (1-\gamma)\,u\left(\left(1+\hat{r}_D^2\right)D\right) = u\left(D\right). \tag{13.72}$$

Inserting the constraint arising from early withdrawal in equation (13.66) into equation (13.72), we can maximize the bank profits in equation (13.71) 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 (13.72). With ξ denoting the Lagrange multiplier, we easily get the first order conditions as

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \hat{L}} &= \pi\left(1+r_L\right) - (1-\gamma)\frac{\partial\left(1+\hat{r}_D^2\right)D}{\partial \hat{L}} \\ &\quad -\xi\left(-\frac{\partial u\left(\left(1+\hat{r}_D^1\right)D\right)}{\partial\left(1+\hat{r}_D^1\right)D}\right. \\ &\quad \left. +(1-\gamma)\frac{\partial\left(1+\hat{r}_D^2\right)D}{\partial \hat{L}}\frac{\partial u\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D}\right) = 0 \\ \frac{\partial \mathcal{L}}{\partial\left(1+\hat{r}_D^2\right)D} &= -(1-\gamma) - \xi(1-\gamma)\frac{\partial u\left(\left(1+\hat{r}_D^2\right)D\right)}{\partial\left(1+\hat{r}_D^2\right)D} = 0. \end{aligned} \tag{13.73}$$

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 (13.69) 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 (13.66) 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 (13.69) 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 (13.70); we obtain

$$\gamma \leq \hat{\gamma}^* = \frac{D-(1-\lambda)\hat{L}^*}{\left(1+\hat{r}_D^{1,*}\right)D}. \tag{13.74}$$

As $\hat{L}^* > L^*$ and from equation (13.66) we obtain that $\left(1+\hat{r}_D^{1,*}\right)D < \left(1+r_D^{1,*}\right)D$, we can now show that

$$\gamma^* < \hat{\gamma}^*. \tag{13.75}$$

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.

14 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 14.1 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 14.2. Similarly, chapter 14.3 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 14.4 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 14.5 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.

14.1 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

$$U = \theta \frac{R - R^*}{D} + (1 - \theta) \Pi_B^0. \tag{14.1}$$

If the bank borrows the amount of M from another bank at rate $\hat{r}_M$, this utility changes to

$$U_B = \theta \frac{(R - R^*) + M}{D + M} + (1 - \theta) \left(\Pi_B^0 - \hat{r}_M M \right) \tag{14.2}$$

and if the bank provides a loan M to another bank, its utility will be

$$U_L = \theta \frac{(R - R^*) - M}{D} + (1 - \theta) \left(\Pi_B^0 + r_M M \right). \tag{14.3}$$

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

$$\begin{aligned} \hat{r}_M \leq \hat{r}_M^* &= \frac{\theta}{1-\theta} \frac{D - (R - R^*)}{D (D + M)}, \\ r_M \geq r_M^* &= \frac{\theta}{1-\theta} \frac{1}{D}. \end{aligned} \tag{14.4}$$

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)

14.2 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 3.1 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

$$\gamma \left(1 + r_D^1\right) D = R, \tag{14.5}$$

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

$$(1 - \gamma)\left(1 + r_D^2\right) D = \pi\,(1 + r_L)\,L. \tag{14.6}$$

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

$$M_H = \gamma_H\left(1+r_D^1\right)D - R. \tag{14.7}$$

Similarly, banks facing the low withdrawal rate γ_L have excess cash reserves to the amount of

$$M_L = R - \gamma_L\left(1+r_D^1\right)D. \tag{14.8}$$

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 (14.7) and (14.8), we recover condition (14.5) 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

$$\begin{aligned}(1+r_M)M_H &= \pi(1+r_L)L - (1-\gamma_H)\left(1+r_D^2\right)D \\ &= \frac{\gamma_H-\gamma}{1-\gamma}\pi(1+r_L)L,\end{aligned} \tag{14.9}$$

where the final equality has been obtained from solving equation (14.6) for $\left(1+r_D^2\right)D$ and inserting here. Solving equation (14.5) for $\left(1+r_D^1\right)D$and inserting this into equation (14.7) we obtain

$$M_H = \frac{\gamma_H-\gamma}{\gamma}R. \tag{14.10}$$

Once we insert this expression into equation (14.9), we get

$$1+r_M = \frac{\gamma}{1-\gamma}\pi(1+r_L)\frac{L}{R}. \tag{14.11}$$

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

$$\Pi_L = \pi(1+r_L)L - (1-\gamma_L)\left(1+r_D^2\right)D + (1+r_M)M_L = 0, \tag{14.12}$$

where the final equation emerges if we insert for $\left(1+r_D^2\right)D$ from equation (14.6), 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 (14.10). 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 $L = D - R$

$$\gamma \geq \frac{R}{D}. \tag{14.13}$$

Hence if the rate of early withdrawals is too low, banks would have an incentive to reduce their cash holdings. Using equation (14.5), we can rewrite equation (14.13) as

$$r_D^1 \leq 0, \tag{14.14}$$

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)

14.3 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

$$\begin{aligned} (14.15) \quad Prob\left(\hat{S} > R - S + M_2 + M_3\right) &= 1 - \frac{R - S + M_2 + M_3}{R}, \\ Prob\left(\hat{S} > R - M_i\right) &= 1 - \frac{R - M_i}{R} \end{aligned}$$

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

$$\begin{aligned} (14.16) \qquad & Prob\left(\hat{S} > R - S + M_2 + M_3\right) + Prob\left(\hat{S} > R - M_2\right) \\ & + Prob\left(\hat{S} > R - M_3\right) = \frac{S}{R}. \end{aligned}$$

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

$$(14.17) \qquad S \leq \frac{R}{p}.$$

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

$$\Pi_B^i = p\left(1-\frac{R-S+M_2+M_3}{R}\right)M_i + p\left(1-\frac{R-M_i}{R}\right)E. \tag{14.18}$$

Minimizing over the size of the interbank loan gives us for bank i the first order condition

$$\frac{\partial \Pi_B^i}{\partial M_i} = \frac{p}{R}\left(M - M_j + E - 2M_i\right) = 0, \tag{14.19}$$

where j indicates the other bank providing the interbank loan. These two conditions for banks 2 and 3 solve for

$$M_2 = M_3 = \frac{S+E}{3}. \tag{14.20}$$

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 (14.18) we obtain the total losses to the banks providing interbank loans as

$$\Pi_B^2 = \Pi_B^3 = \frac{1}{9}\left(S+E\right)^2. \tag{14.21}$$

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 $\frac{1}{9}\left(M+E\right)^2 \leq \frac{E}{3}$. This can easily be shown to require

$$S \leq \sqrt{3E} - E. \tag{14.22}$$

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 (14.20) we see that interbank loan actually given is larger than this minimum if

$$S < 3R + 2E. \tag{14.23}$$

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)

14.4 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

$$\gamma_L D \leq R \leq \gamma_H D \tag{14.24}$$

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

$$(14.25) \qquad \Pi_B^{H,i} = \pi_i \Big((1 + r_L) \left(1 - \alpha_i^H\right) L + \hat{\alpha}_i^H \left(R + \alpha_i^H \lambda_i L\right) - \left(1 + r_M^i\right) M_i^H - (1 - \gamma_H)(1 + r_D) D \Big).$$

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 (14.25) 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

$$\gamma_H D + \hat{\alpha}_i^H \left(R + \alpha_i^H \lambda_i L\right) \leq R + \alpha_i^H \lambda_i L + M_i^H, \tag{14.26}$$

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:

$$\begin{aligned}
M_i^H &\geq 0, && (14.27)\\
\alpha_i^H &\geq 0, && (14.28)\\
\alpha_i^H &\leq 1, && (14.29)\\
\hat{\alpha}_i^H &\geq 0, && (14.30)\\
\hat{\alpha}_i^H &\leq 0. && (14.31)
\end{aligned}$$

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

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial M_i^H} &= -\pi_i \left(1 + r_M^i\right) + \xi_1 + \xi_2 = 0, && (14.32)\\
\frac{\partial \mathcal{L}}{\partial \alpha_i^H} &= -\pi_i \left(1 + r_L\right) L + \pi_i \hat{\alpha}_i^H \lambda_i L && (14.33)\\
&\quad + \xi_1 \left(1 - \hat{\alpha}_i^H\right) \lambda_i L + \xi_3 - \xi_4 = 0, &&\\
\frac{\partial \mathcal{L}}{\partial \hat{\alpha}_i^H} &= \pi_i \left(R + \alpha_i^H \lambda_i L\right) - \xi_1 \left(R + \alpha_i^H \lambda_i L\right) + \xi_5 - \xi_6 = 0. && (14.34)
\end{aligned}$$

Let us now assume that the bank seeks an interbank loan, $M_i^H > 0$ and hence $\xi_2 = 0$ as constraint (14.27) is not binding. Thus the first order condition (14.32) solves for

$$\xi_1 = \pi_i \left(1 + r_M^i\right) > 0 \tag{14.35}$$

and after inserting this expression, the first order condition (14.34) simplifies to

$$-\pi_i r_M^i \left(R + \alpha_i^H \lambda_i L\right) + \xi_5 - \xi_6 = 0. \tag{14.36}$$

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 (14.30) 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 (14.33) as

$$-\pi_i \left((1 + r_L) - \left(1 + r_M^i\right) \lambda_i \right) L + \xi_3 - \xi_4 = 0. \tag{14.37}$$

This implies that for $1 + r_L - \left(1 + r_M^i\right) \lambda_i > 0$, we need $\xi_3 > 0$, therefore from constraint (14.28) that $\alpha_i^H = 0$ and no loans are liquidated if this condition is fulfilled, which can be rewritten as

$$1 + r_M^i \leq \frac{1 + r_L}{\lambda_i}. \tag{14.38}$$

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 (14.38) 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 (14.26) 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, $\left(1 + r_M^H\right) M_i^{L,H}$ $\left(\left(1 + r_M^L\right) M_i^{L,L}\right)$ 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

$$\begin{aligned} \Pi_B^{L,i} = \; & \pi_i \left((1 + r_L) \left(1 - \alpha_i^L\right) L + \hat{\alpha}_i^L \left(R + \alpha_i^L \lambda_i L\right) \right. \\ & + \pi_H \left(1 + r_M^H\right) M_i^{L,H} + \pi_L \left(1 + r_M^L\right) M_i^{L,L} \\ & \left. - (1 - \gamma_L)(1 + r_D) D \right). \end{aligned} \tag{14.39}$$

The maximization of the bank profits in equation (14.25) 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

$$\gamma_L D + \hat{\alpha}_i^L \left(R + \alpha_i^L \lambda_i L\right) + M_i^{L,H} + M_i^{L,L} \leq R + \alpha_i^L \lambda_i L, \tag{14.40}$$

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 $\hat{\xi}_7$:

$$\begin{aligned}
M_i^{L,H} &\geq 0, && (14.41)\\
M_i^{L,L} &\geq 0, && (14.42)\\
\alpha_i^L &\geq 0, && (14.43)\\
\alpha_i^L &\leq 1, && (14.44)\\
\hat{\alpha}_i^L &\geq 0, && (14.45)\\
\hat{\alpha}_i^L &\leq 1. && (14.46)
\end{aligned}$$

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

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial M_i^{L,H}} &= \pi_i \pi_H \left(1 + r_M^H\right) - \hat{\xi}_1 + \hat{\xi}_2 = 0, && (14.47)\\
\frac{\partial \mathcal{L}}{\partial M_i^{L,L}} &= \pi_i \pi_L \left(1 + r_M^L\right) - \hat{\xi}_1 + \hat{\xi}_3 = 0, && (14.48)\\
\frac{\partial \mathcal{L}}{\partial \alpha_i^L} &= -\pi_i \left(1 + r_L\right) L + \hat{\alpha}_i^L \pi_i \lambda_i L + && (14.49)\\
&\quad \left(1 - \hat{\alpha}_i^L\right) \lambda_i L \hat{\xi}_1 + \hat{\xi}_4 - \hat{\xi}_5 = 0, &&\\
\frac{\partial \mathcal{L}}{\partial \hat{\alpha}_i^L} &= \pi_i \left(R + \alpha_i^L \lambda_i L\right) - \hat{\xi}_1 \left(R + \alpha_i^L \lambda_i L\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. && (14.50)
\end{aligned}$$

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 (14.41) that this constraint is not binding and hence $\hat{\xi}_2 = 0$. Then the first order condition in equation (14.47) solves for

$$\hat{\xi}_1 = \pi_i \pi_H \left(1 + r_M^H\right) > 0, \tag{14.51}$$

thus constraint (14.40) is binding.

Inserting for $\hat{\xi}_1$ from equation (14.51) into equation (14.50) we obtain

$$\pi_i \left(R + \alpha_i^L \lambda_i L\right) \left(1 - \pi_H \left(1 + r_M^H\right)\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. \tag{14.52}$$

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 (14.46) 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 (14.42) is binding and hence $\hat{\xi}_3 > 0$, giving us from first order condition (14.48) that

$$\hat{\xi}_1 = \pi_i \pi_L \left(1 + r_M^L\right) + \hat{\xi}_3 > 0. \tag{14.53}$$

Inserting this result into the first order condition (14.50) we get

$$\pi_i \left(R - \alpha_i^L \lambda_i L\right) \left(1 - \pi_L \left(1 + r_M^L\right)\right) - \hat{\xi}_3 \left(R - \alpha_i^L \lambda_i L\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0. \tag{14.54}$$

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 (14.45) that this is binding, thus $\hat{\alpha}_i^L = 0$ and no cash is retained. But from equation (14.53) we know that constraint (14.40) 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

$$\pi_L \left(1 + r_M^L\right) \leq 1 \leq \pi_H \left(1 + r_M^H\right). \tag{14.55}$$

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 (14.41) and (14.42) are not binding. From the first order conditions (14.47) and (14.48) we then immediately get that the returns on interbank loans to the two types of banks are identical,

$$\pi_L \left(1 + r_M^L\right) = \pi_H \left(1 + r_M^H\right) \tag{14.56}$$

and

$$\hat{\xi}_1 = \pi_i \pi_H \left(1 + r_M^H\right) > 0. \tag{14.57}$$

The first order condition (14.50) then becomes

$$\pi_i \left(R + \alpha_i^L \lambda_i L\right) \left(1 - \pi_H \left(1 + r_M^H\right)\right) + \hat{\xi}_6 - \hat{\xi}_7 = 0 \tag{14.58}$$

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 (14.46) 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 (14.56), 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 (14.48) becomes

$$\hat{\xi}_1=\pi_i\pi_L\left(1+r_M^L\right)>0 \tag{14.59}$$

and first order condition (14.50) will be

$$\pi_i\left(R+\alpha_i^L\lambda_i L\right)\left(1-\pi_L\left(1+r_M^L\right)\right)+\hat{\xi}_6-\hat{\xi}_7=0. \tag{14.60}$$

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 (14.47) 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 (14.49) as

$$-\pi_i\left(1+r_L-\lambda_i\pi_H\left(1+r_M^H\right)\right)L+\hat{\xi}_4-\hat{\xi}_5=0. \tag{14.61}$$

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 (14.43) 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

$$\pi_H\left(1+r_M^H\right)=\pi_L\left(1+r_M^L\right)\leq\frac{1+r_L}{\lambda_i} \tag{14.62}$$

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 (14.62)is fulfilled for banks with high repayment rates, it will also be fulfilled for banks with low repayment rates. And similarly, if equation (14.69) 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 (14.62) is fulfilled they do not liquidate any loans, $\hat{\alpha}_i^H = 0$. The constraint on cash reserves from equation (14.40) 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 (14.62) for banks to not liquidating loans to provide more interbank loans and equation (14.38) for banks to seek interbank loans rather than liquidate loans, and see that the latter is more strict and hence we require that

$$1 + r_M^i \leq \frac{1 + r_L}{\lambda_i} \tag{14.63}$$

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 (14.25) and (14.39), obtaining the profits as

$$\begin{aligned} \Pi_B^{H,i} &= \pi_i \left((1 + r_L) L - \left(1 + r_M^i\right) M_i^H - (1 - \lambda_H)(1 + r_D) D \right), \\ \Pi_B^{L,i} &= \pi_i \Big((1 + r_L) L + \pi_H \left(1 + r_M^H\right) M_i^{L,H} \\ &\quad + \pi_L \left(1 + r_M^L\right) M_i^{L,L} - (1 - \lambda_L)(1 + r_D) D \Big). \end{aligned} \tag{14.64}$$

Similarly, the constraints on interbank loans from constraints (14.26) and (14.40), who are binding, become

$$\begin{aligned} M_i^{L,H} + M_i^{L,L} &= R - \lambda_L D, \\ M_i^H &= \lambda_H D - R. \end{aligned} \tag{14.65}$$

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

$$\Pi_B = \gamma_L \left(p\Pi_B^{L,H} + (1-p)\,\Pi_B^{L,L}\right) + \gamma_H \left(p\Pi_B^{H,H} + (1-p)\,\Pi_B^{H,L}\right). \tag{14.66}$$

Inserting equations (14.65) into equation (14.64) and this result into equation (14.66), 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

$$\frac{\partial \Pi_B}{\partial L} = (\gamma_H + \gamma_L)\,\pi\,(1 + r_L) - \pi_H\,(\gamma_H + \gamma_L \pi)\left(1 + r_M^H\right) = 0, \tag{14.67}$$

where $\pi = p\pi_H + (1-p)\,\pi_L$ denotes the average repayment rate of loans. This first order condition solves for

$$1 + r_M^H = \frac{\pi}{\pi_H}\frac{\gamma_H + \gamma_L}{\gamma_H + \pi\gamma_L}\,(1 + r_L)\,. \tag{14.68}$$

Inserting this result into the condition that no interbank loans are liquidated, equation (14.63), we get the requirement that

$$\frac{\pi}{\pi_H}\frac{\gamma_H + \gamma_L}{\gamma_H + \pi\gamma_L} \leq \frac{1}{\lambda_i}. \tag{14.69}$$

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

$$\pi_L \leq \pi_L^* = \frac{\gamma_H \pi_H}{(1-p)\left(\lambda_i\,(\gamma_H + \gamma_L) - \gamma_L \pi_H\right)} - \frac{p}{1-p}\pi_H. \tag{14.70}$$

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 (14.68), this easily solves for

$$\pi_L \geq \pi_L^{**} = \frac{\gamma_H \pi_H}{(1-p)\left((\gamma_H + \gamma_L)\,(1 + r_L) - \gamma_L \pi_H\,(1 + r_D)\right)} - \frac{p}{1-p}\pi_H. \tag{14.71}$$

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 (14.69) 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 (14.26) is not binding, hence $\xi_1 = 0$, and from the first order condition (14.34) we easily get that

$$\pi_H \left(R + \alpha_H^H \lambda_H L\right) + \xi_5 - \xi_6 = 0, \tag{14.72}$$

and as the first term is positive, we require that $\xi_6 > 0$. Hence from constraint (14.31) we obtain that $\hat{\alpha}_H^H = 1$, thus constraint (14.26) becomes $\gamma_H D \leq 0$, which is a contradiction and we will have $\xi_1 > 0$ such that constraint (14.26) is binding and as $R \leq \gamma_H D$, we need $\alpha_H^H > 0$, implying from constraint (14.28) that $\xi_3 = 0$. Inserting this into the first order condition (14.33) we get

$$(\pi_H - \xi_1)\,\hat{\alpha}_H^H \lambda_H L + (\xi_1 \lambda_H - \pi_H (1 + r_L))\, L - \xi_4 = 0. \tag{14.73}$$

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 (14.30) we find $\xi_5 = 0$ and the first order condition (14.34) becomes

$$(\pi_H - \xi_1)\left(R + \alpha_H^H \lambda_H L\right) - \xi_6 = 0. \tag{14.74}$$

From equation (14.35) 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

$$\begin{aligned} M_L^H &= \lambda_H D - R, \\ M_H^H &= 0, \\ M_i^{L,H} + M_i^{L,L} &= R - \lambda_L D, \\ \alpha_H^H \lambda_H L &= \lambda_H D - R. \end{aligned} \tag{14.75}$$

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 (14.75) into the profits of the bank in equation (14.64) 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

$$1 + r_M^H = \frac{\pi(\gamma_H + \gamma_L) - \gamma_H p \pi_H \frac{1}{\lambda_H}}{(1-p)\,\gamma_H + \pi\gamma_L}\,\frac{1 + r_L}{\pi_H}. \tag{14.76}$$

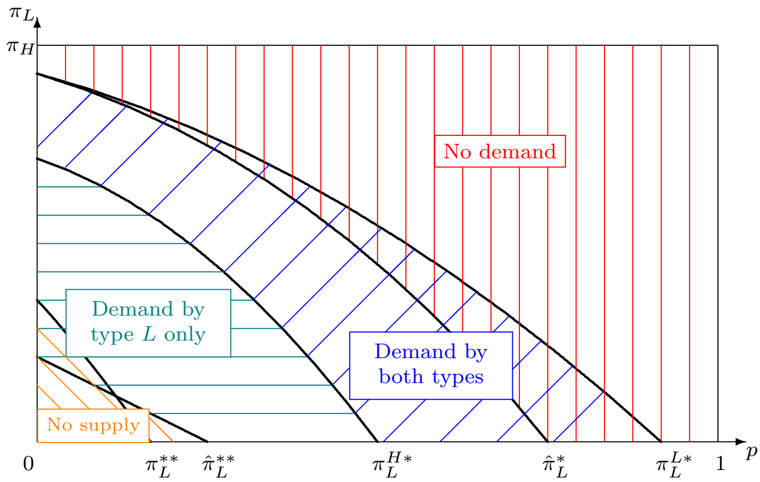


Figure 15: Interbank loans and liquidation

Inserting this interbank loan rate into equation (14.63), we obtain that

$$\pi_L \leq \hat{\pi}_L^* = \frac{\pi_H (1-p) \gamma_H + \frac{\lambda_L \gamma_H p \pi_H}{\lambda_H}}{(1-p)(\lambda_L (\gamma_H + \gamma_L) - \pi_H \gamma_L)} - \frac{p}{1-p} \pi_H. \tag{14.77}$$

Thus if π_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 $\pi_H \left(1 + r_M^H\right) \geq 1 + r_D$

$$\pi_L \geq \hat{\pi}_L^{**} = \frac{\frac{\gamma_H p \pi_H}{\lambda_H} + (1-p) \pi_H}{(\gamma_H + \gamma_L) \frac{1+r_L}{1+r_D} - \gamma_L} - \frac{p}{1-p} \pi_H. \tag{14.78}$$

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 15 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.

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)

14.5 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

$$\begin{aligned} \Pi_B &= \pi (1 + r_L) (D - B + M_0) + (1 + r) B \\ &\quad - (1 + r_M) M_0 - (1 + r_D) D, \end{aligned} \tag{14.79}$$

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

$$\hat{\Pi}_B = (1 + r_M) (D - B) + (1 + r) B - (1 + r_D) D. \tag{14.80}$$

In order for banks to provide loans to the general public, we need $\Pi_B \geq \hat{\Pi}_B$, or

$$\pi \geq \pi^* = \frac{1 + r_M}{1 + r_L}, \tag{14.81}$$

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

$$\Pi_B^S = \pi (1 + r_L) (D - B + M_0) - (1 + r_D) D. \tag{14.82}$$

If repaying the interbank loan, the bank would make the profits as detailed in equation (14.79), 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

$$\Pi_B^R = \Pi_B + \left(\pi\left(1+r_L\right) - \left(1+r_M\right)\right)\left(M - M_0\right), \tag{14.83}$$

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 (14.81) 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 $M_{k+1} = \lambda\pi\left(1+r_L\right)M_k$. The total revenue arising from these k interbank loans is $\pi\left(1+r_L\right)\left(D-B-M_0\right) + \sum_{j=1}^k \left(\pi\left(1+r_L\right)\right)^j \left(D-B-M_0\right)$ and the costs of these loans are $\left(1+r_M\right)\sum_{j=1}^k M_j$. Thus the profits generated to the bank is given by

$$\begin{aligned} \hat{\Pi}_B^R &= \pi\left(1+r_L\right)\left(D-B-M_0\right) + \left(1+r\right)B - \left(1+r_D\right)D \\ &\quad + \sum_{j=1}^k \left(\lambda\pi\left(1+r_L\right)\right)^j \left(D-B-M_0\right) - \left(1+r_M\right)\sum_{j=1}^k M_j. \end{aligned} \tag{14.84}$$

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}, $\lambda\pi\left(1+r_L\right)M_{k-1}$ and do not repay this interbank loan. Hence the profits with strategic default are given by

$$\begin{aligned} \hat{\Pi}_B^S &= \pi\left(1+r_L\right)\left(D-B-M_0\right) + \left(1+r\right)B - \left(1+r_D\right)D \\ &\quad + \sum_{j=1}^{k-1} \left(\lambda\pi\left(1+r_L\right)\right)^j \left(D-B-M_0\right) - \left(1+r_M\right)\sum_{j=1}^{k-1} M_j. \end{aligned} \tag{14.85}$$

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

$$M_k \leq \frac{\left(\lambda\pi\left(1+r_L\right)\right)^k \left(D-B+M_0\right)}{1+r_M}. \tag{14.86}$$

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 $M_k = \lambda\pi\left(1+r_L\right)M_{k-1}$, 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 (14.86) gives us

$$M_1 \leq \frac{\lambda\pi\,(1 + r_L)\,(D - B + M_0)}{1 + r_M}. \tag{14.87}$$

The total amount of interbank lending is now given by summing up all the possible interbank loans, such that

$$\begin{aligned} M &= M_0 + \sum_{k=1}^{+\infty} M_k \\ &= M_0 + \sum_{k=1}^{+\infty} \left(\lambda\pi^*\,(1 + r_L)\right)^{k-1} M_1 \\ &= M_0 + \sum_{k=0}^{+\infty} \left(\lambda\pi^*\,(1 + r_L)\right)^{k} M_1 \\ &= M_0 + \frac{M_1}{1 - \lambda\pi^*\,(1 + r_L)}, \end{aligned} \tag{14.88}$$

where we assume that $\lambda\pi^*\,(1 + r_L) < 1$. 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

$$M_0 = \frac{1 + r}{1 + r_M} B. \tag{14.89}$$

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.

15 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 15.1 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 15.2 we will explore under which condition such a repo run could lead to the failure of the bank.

15.1 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

$$\hat{\mathcal{L}}_1 = \left(\hat{P}_1 - \lambda(1+r)^2 B\right)\left((1+r)^2 B - \hat{P}_1\right), \tag{15.1}$$

whose first order condition for a maximum, $\frac{\partial \hat{\mathcal{L}}_1}{\partial \hat{P}_1} = 0$, yields

$$\hat{P}_1^* = \frac{1}{2} (1 + \lambda) (1 + r)^2 B \leq (1 + r)^2 B, \tag{15.2}$$

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 (15.2), 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

$$\mathcal{L}_1 = \left(P_1 - \hat{P}_1\right) \left((1 + r_B)^2 B - P_1\right), \tag{15.3}$$

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 (15.2) for $\hat{P}_1^*$ that

$$P_1 = \frac{1}{2} \left(\hat{P}_1 + (1 + r)^2 B\right) = \frac{\lambda + 3}{4} (1 + r)^2 B. \tag{15.4}$$

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

$$\mathcal{L}_0^R = \left(P_1 - P_0^R\right) \left(\pi (1 + r_L) P_0^R - P_1\right). \tag{15.5}$$

The first order condition for maximizing the objective function, $\frac{\partial \mathcal{L}_0^R}{\partial P_0^R} = 0$ solves for

$$P_0^R = \frac{\lambda + 3}{4} \frac{\pi (1 + r_L) - 1}{\pi (1 + r_L)} (1 + r)^2 B \leq P_1 \tag{15.6}$$

when inserting for P_1 from equation (15.4).

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

$$\mathcal{L}_0^R=(P_1-P_0)\left(\pi(1+r_L)P_0-(1+r)^2 B\right), \tag{15.7}$$

from which we obtain from the first order condition of maximizing this expression, $\frac{\partial \mathcal{L}_0^R}{\partial P_0}=0$, that

$$P_0=\left(\frac{\lambda+3}{8}-\frac{1}{2\pi(1+r_L)}\right)(1+r)^2 B\leq P_1. \tag{15.8}$$

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

$$\begin{aligned}\Pi_R &= \left(\pi(1+r_L)P_0^R-P_1\right)+\left(P_1-P_0^R\right)\\ &= P_0\left(\pi(1+r_L)-1\right),\end{aligned} \tag{15.9}$$

where we obtain the second equality from inserting the expressions in equations (15.4) and (15.6).

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 (15.4) and (15.8), we get the total surplus when no repurchase agreement is in place as

$$\begin{aligned} \Pi &= \left(\pi\left(1+r_L\right)P_0-\left(1+r\right)^2 B\right)+\left(P_1-P_0\right) \\ &= P_0\left(\pi\left(1+r_L\right)-1\right)-\left(\left(1+r\right)^2 B-P_1\right). \end{aligned} \tag{15.10}$$

As $P_1 \leq \left(1+r_B\right)^2 B$ from our result in equation (15.4), 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 (15.4) and (15.6), we can get the implied interest rate for this repurchase agreement, the Repo rate, as

$$1+r_R=\frac{P_1}{P_0^R}=\frac{\pi\left(1+r_L\right)}{\pi\left(1+r_L\right)-1}, \tag{15.11}$$

where $\pi\left(1+r_L\right)>1$ 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, $\pi\left(1+r_L\right)$. 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)

15.2 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

$$\Pi_B^t = \pi\left(1+r_L\right)B_{t-2} - p\left(1+r_R\right)^2 B_{t-2} - \left(1-p\right)\left(1+r_R\right)B_{t-1}. \tag{15.12}$$

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

$$\hat{\Pi}_B^t = \pi\left(1+r_L\right)^2 B_{t-2} - B_{t-2}. \tag{15.13}$$

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 (15.12) and (15.13) and setting these equal we obtain $(1-p)(1+r_R)^2 + p(1+r_R) = 1$, for which one solution is

$$1 + r_R = 1. \tag{15.14}$$

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

$$\begin{aligned} M &= p(1+r_R)^2 B + (1+r_R) B - \pi (1+r_L)^2 B \\ &= (1 + p - \pi (1+r_L)) B. \end{aligned} \tag{15.15}$$

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 (15.14) 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 (15.15) this condition becomes

$$p \leq \frac{(1 + r_R + \lambda)\pi(1+r_L) - (1+r_R)}{1+r_R}. \tag{15.16}$$

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 (15.15). 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.

16

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 16.1 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 16.2, 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 16.3. We will analyse whether banks, depositors or general taxation should be used to pay for deposit insurance.

16.1 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 16.1.1 we will investigate how deposit insurance not set at the correct price can affect the risk-taking of banks and in chapter 16.1.2 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 16.1.3 will explore how banks make decisions that require bailouts and how deposit insurance can influence such decisions.

16.1.1 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

$$\Pi_B = \pi \left(1 + r_L\right) L - \left(1 + r_D\right) D - E. \tag{16.1}$$

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

$$1 + r_L = \frac{\left(1 + r_D\right) D + E}{\pi \left(D + E\right)}. \tag{16.2}$$

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

$$\hat{\Pi}_B = \pi \left(\left(1 + r_L\right) L - \left(1 + r_D\right) D\right) - P - E. \tag{16.3}$$

If the bank sets the loan rate competitively as obtained in equation (16.2) and we use that $D = L + E$, the bank profits with deposit insurance are given by

$$\hat{\Pi}_B = \left(1 - \pi\right) \left(1 + r_D\right) D - P. \tag{16.4}$$

If we define $\kappa = \frac{E}{D}$ as the equity ratio, we can rewrite the bank profits in equation (16.4) as

$$\begin{aligned} \hat{\Pi}_B &= \left(1 - \pi\right) \left(1 + r_D\right) \frac{1}{\kappa} E - P \\ &= \left(1 - \pi\right) \left(1 + r_D\right) \frac{1}{1 + \kappa} L - P. \end{aligned} \tag{16.5}$$

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 (16.2), 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

$$P = (1 - \pi)(1 + r_D) D = (1 - \pi)(1 + r_D) \frac{1}{1 + \kappa} L. \tag{16.6}$$

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)

16.1.2 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 $\pi(1 + r_L) L$. Hence with deposits repaid at time T, the payment of the deposit insurance is given by

$$P_T = \max\left\{0; \left(1 + r_D\right) D - \pi \left(1 + r_L\right) L\right\} \tag{16.7}$$

This expression represents the payoff of a put option with maturity of time T, where $(1 + r_D) D$ represents the strike price and $\pi (1 + r_L) L$ 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 $\sigma^2 = N\pi (1 - \pi) L_i^2 = \pi (1 - \pi) \frac{L^2}{N}$.

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

$$P_0 = (1 + r_D) D\Phi (d_2) - \pi (1 + r_L) L\Phi (d_1), \tag{16.8}$$

$$d_1 = \frac{1}{\sigma} \ln \frac{1 + r_D}{\pi (1 + r_L)} \frac{D}{L} - \frac{1}{2}\sigma, \tag{16.9}$$

$$d_2 = d_1 + \sigma. \tag{16.10}$$

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 (16.8), 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)

16.1.3 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

$$P = p\left(\left(1 + r_D\right) D - \lambda\pi\left(1 + r_L\right) L\right), \tag{16.11}$$

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

$$\begin{aligned} \hat{P} &= p^2\left(\left(1 + r_D\right) D - \lambda\pi\left(1 + r_L\right) L\right) \\ &\quad + p(1 - p)\left(\left(1 + r_D\right) D - \hat{\lambda}\pi\left(1 + r_L\right) L\right) \\ &= P - p(1 - p)\left(\hat{\lambda} - \lambda\right)\pi\left(1 + r_L\right) L. \end{aligned} \tag{16.12}$$

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

$$\begin{aligned} P^* &= p\left(\left(1+r_D\right)D-\lambda\pi\left(1+r_L\right)L \right. \\ &\quad \left. +\min\left\{C;\left(1-\lambda\right)\pi\left(1+r_L\right)L\right\}\right), \\ \hat{P}^* &= p^2\left(\left(1+r_D\right)D-\lambda\pi\left(1+r_L\right)L \right. \\ &\quad \left. +\min\left\{C;\left(1-\lambda\right)\pi\left(1+r_L\right)L\right\}\right) \\ &\quad +p\left(1-p\right)\left(\left(1+r_D\right)D-\hat{\lambda}\pi\left(1+r_L\right)L\right). \end{aligned} \tag{16.13}$$

If banks are bailed out, they are fully recapitalised and they continue their operation, making profits of $\Pi_B^* = \pi\left(1+r_L\right)-\left(1+r_D\right)D$ 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

$$\Pi_B = \left(1-p\right)\left(\pi\left(1+r_L\right)L-\left(1+r_D\right)D\right)+p\Pi_B^*-P^*. \tag{16.14}$$

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

$$\begin{aligned} \hat{\Pi}_B &= \left(1-p\right)^2\left(\pi\left(1+r_L\right)L-\left(1+r_D\right)D\right) \\ &\quad +p\left(1-p\right)\left(\pi\left(1+r_L\right)L-\left(1+r_D\right)D \right. \\ &\quad \left. +\left(\pi\left(1+r_L\right)-\left(1+r_D\right)\hat{\lambda}\pi\left(1+r_L\right)L\right)\right) \\ &\quad +p^2\Pi_B^*-\hat{P}^*. \end{aligned} \tag{16.15}$$

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

$$\begin{aligned} P^*-\hat{P}^* &\geq p\left(1-p\right)\left(\Pi_B^*-\left(\pi\left(1+r_L\right)L \right.\right. \\ &\quad \left.\left. -\left(1+r_D\right)\hat{\lambda}\pi\left(1+r_L\right)L\right)\right). \end{aligned} \tag{16.16}$$

We can now see that if $C > \left(1-\lambda\right)\pi\left(1+r_L\right)L$, then no bail out happens and hence $\Pi_B^* = 0$. As in this case $P-\hat{P} > 0$ and $\pi\left(1+r_L\right)L - \left(1+r_D\right)\hat{\lambda}\pi\left(1+r_L\right)L > 0$, 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 (16.16) and using the expressions for the deposit insurance from equation (16.13), that for $\Pi_B \leq \hat{\Pi}_B$ we need

$$C \leq (1+r_D)\,D - \left(r_D\hat{\lambda} + \lambda\right)\pi\,(1+r_L)\,L. \tag{16.17}$$

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 (16.16) 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 (16.13), the right-hand side of the condition in equation (16.16) 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.

16.2 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 16.2.1 we will investigate whether a deposit insurance is optimal to provide full or partial coverage of deposits, and chapter 16.2.2 then explores how much of their deposits should be covered.

16.2.1 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

$$\text{(16.18)} \qquad \begin{aligned} \Pi_D^{jj} &= \pi\left(1+r_D^j\right)\hat{D}-\hat{D}-(1-\pi)\hat{D}, \\ \Pi_D^{ji} &= \pi\left(1+r_D^i\right)\hat{D}-\hat{D}-(1-\pi)\hat{D}-cd_i. \end{aligned}$$

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

$$\text{(16.19)} \qquad d_i\leq d_i^*=\pi\frac{\left(1+r_D^i\right)-\left(1+r_D^j\right)}{c}\hat{D},$$

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

$$\text{(16.20)} \qquad \begin{aligned} D_i &= \lambda\left(1+2\pi\frac{\left(1+r_D^j\right)-\left(1+r_D^i\right)}{c}\right)2D \\ &\quad +(1-\lambda)\left(1+\pi\frac{\left(1+r_D^j\right)-\left(1+r_D^i\right)}{c}D\right)D \\ &= (1+\lambda)D+\pi(1+3\lambda)\frac{\left(1+r_D^j\right)-\left(1+r_D^i\right)}{c}D^2, \end{aligned}$$

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

$$\text{(16.21)} \qquad \Pi_B^i=\pi\left((1+r_L)-\left(1+r_D^i\right)\right)D_i.$$

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 (16.20) as

$$1+r_D^* = (1+r_L) - \frac{1-\lambda}{\pi(1+3\lambda)}\frac{c}{D}. \tag{16.22}$$

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 (16.21) back into the profits of the bank in equation (16.21), we easily get that bank profits are given by

$$\Pi_B^* = \frac{(1+\lambda)^2}{1+3\lambda}D. \tag{16.23}$$

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

$$\begin{aligned} \Pi_D^{jj} &= \left(1+r_D^j\right)\hat{D} - \hat{D}, \\ \Pi_D^{ji} &= \left(1+r_D^i\right)\hat{D} - \hat{D} - cd_i. \end{aligned} \tag{16.24}$$

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

$$d_i \leq d_i^{**} = \frac{\left(1+r_D^i\right) - \left(1+r_D^j\right)}{c}\hat{D}. \tag{16.25}$$

Following similar steps to the case of no deposit insurance, we get the total amount of deposits at bank i, comparable to equation (16.20), as

$$D_i = (1+\lambda)D + \frac{\left(1+r_D^i\right) - \left(1+r_D^j\right)}{c}(1+3\lambda)D^2. \tag{16.26}$$

Maximizing bank profits in the same way as without deposit insurance gives the optimal deposit rate as

$$1+r_D^{**} = (1+r_L) - \frac{1+\lambda}{3+\lambda}\frac{c}{D} \tag{16.27}$$

and inserting this deposit rate into equation (16.21) we obtain the bank profits as

$$\Pi_B^{**} = \pi \frac{(1+\lambda)^2}{1+3\lambda} D = \pi \Pi_B^{*}. \tag{16.28}$$

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

$$\begin{aligned} \Pi_D^{jj} &= \left(1+r_D^j\right) D + \pi \left(1+r_D^j\right) D - (1-\pi) D \\ \Pi_D^{jij} &= \left(1+r_D^j\right) D + \left(1+r_D^i\right) D - c d_j \\ \Pi_D^{ji} &= \left(1+r_D^i\right) D + \pi \left(1+r_d^i\right) D - (1-\pi) D - c d_j \end{aligned} \tag{16.29}$$

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

$$d_i \leq d_i^{***} = \frac{\left(1+r_D^j\right) - \pi \left(1+r_D^i\right) + (1-\pi)}{c} D. \tag{16.30}$$

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

$$
\begin{aligned}
(16.31)\quad D_i &= \lambda \left(2D - \frac{\left(1+r_D^j\right) - \pi\left(1+r_D^i\right) + (1-\pi)}{c} D^2 \right. \\
&\quad \left. + \frac{\left(1+r_D^i\right) - \pi\left(1+r_D^j\right) + (1-\pi)}{c} D^2 \right) \\
&\quad + (1-\lambda)\left(D + \frac{\left(1+r_D^i\right) - \left(1+r_D^j\right)}{c} D^2 \right) \\
&= (1+\lambda) D - (1+\lambda\pi)\left(\left(1+r_D^j\right) - \left(1+r_D^i\right) \right) \frac{D^2}{c}.
\end{aligned}
$$

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 (16.21) and maximizing profits by choosing the optimal deposit rate, we obtain, again using only symmetric equilibria, that

$$
1 + r_D^{***} = (1 + r_L) - \frac{1+\lambda}{1+\pi\lambda} \frac{c}{D} \tag{16.32}
$$

and hence when inserting this expression back into the profits of the banks in equation (16.21), we obtain

$$
\Pi_B^{***} = \pi \frac{(1+\lambda)^2}{1+\pi\lambda} D. \tag{16.33}
$$

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)

16.2.2 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

$$1+r_D^U=\frac{1+r_D^I}{\pi}. \tag{16.34}$$

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

$$1+r_D^U=\frac{\left(1+r_D^I\right)^2 D_I}{\pi\,(1+r_L)\,L-\left(1+r_D^I\right)D_U}\geq\frac{1+r_D^I}{\pi} \tag{16.35}$$

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

$$\begin{aligned} D_I &= D-D_U>D-\frac{\pi\,(1+r_L)}{1+r_D^I}\,(D-P) \\ &= \frac{\pi\,(1+r_L)}{1+r_D^I}P-\frac{\pi\,(1+r_L)-\left(1+r_D^I\right)}{1+r_D^I}D. \end{aligned} \tag{16.36}$$

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

$$P = \pi \max\left\{0; \left(1+r_D^I\right) D_I - \left(1 - \frac{\left(1+r_D^U\right) D_U}{\left(1+r_D^I\right) D_I + \left(1+r_D^U\right) D_U}\right) \left(1+r_L\right) L\right\} + \left(1-\pi\right)\left(1+r_D^I\right) D_I. \tag{16.37}$$

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

$$P = (1-\pi)\left(1+r_D^I\right) D_I. \tag{16.38}$$

Inserting for the deposit rate of uninsured deposits from equation (16.34) and the deposit insurance premium from equation (16.38) into this constraint, while noting that $L = D - P$ and $D = D_I + D_U$, this solves for

$$D_I \leq \frac{\pi (1+r_L) - \left(1+r_D^I\right)}{(1-\pi)\left(1+r_D^I\right)\left(\pi(1+r_L)-1\right)} D. \tag{16.39}$$

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 (16.35) into the deposit insurance premium from equation (16.37) and obtain

$$P = \left(1+r_D^I\right) D - \pi(1+r_L) L, \tag{16.40}$$

which using that $L = D - P$ becomes

$$P = \frac{\pi(1+r_L) - \left(1+r_D^I\right)}{\pi(1+r_L)-1} D \tag{16.41}$$

Inserting this deposit insurance premium and the deposit rate for uninsured deposits from equation (16.35) into the constraint, we obtain that

$$(1-\pi)\left(1+r_D^I\right) D_I < \frac{1+r_D^I}{\pi(1+r_L)-1}\left(\pi r_L - 1\right) D. \tag{16.42}$$

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

$$\begin{aligned}(16.43)\qquad \Pi_B &= \pi(1+r_L)L - \left(1+r_D^I\right)D_I - \left(1+r_D^U\right)D_U \\ &= \left(\pi(1+r_L) - \frac{1+r_D^I}{\pi}\right)D \\ &\quad + \frac{(1-\pi)\left(1+r_D^I\right)}{\pi}\left(1-\pi^2(1+r_L)\right)D_I\end{aligned}$$

when inserting from equation (16.34) for the deposit rate of uninsured deposits and equation (16.38) 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 $\pi^2(1+r_L) < 1$ and decreasing otherwise. Thus, using the constraint on insured deposits from equation (16.39), we get

$$(16.44)\qquad D_I = \begin{cases} \frac{\pi(1+r_L)-\left(1+r_D^I\right)}{(1-\pi)\left(1+r_D^I\right)(\pi(1+r_L)-1)} & \text{if} \quad \pi^2 < \frac{1}{1+r_L} \\ 0 & \text{if} \quad \pi^2 > \frac{1}{1+r_L} \end{cases}.$$

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.

16.3 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

$$\begin{aligned} \Pi_B^L &= \pi_L \left((1 + r_L) L - (1 + r_D) D \right) \\ &= \pi_L \left((1 + r_L) + (r_L - r_D) (\kappa - 1) \right) E, \end{aligned} \tag{16.45}$$

where r_D denotes the deposit rate and we used that $L = \kappa E$ and $D = (\kappa - 1) E$. 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

$$\begin{aligned} \Pi_B^H &= \pi_H \left((1 + r_L) L - (1 + r_D) D \right) - cL \\ &= \left(\pi_H \left((1 + r_L) + (r_L - r_D) (\kappa - 1) \right) - c\kappa \right) E. \end{aligned} \tag{16.46}$$

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

$$r_L - r_D \geq \frac{c\kappa - (\pi_H - \pi_L)(1 + r_L)}{(\pi_H - \pi_L)(\kappa - 1)}. \tag{16.47}$$

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

$$\begin{aligned} \Pi_D &= \pi_H (1 + r_D) \rho W + \pi_L (1 + r_L)(1 - \rho) W - W \\ &= \pi_H \left((1 + r_L) - (r_L - r_D) \right) \rho W + \pi_L (1 + r_L)(1 - \rho) W - W. \end{aligned} \tag{16.48}$$

If individuals we to invest all their wealth into the company directly, they would obtain

$$\hat{\Pi}_D = \pi_L (1 + r_L) W - W. \tag{16.49}$$

It is more profitable to invest some of their wealth into deposits if $\Pi_D \leq \hat{\Pi}_D$, which easily becomes

$$r_L - r_D \leq \frac{(\pi_H - \pi_L)(1 + r_L)}{\pi_H}. \tag{16.50}$$

Combining equations (16.47) and (16.50), we easily see that for a viable solution, we need the leverage ratio to be not too high as we obtain that we require

$$\kappa \leq \kappa^* = \frac{\pi_L (\pi_H - \pi_L)(1 + r_L)}{c\pi_H - (\pi_H - \pi_L)^2 (1 + r_L)}. \tag{16.51}$$

We can easily see from equation (16.46) 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 (16.47) and (16.50) become equalities and hence solving equation (16.50) we obtain the deposit rate to be

$$1 + r_D = \frac{\pi_L}{\pi_H} (1 + r_L). \tag{16.52}$$

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

$$c < \frac{\pi_H - \pi_L}{\pi_H} (1 + r_L). \tag{16.53}$$

From equation (16.51) we also require that

$$c \geq \frac{(\pi_H - \pi_L)^2}{\pi_H} (1 + r_L). \tag{16.54}$$

Combining these two requirements we obtain that

$$\frac{(\pi_H - \pi_L)^2}{\pi_H} (1 + r_L) \leq c < \frac{\pi_H - \pi_L}{\pi_H} (1 + r_L). \tag{16.55}$$

As a further constraint, we need that the leverage is exceeding 1, $\kappa^* > 1$ as otherwise banks do not use deposits. Hence from equation (16.51) we require

$$c < (\pi_H - \pi_L)(1 + r_L), \tag{16.56}$$

which is more restrictive than the upper constraint in equation (16.55) and this constraint becomes

$$\frac{(\pi_H - \pi_L)^2}{\pi_H}(1 + r_L) < c \leq (\pi_H - \pi_L)(1 + r_L). \tag{16.57}$$

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

$$P = \tau_D D + \tau_E E + \tau_W (1 - \rho) W. \tag{16.58}$$

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

$$\rho = (\kappa - 1)\frac{1 - \tau_E}{1 - \tau_D}\frac{E}{W}. \tag{16.59}$$

Inserting this expression into equation (16.58) gives us for the deposit insurance premium

$$P = \left(\tau_E + (\kappa - 1)\frac{1 - \tau_E}{1 - \tau_D}(\tau_D - \tau_W)\right) E + \tau_W W. \tag{16.60}$$

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

$$\begin{aligned} \Pi_D &= \pi_H (1 + r_D)(1 - \tau_D) D + \pi_L (1 + r_L)(1 - \tau_W)(1 - \rho) W \\ &\quad + \pi_L (1 + r_L) P \\ &= (\kappa - 1)(1 - \tau_E)((\pi_H - \pi_L)(1 + r_L) - \pi_H (r_L - r_D)) E \\ &\quad + \pi_L (1 + r_L)(W + \tau_E E). \end{aligned} \tag{16.61}$$

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 (16.58).

If depositors invest directly into the company, they obtain

$$\hat{\Pi}_D = \pi_L (1 + r_L)(1 - \tau_W) W. \tag{16.62}$$

Again, deposits are provided if $\hat{\Pi}_D \geq \Pi_L$, which solves for

$$r_L - r_D \leq \frac{(\kappa - 1)(\pi_H - \pi_L)(1 + r_L)(1 - \tau_E) E + \pi_L (\tau_W W + \tau_E E)}{\pi_H (\kappa - 1)(1 - \tau_E) E}. \tag{16.63}$$

The incentives for the bank are unchanged, merely the equity reduces to $(1 - \tau_E) E$, thus the constraint in equation (16.47) remains valid and combining this with equation (16.63), we obtain a viable solution is available if the leverage of the bank is below

$$\kappa \leq \kappa^{**} = \frac{\tau_W W + E}{(1 - \tau_E) E} \kappa^*, \tag{16.64}$$

where κ^* is the leverage ratio banks choose in the absence of deposit insurance, it was defined in equation (16.51). 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

$$\tau_E = \tau_D = 0, \tag{16.65}$$

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

$$\begin{aligned}(16.66) \qquad \tau_W &= \frac{\pi_H c - (\pi_H - \pi_L)^2 (1+r_L)}{\pi_L (\pi_H - \pi_L)(1+r_L)} \frac{E+W}{W} - \frac{E}{W} \\ &= \frac{1}{\kappa^*}\frac{E+W}{W} - \frac{E}{W}.\end{aligned}$$

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 (16.63), as an equality and inserting the leverage ratio κ^{**}, the fee $\tau_E = 0$ and τ_W from equation (16.66) and noting that $(\kappa^{**} - 1)E = D = W$, we get the deposit rate as

$$(16.67) \qquad 1 + r_D = \frac{\pi_L}{\pi_H}(1+r_L)(1+\tau_W).$$

This deposit rate in higher than without deposit insurance, as comparison with equation (16.52) 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 (16.58), 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 (16.64) 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

$$(16.68) \qquad L = (1-\tau_E)E + (1-\tau_D)D = (1-\tau_E)\overline{\kappa}E$$

if we insert for $D = \rho W$ and for ρ from equation (16.59) 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 (16.66) that the fee charged on non-deposits is given by

$$(16.69) \qquad \tau_W = \frac{\overline{\kappa}\pi_H c - (\pi_H - \pi_L)(1 + r_L)((\pi_H - \pi_L)\overline{\kappa} + \pi_L)}{\pi_L (\pi_H - \pi_L)(1 + r_L)} \frac{E}{W} > 0.$$

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

$$(16.70) \qquad (1 - \pi_H)(1 + r_D) D = \pi_L (1 + r_L) P.$$

Inserting all expressions for deposits D and deposit insurance premium P, this expression solves for

$$(16.71) \quad \frac{E}{W} = \frac{(1 - \tau_D)\tau_W \pi_L (1 + r_L)}{(1 - \pi_H)(1 + r_D) - \pi_L (1 + r_L)(\tau_D (1 - \tau_D) - \tau_W)} \frac{1}{\overline{\kappa} - 1}.$$

We can now insert for τ_W from equation (16.69) 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.

17

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 17.3 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 17.1, 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 17.2 we will therefore investigate the issue and management of this payment method.

17.1 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 17.1.1 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 17.1.2 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 17.1.3 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 17.1.4.

17.1.1 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

$$\Pi_D^i = \left(1 + r_D^i\right) D - c d_i D - p \lambda d_N D. \tag{17.1}$$

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

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}. \tag{17.2}$$

Similarly we get for the decision between banks i and bank k that

$$d_i \leq \hat{d}_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^k\right)}{2c}. \tag{17.3}$$

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

$$\Pi_B^i = \left(\pi \left(1 + r_L\right) - \left(1 + r_D^i\right)\right) \left(d_i + \hat{d}_i\right) D. \tag{17.4}$$

Banks choose the deposit rate optimally such that it maximizes their profits, which after inserting from equations (17.2) and (17.3) for the market share of depositors gives us the first order condition as

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial \left(1 + r_D^i\right)} &= -\left(1 + \frac{2\left(1 + r_D^i\right) - \left(1 + r_D^j\right) - \left(1 + r_D^k\right)}{2c}\right) D \\ &\quad + \left(\pi \left(1 + r_L\right) - \left(1 + r_D^i\right)\right) \frac{D}{c} = 0. \end{aligned}$$

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

$$1 + r_D = \pi (1 + r_L) - c. \tag{17.5}$$

Inserting this result, we obtain that the market shares of banks from equations (17.2) and (17.3) becomes $d_i = \hat{d}_i = \frac{1}{2}$ and we get from equstion (17.4) that

$$\Pi_B^N = cD. \tag{17.6}$$

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

$$\Pi_D^i = \left(1 + r_D^i\right) D - (1-\lambda) d_i cD - \lambda \min\{d_i, d_j, d_k\} cD - p\lambda c d_C D. \tag{17.7}$$

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

$$\begin{aligned} 1 + r_D &= \pi (1 + r_L) - (1-\lambda) c, \\ \Pi_B^C &= (1-\lambda) cD. \end{aligned} \tag{17.8}$$

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

$$\Pi_D^i = \left(1 + r_D^i\right) D - (1-\lambda) c d_i D - \lambda \min\{d_i, d_j\} cD - p\lambda c d_P D, \tag{17.9}$$

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

$$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}. \tag{17.10}$$

If bank k is not cooperating with the other two banks, then their depositors obtain profits as given in equation (17.1) 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 (17.9). As the cooperating banks are identical and we only consider symmetric equilibria such that $r_D^i = r_D^j$ and hence from equation (17.10) 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 $\min\{d_i, d_j\} = d_i$ and obtain that depositors access all services at their own bank and hence $\Pi_D^i = \left(1 + r_D^i\right) D - cd_iD - p\lambda cd_PD$. 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

$$d_i \leq \hat{d}_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^k\right) - p\lambda c\left(d_N - d_P\right)}{2c}. \tag{17.11}$$

The profits of the cooperating bank is then given by

$$\Pi_B^i = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^i\right) D\right)\left(d_i^* + \hat{d}_i^*\right). \tag{17.12}$$

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

$$\begin{aligned} 1 + r_D^i \quad &= \quad \frac{(2 - \lambda)\, c}{(3 - 2\lambda)\, c}\pi\left(1 + r_L\right) \\ &\quad - \frac{(1 - \lambda)\, c}{(3 - 2\lambda)\, c}\left(2c - \left(1 + r_D^k\right) + \frac{p\lambda c}{3}\right). \end{aligned} \tag{17.13}$$

For the non-cooperating bank we then have similarly their profits given by

$$\Pi_B^k = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^k\right) D\right)\left(\left(1 - \hat{d}_i^*\right) + \left(1 - \hat{d}_j^*\right)\right). \tag{17.14}$$

		Bank j	
		Cooperating	Not cooperating
Bank i	Cooperating	Π_B^P, Π_B^P, $\hat{\Pi}_B^P$	Π_B^N, Π_B^N, Π_B^N
	Not cooperating	Π_B^N, Π_B^N, Π_B^N	Π_B^N, Π_B^N, Π_B^N

(a) Bank k not cooperating

		Bank j	
		Cooperating	Not cooperating
Bank i	Cooperating	Π_B^C, Π_B^C, Π_B^C	Π_B^P, $\hat{\Pi}_B^P$, Π_B^P
	Not cooperating	$\hat{\Pi}_B^P$, Π_B^P, Π_B^P	Π_B^N, Π_B^N, Π_B^N

(b) Bank k cooperating

Figure 16: Cooperation game for access to cash services

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

$$(17.15) \qquad 1 + r_D^k = \frac{1}{2}\pi\left(1 + r_L\right) - \frac{1}{2}c + \frac{1}{2}\left(1 + r_D^i\right) + \frac{p\lambda c}{6}.$$

Combining equations (17.14) and (17.15), we solve the deposit rates as

$$(17.16) \qquad \begin{aligned} 1 + r_D^i &= \pi\left(1 + r_L\right) - \frac{1 + 4c + \frac{p\lambda c}{3}}{5 - 3\lambda}\left(1 - \lambda\right), \\ 1 + r_D^k &= \pi\left(1 + r_L\right) - \frac{1}{2}\frac{2c + \left(7c + 1\right)\left(1 - \lambda\right) - \frac{2}{3}\left(2 - \lambda\right)\lambda pc}{5 - 3\lambda}. \end{aligned}$$

Using this result, we can insert the deposit rates into equations (17.10) and (17.11) and then obtain the profits of the cooperating banks in equation (17.12) and of the non-cooperating bank in equation (17.15). 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 16, 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 17 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.

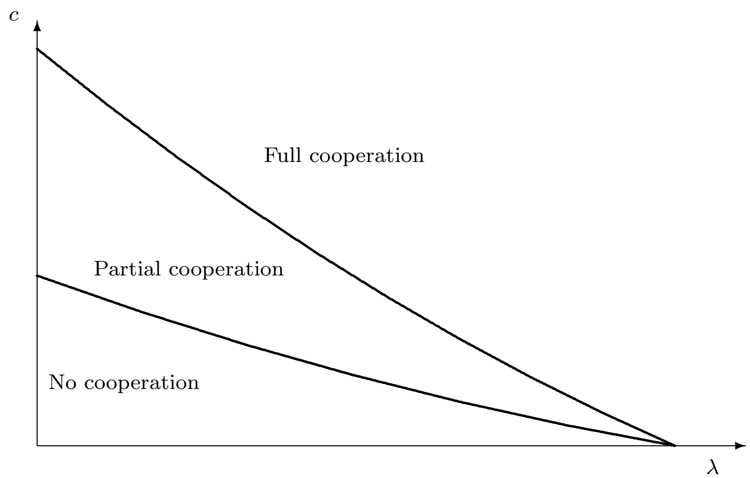


Figure 17: Equilibrium cooperation for cash services without access fees

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

$$\Pi_B^N = cD. \tag{17.17}$$

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 (17.7) by
(17.18)
$\Pi_D^i = \left(1 + r_D^i\right) D - (1-\lambda)\, d_i c D - \min\left\{\lambda c d_i, \lambda c d_j + f, \lambda c d_k + f\right\} D - p\lambda c d_C D.$

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

(17.19) $$d_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}.$$

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 $p\frac{3-d_i-\hat{d}_i}{3} fD$, 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

(17.20) $$\Pi_B^i = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^i\right) D\right)\left(d_i + \hat{d}_i\right) + p\frac{3 - d_i - \hat{d}_i}{3} fD.$$

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 (17.19), 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

(17.21) $$1 + r_D = \pi\left(1 + r_L\right) - c - \frac{pf}{3}.$$

Inserting this deposit rate into the bank profits of equation (17.20) gives us

$$\Pi_B^C = (c + pf)\, D > \Pi_B^N. \tag{17.22}$$

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 (17.10) and (17.11). 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

$$\begin{aligned} d_i \leq d_i^* &= \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c}, \\ d_i \leq \hat{d}_i^* &= \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^k\right) - p\lambda\left(d_N - d_P\right)}{2c}. \end{aligned} \tag{17.23}$$

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

$$d_P = \frac{5}{12} - \frac{2\lambda c + f}{6\lambda c} f. \tag{17.24}$$

The profits of the cooperating banks are then given by

$$\Pi_B^i = \left(\pi\left(1 + r_L\right) L - \left(1 + r_D^i\right) D\right)\left(d_i + \hat{d}_i\right) + p\frac{2\lambda c + f}{6\lambda c} f D, \tag{17.25}$$

where the last term accounts for the fee income. Using equation (17.23), we easily solve the first order condition $\frac{\partial \Pi_B^i}{\partial\left(1 + r_D^i\right)} = 0$ as

$$1 + r_D^i = \frac{2}{3}\pi\left(1 + r_L\right) - \frac{2}{3}c + \frac{1}{3}\left(1 + r_D^k\right) - \frac{p\lambda c}{3}\left(d_N - d_P\right). \tag{17.26}$$

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 (17.15)

$$(17.27) \qquad 1 + r_D^k = \frac{1}{2}\pi\left(1 + r_L\right) - \frac{1}{2}c + \frac{1}{2}\left(1 + r_D^i\right) + \frac{p\lambda c}{2}\left(d_N - d_P\right).$$

Solving equations (17.26) and (17.27) gives us the deposit rates of the cooperating and non-cooperating banks, respectively, as

$$(17.28) \qquad \begin{aligned} 1 + r_D^i &= \pi\left(1 + r_L\right) - c + \frac{4p\lambda c}{15}\left(d_N - d_P\right), \\ 1 + r_D^k &= \pi\left(1 + r_L\right) - c + \frac{p\lambda c}{5}\left(d_N - d_P\right). \end{aligned}$$

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 (17.28) into equation (17.23) and subsequently equation (17.25), we can show that $\Pi_B^P > \Pi_B^N$. Determining the equilibrium of the strategic game in figure 16, 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)

17.1.2 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

$$\begin{aligned}(17.29)\quad \Pi_B^i &= \pi\left(1+r_L\right)L-\left(1+r_D^i\right)\frac{D}{M}\\ &\quad -f\left(1-\frac{N_i}{N}\right)\frac{D}{M}+f\frac{N_i}{N}\frac{M-1}{M}D-CN_i\\ &= \left(\pi\left(1+r_L\right)-\left(1+r_D^i-f\right)\right)\frac{D}{M}+f\frac{N_i}{N}D-CN_i,\end{aligned}$$

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

$$(17.30)\qquad 1+r_D^i=\pi\left(1+r_L\right)-f.$$

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 (17.30), we can rewrite the profits of banks in equation (17.29) as

$$(17.31)\qquad \Pi_B^i=f\frac{N_i}{N}D-CN_i.$$

Noting that $N=\sum_{j=1}^{M}N_j$, we get the first order condition for the optimal number of cash access points as

$$(17.32)\qquad \frac{\partial\Pi_B^i}{\partial N_i}=f\frac{N-N_i}{N^2}D-C=0,$$

from which we easily obtain that

$$(17.33)\qquad N_i=N-\frac{C}{fD}N^2.$$

As all banks are equal, all banks will have the same number of access points such that $N=MN_i$. Multiplying equation (17.33) by M and solving for the total number of access points, N, we obtain

$$(17.34)\qquad N=\frac{M-1}{M}\frac{f}{C}D.$$

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 (17.31) that bank profits are given by

$$\Pi_B^i = \frac{f}{M^2} D \tag{17.35}$$

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

$$\begin{aligned} \Pi_D^i &= \left(1 + r_D^i\right) D - c d_i D - D \\ &= r_D^i D - \frac{M}{M-1} \frac{c}{f} C \\ &= \left(\pi \left(1 + r_L\right) - 1\right) D - fD - \frac{M}{M-1} \frac{c}{f} C, \end{aligned} \tag{17.36}$$

where the final equality merges when inserting for the deposit rate from equation (17.30). For deposits to be provided we require that $\Pi_D^i \geq 0$, which solves for

$$f \leq f^* = \frac{1}{2} \left(\pi \left(1 + r_L\right) - 1\right) + \sqrt{\frac{1}{4} \left(\pi \left(1 + r_L\right) - 1\right)^2 - \frac{M}{M-1} \frac{cC}{D}}. \tag{17.37}$$

Banks can charge higher interchange fees if they were to extract all surplus from depositors by adjusting the number of access points such that

$$\Pi_D^i = \left(\pi \left(1 + r_L\right) - 1\right) D - fD - c d_i D = 0. \tag{17.38}$$

Hence banks would have to provide

$$N = \frac{c}{2 \left(\pi \left(1 + r_L\right) - 1 - f\right)} \tag{17.39}$$

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 (17.31) become

$$\Pi_B^i = f \frac{D}{M} - \frac{cC}{2M \left(\pi \left(1 + r_L\right) - 1 - f\right)} \tag{17.40}$$

and hence the optimal interchange fee is given by the first order condition

$$\frac{\partial \Pi_B^i}{\partial f} = \frac{D}{M} - \frac{cC}{2M\left(\pi\left(1+r_L\right)-1-f\right)^2} = 0. \tag{17.41}$$

This solves for the optimal interchange fee to be

$$f^{**} = \pi\left(1+r_L\right) - 1 - \sqrt{\frac{cC}{2D}} \tag{17.42}$$

and inserting this into equation (17.39), the optimal number of access points is given by

$$N = \sqrt{\frac{cD}{2C}} \tag{17.43}$$

and bank profits are then

$$\Pi_B^i = \frac{f^{**}}{M}D - \frac{1}{M}\sqrt{\frac{cC}{2D}}. \tag{17.44}$$

Whether a bank would choose interchange fee f^* or interchange fee f^{**} will depend on the profits these generate, as determined by equation (17.35) and equation (17.44), 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

$$\Pi_W^i = cd_iD + CN = \frac{c}{2N}D + CN, \tag{17.45}$$

which leads to the first order condition for minimum costs that

$$\frac{\partial \Pi_W^i}{\partial N} = -\frac{c}{2N^2}D + C = 0, \tag{17.46}$$

from which we easily obtain that

$$N = \sqrt{\frac{cD}{2C}}. \tag{17.47}$$

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 (17.44) exceed those in equation (17.35).

Solving equation (17.42) for $\pi(1+r_L)-1$ and inserting the resulting expression into equation (17.37), we get the relationship between the two interchange fees as

$$f^* = \frac{1}{2}f^{**} + \frac{1}{2}\sqrt{\frac{cC}{2D}} + \sqrt{\frac{1}{4}\left(f^{**} + \sqrt{\frac{cC}{2D}}\right)^2 - \frac{M}{M-1}\frac{cC}{2D}}. \tag{17.48}$$

From this relationship we can easily determine that

$$\begin{aligned} \frac{\partial f^*}{\partial f^{**}} &= \frac{1}{2} + \frac{1}{4}\frac{f^{**} + \sqrt{\frac{cC}{2D}}}{\sqrt{\frac{1}{4}\left(f^{**} + \sqrt{\frac{cC}{2D}}\right)^2 - \frac{M}{M-1}\frac{cC}{2D}}} \\ &= \frac{1}{2} + \frac{1}{4}\frac{\pi(1+r_L)-1}{\sqrt{\frac{1}{4}\left(f^{**} + \sqrt{\frac{cC}{2D}}\right)^2 - \frac{M}{M-1}\frac{cC}{2D}}} \\ &> 1, \end{aligned} \tag{17.49}$$

where the final inequality arises when using equation from (17.42) 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 (17.44) exceed those in equation (17.35), this requires

$$f^* \leq Mf^{**} - \frac{M}{D}\sqrt{\frac{cC}{2D}}. \tag{17.50}$$

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 (17.49), the interchange fee f^* reduces more. This makes it more likely that the condition in equation (17.50) 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 (17.50) 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)

17.1.3 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

$$\Pi_D^i = \left(1 + r_D^i\right) D - (N + N^*)\, c d_i D - \hat{N} c D. \tag{17.51}$$

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

$$d_i = \hat{d}_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2c\,(N + N^*)}, \tag{17.52}$$

where $\hat{d}_i$ denotes the market share on the other side of the bank. Using that banks profits are given by

$$\begin{aligned} \Pi_B^i &= \pi\,(1 + r_L)\, L - \left(1 + r_D^i\right) D_i \\ &= \left(\pi\,(1 + r_L) - \left(1 + r_D^i\right)\right) \left(d_i + \hat{d}_i\right) D, \end{aligned} \tag{17.53}$$

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 (17.52) 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

$$1 + r_D^i = \frac{1}{2}\left(\pi\,(1 + r_L) - c\,(N + N^*) + \left(1 + r_D^j\right)\right). \tag{17.54}$$

Similarly we get for the other bank with profits $\Pi_B^j = \left(\pi\,(1 + r_L) - \left(1 + r_D^j\right)\right)\left((1 - d_j) + \left(1 - \hat{d}_j\right)\right) D$ that

$$1 + r_D^j = \frac{1}{2}\left(\pi\,(1 + r_L) - cN + \left(1 + r_D^i\right)\right). \tag{17.55}$$

Combining equations (17.54) and (17.55) we get that

$$1 + r_D^i = 1 + r_D^j = \pi\,(1 + r_L) - c\,(N + N^*)\,. \tag{17.56}$$

Inserting equations (17.52) and (17.56) into equation (17.53), we get the bank profits without remote access as

$$\Pi_B^B = c\,(N + N^*)\, D. \tag{17.57}$$

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

$$\Pi_D^i = \left(1 + r_D^i\right) D - Ncd_i D. \tag{17.58}$$

Again, depositors prefer using banks i over banks j if $\Pi_D^i \geq \Pi_D^j$, from which we obtain

$$d_i = \hat{d}_i \leq d_i^* = \frac{1}{2} + \frac{\left(1 + r_D^i\right) - \left(1 + r_D^j\right)}{2cN}. \tag{17.59}$$

The profits of banks remain given by equation (17.53) and after inserting from equation (17.59), the first order condition for the optimal deposit rate, $\frac{\partial \Pi_B^i}{\partial\left(1+r_D^i\right)} = 0$ solves for

$$1 + r_i^D = \frac{1}{2}\left(\pi\left(1 + r_L\right) - cN + \left(1 + r_D^j\right)\right) \tag{17.60}$$

and similarly for the other bank

$$1 + r_D^j = \frac{1}{2}\left(\pi\left(1 + r_L\right) - cN + \left(1 + r_D^i\right)\right) \tag{17.61}$$

such that when combining these two deposit rates we obtain that

$$1 + r_D^i = 1 + r_D^j = \pi\left(1 + r_L\right) - cN. \tag{17.62}$$

Inserting all results from equations (17.59) and (17.62) into the banks profits of equation (17.53), we obtain

$$\Pi_B^R = cND \leq \Pi_B^B. \tag{17.63}$$

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 (17.58) and for those with bank j not offering remote access are given by equation (17.51). Depositors prefer bank the bank offering remote access of $\Pi_D^i \geq \Pi_D^j$, from which we obtain that

$$(17.64) \qquad d_i = \hat{d}_i \leq d_i^* = \frac{\left(1+r_D^i\right) - \left(1+r_D^j\right) + c\left(N + N^* + \hat{N}\right)}{c\left(2N + N^*\right)}.$$

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

$$(17.65) \qquad \begin{aligned} 1 + r_D^i &= \pi\left(1+r_L\right) - \frac{1}{3}c\left(3N + 2N^* + \hat{N}\right), \\ 1 + r_D^j &= \pi\left(1+r_L\right) - \frac{1}{3}c\left(3N + N^* + \hat{N}\right). \end{aligned}$$

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

$$(17.66) \qquad \begin{aligned} \hat{\Pi}_B^R &= 2\frac{\left(N + \frac{2}{3}N^* + \frac{1}{3}\hat{N}\right)^2}{2N_B + N_O} cD, \\ \hat{\Pi}_B^B &= 2\frac{\left(N + \frac{1}{3}N^* - \frac{1}{3}\hat{N}\right)^2}{2N_B + N_O} cD. \end{aligned}$$

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

$$\hat{N} \leq \hat{N}^* = 3\left(\sqrt{\frac{1}{2}\left(2N + N^*\right)\left(N + N^*\right)} - \left(N + \frac{2}{3}N^*\right)\right). \tag{17.67}$$

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

$$\hat{N} \leq \hat{N}^{**} = 3\left(\left(N + \frac{1}{3}N^*\right) - \sqrt{\frac{1}{2}N\left(2N + N^*\right)}\right), \tag{17.68}$$

where we easily see that $\hat{N}^* \leq \hat{N}^{**}$.

Figure 18 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 (17.67) and (17.68) 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 (17.67) and (17.68) 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.

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.

		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

Figure 18: Strategic game to provide remote access

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)

17.1.4 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

$$\Pi_C^i = \left(p\pi\left((1+R) - \left(1+r_L^i\right)\right) - cd_i - f_i\right)L, \tag{17.69}$$

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

$$d_i \leq d_i^* = \frac{1}{2} + \frac{f_j - f_i + p\pi\left(\left(1+r_L^j\right) - \left(1+r_L^i\right)\right)}{2c}. \tag{17.70}$$

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

$$\Pi_B^i = \left(p\left(\pi\left(1+r_L^i\right) - (1+r_D)\right) + f_i\right)d_i^* L, \tag{17.71}$$

Inserting from equation (17.70) for the market share of the bank, we get the first order conditions for the optimal fee and loan rate as

$$\begin{aligned} \frac{\partial \Pi_B^i}{\partial f_i} &= d_i^* L - \frac{p\pi\left(1+r_L^i\right) - p(1+r_D) + f_i}{2c} L = 0 \\ \frac{\partial \Pi_B^i}{\partial\left(1+r_L^i\right)} &= p\pi d_i^* L - \frac{p\pi\left(1+r_L^i\right) - p(1+r_D) + f_i}{2c} p\pi L = 0, \end{aligned} \tag{17.72}$$

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

$$
\begin{aligned}
2cd_i^* - p\pi\left(1+r_L^i\right) - f_i + p\left(1+r_D\right) &= 0, \\
2c\left(1-d_i^*\right) - p\pi\left(1+r_L^j\right) - f_j + p\left(1+r_D\right) &= 0.
\end{aligned} \tag{17.73}
$$

Setting these two expressions equal and inserting for d_i^* from equation (17.70), we can obtain a relationship between the two account fees, which becomes

$$
f_j = f_i - p\pi\left(\left(1+r_L^j\right) - \left(1+r_L^i\right)\right). \tag{17.74}
$$

Using this expression in equation (17.70) for the market share of bank i and inserting this into the first line of the first order condition in equation (17.73), we obtain

$$
f_i = c - p\left(\pi\left(1+r_L^i\right) - \left(1+r_D\right)\right). \tag{17.75}
$$

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 (17.75) 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)

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.

17.2 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 17.2.1 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 17.2.2, focussing the interchange fee that banks are charged.

17.2.1 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

$$\begin{aligned} \Pi_D &= \text{Prob}(W_1 \geq P)(V - P) \\ &\quad + \text{Prob}(W_1 + W_2 \geq 2P)(V - P) \\ &= (2 - G(P) - G(2P - W_1))(V - P). \end{aligned} \tag{17.76}$$

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

$$\begin{aligned} \hat{\Pi}_D &= (V - P) + \text{Prob}(W_1 + W_2 \geq 2P)(V - P) \\ &= (2 - G(2P - W_1))(V - P). \end{aligned} \tag{17.77}$$

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

$$\begin{aligned} \Pi_C &= (1 - G(P)) P \\ &\quad + (1 - G(P))(1 - G(2P - W_1)) P \\ &\quad + G(P)(1 - G(P - W_1)) P, \\ \hat{\Pi}_C &= (1 - G(W^*))(P - F) \\ &\quad + (1 - G(W^*))(1 - G(2P - W_1)) P \\ &\quad + G(W^*)(1 - G(P - W_1)) P. \end{aligned} \tag{17.78}$$

Merchants will accept credit card payments if it is more profitable to do so, $\Pi_C \geq \hat{\Pi}_C$, which solves for

$$F \leq F^* = \frac{G(P) - G(W^*)}{1 - G(W^*)} (1 - G(2P - W_1) + G(2P - W_1)) P. \tag{17.79}$$

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

$$G(W|W_1) = Prob(W_1 + W_2 \leq W) = G(W - W_1) \tag{17.80}$$

and we then obtain

$$G(W|W_1 > W^*) = \int_W^{+\infty} G(W - W_1) \frac{dG(W_1)}{1 - G(W)} \tag{17.81}$$

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 $1 - G(P|W_1 > W^*)$; 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

$$\begin{aligned} \Pi_B &= (1 - G(W^*))\Big((F - P) + (1 - G(P|W_1 > W^*))P \\ &\quad + \int_{\min\{W^*;P\}}^{P} (W_1 + W_2)\, dG(W_1 + W_2|W_1 > W^*)\Big), \end{aligned} \tag{17.82}$$

Firstly we see that for the fee the bank charges merchants, we find

$$\frac{\partial \Pi_B}{\partial F} = (1 - G(W^*)) > 0 \tag{17.83}$$

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 (17.79) an equality and the fee charged is F^*. We then get after inserting from equation (17.79) for the fee, that the profits are changing with the threshold for issuing credit cards according to

$$\begin{aligned} \frac{\partial \Pi_B}{\partial W^*} &= g(W^*)\Big(-P + PG(P - W^*) \\ &\quad + \int_{\min\{W^*;P\}}^{P} (W_1 + W_2)\, dG(W_1 + W_2 - W^*)\Big). \end{aligned} \tag{17.84}$$

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 (17.84) becomes zero and $G(P - W^*) = 0$, hence

$$\frac{\partial \Pi_B}{\partial W^*} = -G(W^*) P < 0, \tag{17.85}$$

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 (17.84), at $W^* = P$ the expression is negative if $P > 0$, hence we decrease W^* until equation (17.84) 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 (17.79) we have

$$\frac{\partial F^*}{\partial W^*} = -g(W^*) \frac{1 - G(P)}{(1 - G(W^*))^2} P < 0, \tag{17.86}$$

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)

17.2.2 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

$$\Pi_B^D = F_D + F_I. \tag{17.87}$$

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

$$\Pi_B^C = F_C - F_I, \tag{17.88}$$

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

$$\frac{\partial T}{\partial F_I} = -\frac{\partial F_C}{\partial F_I} g(F_C)(1 - H(F_D)) - \frac{\partial F_D}{\partial F_I} h(F_D)(1 - G(F_C)) = 0. \tag{17.89}$$

Looking at equations (17.87) and (17.88), 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 (17.89) that we require

$$\frac{g(F_C)}{1 - G(F_C)} = \frac{h(F_D)}{1 - H(F_D)} \tag{17.90}$$

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 (17.87) and (17.88), 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

$$\frac{\partial \Pi}{\partial F_I} = \left(\Pi_B^D + \Pi_B^C\right) \frac{\partial T}{\partial F_I} + \left(\frac{\partial F_C}{\partial F_I} + \frac{\partial F_D}{\partial F_I}\right) T = 0. \tag{17.91}$$

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 (17.91) 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 (17.88), while increasing profits from depositors as equation (17.87) 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.

17.3 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 17.3.1 using a specific format, gross settlement, before then comparing it to an alternative net settlement mechanism in chapter 17.3.2, 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 17.3.3 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 17.3.4 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 17.3.5 where we discuss the incentives to delay payments in the face of a bank facing a liquidity shortage.

17.3.1 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

$$M_i^1 = \lambda_i R_i^0 \tag{17.92}$$

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,

$$R_i^1 = (1 - \lambda_i) R_i^0 + \lambda_j R_j^0. \tag{17.93}$$

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

$$M_i^2 = \min \left\{ R_i^1; M - \lambda_i R_i^0 \right\}. \tag{17.94}$$

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

$$R_i^2 = \max\left\{R_i^0 + \lambda_j R_j^0 - M; 0\right\} + \min\left\{M - \lambda_j R_j^0; R_j^1\right\}. \tag{17.95}$$

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

$$\Pi_B = f \min\left\{M; R_i^0 + \lambda_j R_j^0\right\} - cR_i^0 - \hat{c} \max\left\{M - R_i^0 - \lambda_j R_j^0; 0\right\}^2. \tag{17.96}$$

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 (17.96) can be rewritten as

$$\Pi_B = f \min\left\{M; 2R^0\right\} - cR^0 - \hat{c} \max\left\{M - 2R^0; 0\right\}^2. \tag{17.97}$$

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 (17.95) 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 (17.97) 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

$$R^0 = \frac{1}{2}M - \frac{c - 2f}{8\hat{c}}, \tag{17.98}$$

which then gives the bank profits of

$$\Pi_B^{**} = \frac{1}{2}\left(2f - c\right)M + \frac{\left(c - 2f\right)^2}{16\hat{c}}. \tag{17.99}$$

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 (17.95) 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 (17.98) 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)

17.3.2 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

$$\begin{aligned}(17.100)\ \Pi_D &= \lambda u\left(\left(1+r_D^1\right)D\right) \\ &\quad + (1-\lambda)\left(\gamma u\left(\left(1+r_D^2\right)D\right) + (1-\gamma)\,u\left(\left(1+\hat{r}_D^2\right)D\right)\right).\end{aligned}$$

The amount of deposits that are withdrawn after the first time period is given by

$$R = \lambda\left(1+r_D^1\right)D + (1-\lambda)(1-\gamma)\left(1+\hat{r}_D^2\right)D, \tag{17.101}$$

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

$$(1-\lambda)\gamma\left(1+r_D^2\right)D = \pi\left(1+r_L\right)(D-R). \tag{17.102}$$

We now maximize the utility of depositors, equation (17.100), 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 (17.101) and (17.102). Once these are determined, equation (17.102) 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

$$\begin{aligned}
\frac{\partial \Pi_D}{\partial (1+r_D^1) D} &= \lambda \frac{\partial u\left(\left(1+r_D^1\right) D\right)}{\partial (1+r_D^1) D} \\
&\quad - (\xi_1 - \xi_2 \pi (1+r_L)) \lambda = 0, \\
\frac{\partial \Pi_D}{\partial (1+\hat{r}_D^2) D} &= (1-\lambda)(1-\gamma) \frac{u\left(\left(1+\hat{r}_D^2\right) D\right)}{\partial (1+\hat{r}_D^2) D} \\
&\quad - (\xi_1 - \xi_2 \pi (1+r_L)) (1-\lambda)(1-\gamma) = 0,
\end{aligned} \tag{17.103}$$

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 (17.101) we get the required cash reserves as

$$R = (1-\gamma(1-\lambda))\left(1+r_D^1\right) D. \tag{17.104}$$

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

$$R^* = \lambda\left(1+r_D^1\right) D. \tag{17.105}$$

The objective function for depositors, equation (17.100) and the constraint from equation (17.102) 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 (17.104) and (17.105).

If we insert equation (17.104) into equation (17.101), and this in turn into the utility of depositors in equation (17.100), we easily get the first order condition for a maximizing this utility as

$$\begin{aligned}
\frac{\partial \Pi_D}{\partial (1+r_D^1) D} &= \lambda \frac{\partial u\left(\left(1+r_D^1\right) D\right)}{\partial (1+r_D^1) D} \\
&\quad -\lambda \pi (1+r_L) \frac{\partial u\left(\left(1+r_D^2\right) D\right)}{\partial (1+r_D^2) D} = 0.
\end{aligned} \tag{17.106}$$

Solving equations (17.102) and (17.104) for $\left(1+r_D^2\right) D$ and $\left(1+r_D^1\right) D$, respectively, we can obtain

$$\text{(17.107)}\quad \begin{aligned} \frac{\partial\left(1+r_D^1\right)D}{\partial\lambda} &= -\frac{R}{\lambda^2}+\frac{1}{\lambda}\frac{\partial R}{\partial\lambda}, \\ \frac{\partial\left(1+r_D^2\right)D}{\partial\lambda} &= \frac{\pi\left(1+r_L\right)}{\left(1-\lambda\right)^2}\left(D-R\right)-\frac{\pi\left(1+r_L\right)}{1-\lambda}\frac{\partial R}{\partial\lambda}. \end{aligned}$$

Using our first order condition from equation (17.106) we then obtain

$$\text{(17.108)}\quad \begin{aligned} \frac{\partial^2\Pi_D}{\partial\left(1+r_D^1\right)D\partial\lambda} &= \frac{\partial^2 u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D^2}\frac{\partial\left(1+r_D^1\right)}{\partial\lambda} \\ &\quad -\pi\left(1+r_L\right)\frac{\partial^2 u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D^2}\frac{\partial\left(1+r_D^2\right)D}{\partial\lambda} \\ &= 0, \end{aligned}$$

which after inserting from equation (17.107) can be solved for

$$\text{(17.109)}\quad \frac{\partial R}{\partial\lambda}=\frac{\frac{\partial^2 u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D^2}\frac{R}{\lambda^2}+\left(\pi\left(1+r_L\right)\right)^2\frac{\partial^2 u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D^2}\frac{D-R}{\left(1-\lambda\right)^2}}{\frac{\partial^2 u\left(\left(1+r_D^1\right)D\right)}{\partial\left(1+r_D^1\right)D^2}\frac{1}{\lambda^2}+\left(\pi\left(1+r_L\right)\right)^2\frac{\partial^2 u\left(\left(1+r_D^2\right)D\right)}{\partial\left(1+r_D^2\right)D^2}\frac{1}{\left(1-\lambda\right)^2}}>0,$$

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 (17.102) and (17.105) 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 (17.102) we then obtain the repayment to depositors retaining their deposits at their bank as

$$\left(1 + r_D^{2,ii}\right) D = \frac{\pi_i \left(1 + r_L\right)}{1 - \lambda} \left(D - R\right). \tag{17.110}$$

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, $(1 - \lambda) \left(1 + \hat{r}_D^{2,LH}\right) D$. 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, $(1 - \lambda) (1 - \gamma) \left(1 + \hat{r}_D^{2,HL}\right) D$, those not transferring deposits as the other bank is of high risk, $(1 - \lambda) \left(1 + \hat{r}_D^{2,HL}\right) D$, and those who transferred deposits, $(1 - \lambda) \left(1 + \hat{r}_D^{2,HL}\right) D$. 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

$$\begin{aligned}(17.111)\quad \pi_H(1+r_L)(D-R)+(1-\lambda)\left(1+\hat{r}_D^{2,LH}\right)D &= (1-\lambda)(1-\gamma)\left(1+\hat{r}_D^{2,HL}\right)D \\ &\quad +(1-\lambda)\gamma(1-\lambda)\left(1+\hat{r}_D^{2,HL}\right)D+(1-\lambda)\left(1+\hat{r}_D^{2,HL}\right)D \\ &= 2(1-\lambda)\left(1+\hat{r}_D^{2,HL}\right)D.\end{aligned}$$

For the high-risk bank, we get similarly that the resources available are from the loan they have provided, $\pi_L(1+r_L)(D-R)$, and the from those deposits that had to be transferred, $(1-\lambda)(1-\gamma)\left(1+\hat{r}_D^{2,HL}\right)D$. These resources are then used to repay depositors that have transferred deposits into this bank as they were required to, $(1-\lambda)(1-\gamma)\left(1+\hat{r}_D^{2,LH}\right)D$, and all of their remaining depositors leaving due to them being high risk, $(1-\lambda)\left(1+\hat{r}_D^{2,LH}\right)D$. Thus

$$\begin{aligned}(17.112)\quad \pi_L(1+r_L)(D-R)+(1-\lambda)(1-\gamma)\left(1+\hat{r}_D^{2,HL}\right)D &= (1-\lambda)\left(1+\hat{r}_D^{2,LH}\right)D+(1-\lambda)(1-\gamma)\left(1+\hat{r}_D^{2,LH}\right)D \\ &= (1-\lambda)(2-\gamma)\left(1+\hat{r}_D^{2,LH}\right)D.\end{aligned}$$

Solving equations (17.111) and (17.112) we get the repayments to depositors by the low-risk and high-risk banks, respectively, as

$$\begin{aligned}(17.113)\quad \left(1+\hat{r}_D^{2,HL}\right)D &= \frac{\pi_H(2-\gamma)+\pi_L}{(1-\lambda)(3-\gamma)}(1+r_L)(D-R), \\ \left(1+\hat{r}_D^{2,LH}\right)D &= \frac{\pi_H(1-\gamma)+2\pi_L}{(1-\lambda)(3-\gamma)}(1+r_L)(D-R).\end{aligned}$$

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

$$pu\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \geq pu\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right), \tag{17.114}$$

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

$$pu\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) > u\left(\left(1+r_D^1\right)D\right), \tag{17.115}$$

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

$$pu\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LL}\right)D\right) > pu\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LL}\right)D\right). \tag{17.116}$$

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 (17.116) is always fulfilled and deposits in banks exhibiting high risks will be transferred.

We similarly to equation (17.115) assume that

$$pu\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)+(1-p)\,u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right) > u(\left(1+r_D^1\right)D), \tag{17.117}$$

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 (17.117) is fulfilled, then the condition in equation (17.115) will be also be fulfilled. Hence the constraint in equation (17.117) 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

$$\begin{aligned} \Pi_D = \; & \lambda u\left(\left(1+r_D^1\right)D\right) \\ & + (1-\lambda)\left(p^2 u\left(\left(1+\hat{r}_D^{2,HH}\right)D\right)\right. \\ & + (1-p)^2\, u\left(\left(1+\hat{r}_D^{2,LL}\right)D\right) \\ & + p\,(1-p)\,u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right) \\ & \left. + (1-p)\,pu\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)\right), \end{aligned} \tag{17.118}$$

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

$$\begin{aligned} \Pi_D^* = \; & (1-p)\,u\left(\left(1+r_D^1\right)D\right) \\ & + p\left((1-\gamma\,(1-\lambda))\,u\left(\left(1+r_D^1\right)D\right)\right. \\ & \left. + (1-\lambda)\,\gamma u\left(\left(1+r_D^2\right)D\right)\right), \end{aligned} \tag{17.119}$$

where $\left(1+r_D^2\right)D$ is given from equation (17.102). If we now define

$$\Delta\Pi_D = \Pi_D^* - \Pi_D \tag{17.120}$$

as the difference in the utility of depositors in the gross and net settlement systems, we see that

$$
\begin{aligned}
(17.121)\quad \frac{\partial \Delta \Pi_D}{\partial \pi_L} &= -\frac{\partial \Pi_D}{\partial \pi_L} \\
&= -(1-\lambda)(1+r_L)(D-R) \\
&\quad \times \left((1-p)^2 \frac{\partial u\left(\left(1+\hat{r}_D^{2,LHL}\right)D\right)}{\partial\left(1+\hat{r}_D^{2,LHL}\right)D} \frac{1}{1-\lambda} \right. \\
&\quad + p(1-p) \frac{\partial u\left(\left(1+\hat{r}_D^{2,HL}\right)D\right)}{\partial\left(1+\hat{r}_D^{2,HL}\right)D} \frac{1}{(1-\lambda)(3-\gamma)} \\
&\quad \left. + p(1-p) \frac{\partial u\left(\left(1+\hat{r}_D^{2,LH}\right)D\right)}{\partial\left(1+\hat{r}_D^{2,LH}\right)D} \frac{2}{(1-\lambda)(2-\gamma)} \right) \\
&< 0,
\end{aligned}
$$

after inserting from equations (17.110) and (17.113). Similarly we can obtain that

$$
(17.122)\quad \begin{aligned}
\frac{\partial \Delta \Pi_D}{\partial \gamma} &> 0, \\
\frac{\partial \Delta \Pi_D}{\partial p} &< 0.
\end{aligned}
$$

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)

17.3.3 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

$$\Pi_B^* = \frac{1}{2}\lambda M - cM, \tag{17.123}$$

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

$$c \leq c^* = \frac{1}{2}\lambda. \tag{17.124}$$

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

$$\Pi_B^{**} = \frac{1}{2}pM - \hat{c}M, \tag{17.125}$$

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 (17.123) and (17.125) we easily obtain that

$$c \geq c^{**} = \hat{c} - \frac{1}{2} p (1 - \lambda) . \tag{17.126}$$

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

$$\hat{c} \leq \hat{c}^* = \frac{1}{2} p. \tag{17.127}$$

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

$$\Pi_B^{***} = \frac{1}{2} pM - \hat{c}M + \frac{1}{2} (1 - p) \lambda M - cM. \tag{17.128}$$

Using both mechanisms is preferred to using the direct settlement only if $\Pi_B^{***} \geq \Pi_B^*$, which after inserting from equations (17.123) and (17.128) becomes

$$\hat{c} \leq \hat{c}^{**} = \frac{1}{2} p (1 - \lambda) . \tag{17.129}$$

If this condition is fulfilled, we see from equation (17.126) 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 (17.125) and (17.128) becomes

$$c \leq c^{***} = \frac{1}{2} \lambda (1 - p) . \tag{17.130}$$

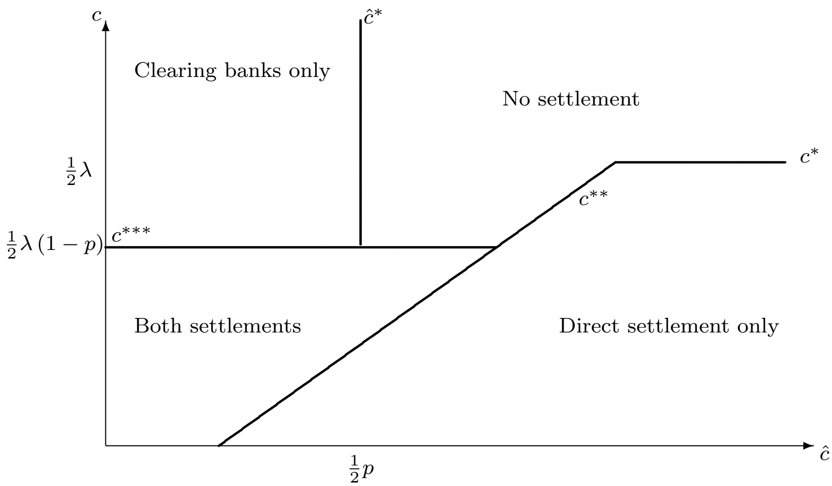


Figure 19: Settlement with clearing banks

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 (17.131) that

$$c \leq c^{****} = -\hat{c} + \frac{1}{2}\left(p + \lambda\left(1 - p\right)\right). \tag{17.131}$$

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.

Figure 19 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)

17.3.4 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

$$\Pi_B^{EE} = \hat{C} + (1 - p_1)\,\hat{C} + p_2\hat{C}. \quad (17.132)$$

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

$$\Pi_B^{LE} = C + (1 - p_1)\,\hat{C} + p_2\hat{C}. \tag{17.133}$$

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

$$\Pi_B^{EL} = 2\hat{C} + p_2\hat{C}. \tag{17.134}$$

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

$$\Pi_B^{LL} = C + \hat{C} + p_2\hat{C}. \tag{17.135}$$

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 20a, 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.

		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

Figure 20: Strategic payment delays

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

$$
\begin{aligned}
(17.136)\quad \Pi_B^{EE} &= (1-p_1)\hat{C} + (1-p_1)(1-p_2(1-p_2))\hat{C} \\
&\quad + p_2(1-p_2)\hat{C}, \\
\Pi_B^{EL} &= \hat{C} + (1-p_1)(1-p_2(1-p_2))\hat{C} + p_2(1-p_2)\hat{C}, \\
\Pi_B^{LE} &= C + (1-p_1)(1-p_2(1-p_2))\hat{C} + p_2(1-p_2)\hat{C}, \\
\Pi_B^{LL} &= C + (1-p_1)(1-p_2(1-p_2))\hat{C} + p_2(1-p_2)\hat{C}.
\end{aligned}
$$

The resulting strategic game is shown in figure 20b, 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)

17.3.5 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

		Bank 1 early	Bank 1 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	$\hat{C},\ 0$	$p\hat{C}+(1-p)\,(M+C),\ \frac{1}{2}pM$

Figure 21: Strategic interactions over payment delays

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 $\frac{1}{2}pM$.

Figure 21 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.

Interlude
Ethical considerations in banking

On first sight, economic models seem to not make any judgements on whether an action taken is just. A typical chain or arguments is that an action A induced a utility change Δu_i in individual i. If we find that for all N individuals together $\sum_{i=1}^{N} \Delta u_i \geq 0$, the action A is beneficial. Should there be an individual j for who $\Delta u_j < 0$, then at least 1 individual must have $\Delta u_i > 0$ and it would be possible to transfer utility from such individuals i to individual j to ensure that $\Delta u_j \geq 0$, while maintaining that $\Delta u_i \geq 0$; such a transfer usually takes the form of monetary compensation. After this transfer, all individuals are gaining from taking action A and are willing to participate. In many of the models we discussed, such transfers were not explicitly conducted, but a constraint imposed that all individuals must be better off. Such a constraint is then enforced by the use of a specific type of contract that ensures certain decisions are made by individuals that ensure action A improves the utility of all individuals.

Many different choices, apart from A, can be made and which action is chosen will depend on who makes this choice, as this individual would choose the action that increases its utility the most, while ensuring that no one is worse off. The aim of the decision-making process is to maximize the utility of the decision-maker, while ensuring that other individuals that have to participate in the action are willing to participate. This is ensured by requiring that their utility levels are not reducing. Despite all individuals increasing their utility level from taking action A, this action might be seen as unjust and morally not acceptable.

In ethics, a number of fundamental theories of just behaviour has been developed that can give additional insights into a moral assessment of any decisions that are made, in our context especially the decisions of banks with respect to who the lend to and whose deposits to accept, but also the compensation of employees. Making decisions that are just, we here call ethical behaviour.

Teleology One can argue that an action is just if it makes everyone better off. The end of the decision, an increase in utility, justifies the means, the action taken. This teleological approach to identifying just actions is the common technique used in economics to identify acceptable actions, and then the best action to choose. More specifically, economics is using utilitarianism in its assessment of actions and decisions; utilitarianism specifically suggests that actions should be maximizing the happiness (utility) of all individuals affected by an action.

Taking such a teleological approach is not without its critics and its conclusions on just actions have often been questioned. Consider the model from chapter 10.2.3, where companies were giving additional loans by a bank already having provided with loans in the past, despite the bank knowing that the company is unlikely to repay these new loans due to the financial distress they are in. Banks whoever, might be better off from making such a decision as companies could generate additional profits that can be used to repay this loan and some parts of the already existing loans, if it is successful. If the new investment is not successful, the bank would suffer higher losses, but these additional losses are outweighed by the possibility of the company being able to repay some of the existing loans. The company and its employees are also better off as the companies continues to operate and provide employment. Using teleological ethics, the end reducing the losses to the bank justifies the means, the provision of an additional loan to an otherwise bankrupt company. As we will see below, using alternative criteria that might include a fair assessment of the prospects of a company, this decision will no longer be seen as just.

Another criticism that might arise is that while teleology may ensure that every individual is better off and that contracts are set up such that this key property is ensured, how these gains are distributed might not be seen as just. It is often that one of the individuals involved, often the bank, can use these contracts to extract all benefits from other individuals, such as companies or depositors, and themselves gain the most from providing loans or accepting deposits. It could also be possible that due to competition between banks, companies might be able to extract all benefits and gain most, while banks make no net gain. Teleology in general, and utilitarianism specifically, do not provide criteria to decide whether such extraction of all surplus from other individuals can be seen as just.

Another concern when using teleology is that what might be individually rational in the sense of increasing the utility of an individual, might be a waste of resources. While using this approach can be used to justify the high compensation bankers obtain, be it due to their high productivity or the need to provide incentives to exert effort, the reliance on bonuses in this framework can also lead to morally questionable behaviour. If bonuses are determined subjectively on the basis of an assessment managers make regarding the individual contribution to profits, bankers might be tempted to spend resources and effort on presenting themselves in the best light. This is not only using resources, such as time, that could be used to assess companies applying for loans, but might also affect their decision-making as we will see in chapter 25. If only considering those individuals that are involved in the decision-making, the effect on other individuals can easily be overlooked.

A situation in which the effect of an action on individuals that are not part of the decision-making is ignored, is often referred to as an external effect. The ethics literature often refers to stakeholders as all those individuals, or groups of individuals, that are affected by any decision, whether they have been involved in the process or not. While teleology would require the assessment of an action to consider the effect of all stakeholders to assess whether it is a just action, the reality is that often only those are considered who are involved in the decision-making, and in some instances those who the decision is directed at. The wider range of stakeholders is commonly ignored. While this is not a limitation of the teleological approach itself, but more the way this is often implemented, particularly in economics, alternative ethical theories might provide an approach that puts the need to consider the wider impact of decisions at the forefront.

Deontology While teleology focussed on the outcome to assess whether an action is just, deontology investigates the action itself. Rather than assessing an action only if the end is just, hence an action is just if all individuals affected benefit, the means to obtain an end itself should be just. An action is considered just if it complies with a set of rules that have been judged to lead to just actions.

The most widely known set of such rules has been proposed by Immanuel Kant and is generally known as Kant's categorical imperative, Kant (1783). The tests proposed to judge an action can be formulated in different ways, one of which states that each individual has a duty to look after all other individuals. Implied in this formulation is that when taking an action, the decision-maker should consider the impact this action has on others and not take this action if they are affected negatively, or do not participate sufficiently in any gains. This criterion for a just action, and hence what is generally interpreted as ethical behaviour, addresses the limitation of the teleological approach that an individual could extract all surplus for their own benefit. While no one is worse off and the teleological criterion of a just action would be fulfilled, it would fall foul of this interpretation of Kant's categorical imperative.

An alternative formulation of Kant's categorical imperative proposes to assess that if everyone took this action or behaved in this way, would the outcome still be beneficial. It might be beneficial for a bank to provide loans only to very safe companies, but if every bank would act like this, no risky, but innovative investments would be financed by banks. Whether such an outcome is desirable and beneficial for companies, is doubtful. Hence, while providing only safe loans is ethical in the assessment of teleological criteria, it would not be an ethical behaviour using the deontological approach.

The so-called Golden Rule, which can be stated as 'do only what you others want to do to you', is not based on Kant's categorical imperative, but largely compatible with his suggested criterion. A direct implication of the Golden Rule in our context is that if a bank seeks to obtain surplus from providing loans or accepting deposits, borrowers and depositors should also obtain some surplus, thus banks would not extract all surplus for their won benefit and instead share the surplus generated. This would suggest that all stakeholders of an action are considered and their respective interests balanced against each other. Only if the allocation of surplus across all stakeholders is generally accepted, is an action deemed to be just and the behaviour ethical.

Apart from Kant's categorical imperative, John Rawls, Rawls (1971), focuses directly on the action taken. Considering the context of banking, he proposes that all individuals should have equal opportunities to achieve surplus from any decisions and not be excluded from it by their role in the economy. This would imply a negotiation between all individuals affected by an action, taking into account the specific circumstances of each individual as equally valid. All employees should have the same criteria applied to gain promotions or bonuses and this should be based on merit alone, not on their ability for self-promotion, unless this is part of the criterion, or their personal connections with more senior managers.

Furthermore, the least advantaged individual should also participate in any surplus. In terms of the compensation policy of banks this implies that all employees should be benefitting from the bonuses available, not only those employees who have directly contributed to the profits; other employees will have made their contribution more indirectly through cleaning offices, providing secretarial support, maintaining the IT infrastructure, or supplying food in the canteen, and should also be rewarded for their roles.

In addition, Rawls also required that all this is achieved with minimal restrictions on the decisions that individuals make, implying that regulatory interference by governments in achieving moral aims should be minimal, but also constraints imposed on employees by banks should be limited. Therefore, imposing ethical behaviour through restrictions is not in itself an ethical action, ethical behaviour should be coming out of the individuals own values and not be the result of constraints that are followed.

Virtue ethics While teleology assesses ethical behaviour from the outcome of actions, deontology focuses on whether the way actions are determined follows a just process. Virtue ethics goes one step further and judges the motivations for such actions. The outcomes of actions or whether they are actually just by some criteria are irrelevant for the assessment whether an action is classified as just and hence represents ethical behaviour. Virtues that should be reflected in any action go back to the Ancient Greek times of Aristotle and include honesty and modesty, amongst many others. In modern times such virtues have been extended to include the support for equality rights.

The same action can be seen as virtuous, and hence ethical, and unethical, depending on the context in which it is taken. Consider the competition between banks; it can be ethical if the banks compete with the aim of providing loans at lower costs to companies or pay higher interest on deposits. The same competition would be judged unethical in virtue ethics if the motivation for such competition was to harm a competitor, for example by driving him out of the market. Dishonest behaviour would be unethical in any circumstances, but we note that in the framework of the model, a teleological approach would judge fraud as ethical if it benefits the employee, neglecting here the external effects on the shareholders of the bank. Similarly, the award of bonuses can be unethical if the aim is not to reward contributions to the profits of the bank, but if they are the result of employees not criticising their managers or if this allows managers to promote their own goals by claiming to have managed their staff so successfully that they have earned large bonuses, and therefore he himself deserves being awarded a large bonus.

Rights and duties As has become clear in the previous discussion, the different approaches to assessing ethical behaviour - teleology, deontology, and virtue ethics - can give rise to different results. Teleology seeks a desirable outcome, deontology a just cause, and virtue ethics assesses the right motives for the action taken. As the arbitrary examples in figure 22 show, it is possible to identify actions that meet all the criteria for ethical behaviour, but in many cases an action would be judged ethical by some criteria, but not by others. As the criteria for assessing ethical behaviour are not generally agreed and banks are exposed to many different stakeholders, with often conflicting ethical assessments, it will be easy to accuse banks of acting unethical by violating one of the criteria.

Banks are often criticised for their unethical behaviour and the behaviour that is criticised is changing according to the circumstances of the time. Banks have been criticised for not providing enough loans, especially to small businesses, thus stifling innovation and technological progress. Then at other times, there has been criticism that banks do not apply sufficiently strict criteria when deciding on loan applications, resulting in large losses to them, which then requires a contraction on loans in the future once losses from defaults accumulate. Taking risks by investing into credit derivatives before the financial crisis 2007/8 has also been criticised as having caused this crisis, while at the same time urging banks to become more profitable. Another frequent criticism is that banks are not paying sufficient interest on deposits, while at the same time criticising high loan rates. The reaction to such criticism is often that regulation is imposed on banks that induces banks to make different decisions. Regulatory changes that have been brought in were changes to capital requirements, the requirements in

Entries in **bold** indicate that the assessment suggests unethical behaviour.

Action	Teleology (desirable outcome)	Deontology (just cause)	Virtue ethics (just motive)
Obtain information on companies' risks	Deposits are safe , company obtains loan, banks make profits	Allow the company to obtain loans at low interest	Grow the economy through investment
Securitise loans	**Risks become concentrated in a small number of buyers and banks may require bail-outs**	Move risks from banks who do not want to take it	Allow risks to be distributed according to preferences
Lowering fees and initiating a price war	After banks merge economies of scale reduce costs and fees	Lower the costs by concentrating banks	**Cause losses to competitors to cause them to fail or merge with them at low costs**
Advise company in distress to merge with another company	Loan is repaid and company survives	**Bank seeks to secure its own loans, not the company's survival**	Ensure the company survives

Figure 22: Examples of conflicting ethical assessments

the area of open banking, or codes of conducts in the treatment of applicants for loans, borrowers, and depositors.

Such regulatory interventions are seeking to address what is seen at that time as unethical behaviour and thereby impose more ethical behaviour on banks through the requirement to comply with the imposed regulation. As I have pointed out above, this in itself might not be ethical behaviour as it imposes additional constraints on decisions individuals make. Especially virtue ethics can be very restrictive on what constitutes ethical behaviour as it imposes a strict set of norms on the motivation of decisions, that cannot easily be mitigated. In contrast, teleology allows to offset any losses to some individuals by sharing the surplus other individuals obtain, and is therefore more flexible in what constitutes ethical behaviour.

Whenever multiple goals and criteria should be met, compromises have to be found; this compromise could be classified as a balance of rights and duties. In the context of banking, there is the right of the bank to benefit from their expertise and knowledge by generating profits. This is covered by teleology as the bank benefits from these profits, but also deontology where generating profits to maintain the services in the future could be regarded as ethically desirable. Unless profits themselves are seen as unethical, virtue ethics would also allow for banks to make profits, although the size of these profits should not be too high to comply with the virtue of modesty. These rights, however, have to be balanced against the duties of banks, most notably that of providing them with loans to finance their investments and accepting deposits to allow depositors a return on their investment. With their customers as stakeholders, the teleological approach suggests that they also need to obtain benefits from borrowing or making deposits, which is also covered by a deontological approach and similarly virtue ethics would not object to customers benefitting. The ethical judgement comes into finding the right balance between banks and their customers, and other stakeholders, as appropriate. Ideally banks would seek to comply with all three approaches of assessing ethical behaviour, but as we have seen, while some criteria can easily be met, an action might fall short of another criterion and be judged as unethical by those applying this theory.

Ethical behaviour cannot be imposed by regulation, whether from politics or internal ethics codes, it must be internalised by employees and be applied as a matter of course.

Reading Reynolds & Newell (2011, Chs. 3, 5, 7-8)

References

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.

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.

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.

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.

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.

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.

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)

Cao, J. (2022). *The economics of banking.* Abingdon: Routledge.

Carletti, E., Leonello, A., & Marquez, R. (2023). Loan guarantees, bank underwriting policies and financial stability. *Journal of Financial Economcis*, *149*, 260-295.

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.

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.

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.

Detragiache, E., Garella, P., & Guiso, L. (2000). Multiple versus single banking relationships: Theory and evidence. *Journal of Finance*, *55*(3), 1133-1161.

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.

Donze, J., & Dubec, I. (2006). The role of interchange fees in atm networks. *International Journal of Industrial Organization*, *24*, 29-43.

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.

Faria-e-Castro, M., Paul, P., & Sánchez, J. M. (2024). Evergreening. *Journal of Financial Economics*, *153*, 103778.

Foote, E. (2014). Information asymmetries and spillover risk in settlement systems. *Journal of Banking & Finance*, *42*, 179-190.

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., & Rochet, J.-C. (2008). *Microeconomics of banking* (2nd ed.). Cambridge, MA: 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.

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.

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.

Hauswald, R., & Marquez, R. (2006). Competition and strategic information acquisition in credit markets. *Review of Financial Studies*, *19*, 967-1000.

He, Z., & Manela, A. (2016). Information acquisition in rumor-based bank runs. *Journal of Finance*, *71*(3), 1113-1158.

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.

Hu, Y., & Varas, F. (2021). A theory of zombie lending. *Journal of Finance*, *76*(4).

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.

Kant, I. (1783). *Prolegomena zu einer jeden künftigen metaphysik, die als wissenschaft wird auftreten können.* Riga: Johann Friedrich Hartknoch.

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. (2016). *The economics of banking.* London: Palgrave Macmillan.

Krause, A. (2022a). The optimal exclusion length of borrowers after default. *Economics Letters*, *220*, 110881.

Krause, A. (2022b). Straetic default and optimal audit resources with costly state verification. *Research in Economcis*, *76*, 413-421.

Leland, H. E., & Pyle, D. H. (1977). Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance*, *32*, 371-387.

Martin, A., Skeie, D., & von Thadden, E.-L. (2014). Repo runs. *Review of Financial Studies*, *27*(4), 957-989.

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. (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.

Morrison, A. D., & White, L. (2011). Deposit insurance and subsidized recapitalizations. *Journal of Banking & Finance*, *35*, 3400-3416.

Padilla, A. J., & Pagano, M. (2000). Sharing default information as a borrower discipline device. *European Economic Review*, *44*(10), 1951-1980.

Park, H., & Kahn, C. M. (2019). Collateral, rehypothecation, and efficiency. *Journal of Financial Intermediation*, *39*, 34-46.

Rawls, J. (1971). *A theory of justice.* Cambridge, MA: Harvard University Press.

Reynolds, J. N., & Newell, E. (2011). *Ethics in investment banking.* Basingstoke and New York, NY: Palgrave MacMillan.

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.

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance*, *49*, 1069-1087.

Shy, O., Stenbacka, R., & Yankov, V. (2016). Limited deposit insurance coverage and bank competition. *Journal of Banking & Finance*, *71*, 95-108.

Spulber, D. F. (1999). *Market microstructure: Intermediaries and the theory of the firm.* Cambridge: Cambridge University Press.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, *71*, 393-410.

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.

Vesala, T. (2007). Switching costs and relationship profits in bank lending. *Journal of Banking & Finance*, *31*, 477-493.

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.

Index

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.

Andreas Krause has completed his apprenticeship at the National-Bank AG in Essen, Germany in 1993, before commencing his degree in economics at the University of Fribourg, Switzerland, which he completed in 2000 with a doctorate on financial markets theory. Since then, he has worked at the University of Bath, Great Britain, teaching and researching in banking and finance. He has published a large number of journal articles and book chapters on a wide range of topics, including payment cards, the systemic risk of banks, interbank markets, and strategic default of borrowers, amongst many others.