ELSEVIER

Journal of Corporate Finance 9 (2003) 333–352

Journal of CORPORATE FINANCE

www.elsevier.com/locate/econbase

Firm performance and compensation structure: performance elasticities of average employee compensation

Bruce A. Rayton*

University of Bath School of Management, Claverton Down, Bath BA2 7AY, UK

Accepted 16 April 2002

Abstract

Agency costs are a cost of production, and firms that do a better job of minimizing these costs should exhibit better performance. This paper tests this hypothesis by calculating the performance elasticity of average employee hourly compensation for U.S. manufacturing firms. This elasticity indicates the degree of alignment between employee and shareholder objectives. The estimated elasticity is indistinguishable from zero in low performance firms, and it equals 0.193 in high performance firms. While it is difficult to know whether an elevated performance sensitivity *causes* better firm performance, clearly the best performers in manufacturing industries link average employee pay to performance.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Firm performance; Compensation structure; Performance elasticity; Average employee compensation

Agency costs, broadly defined as the costs of poorly aligned incentives, are a cost of production. As such, firms minimize agency costs subject to a variety of technological and institutional factors. Decreased agency costs result in increased firm performance, all else equal, and this relationship is the subject of an expanding literature on the way firms forge agreements and contracts to minimize agency costs. This literature questions the assumptions about contracting in traditional economic theory.

Contracting has little role in traditional economic theory: efficient contracts are simply assumed to exist. The neoclassical model assumes transaction costs are zero, and this means

* Tel.: +44-1225-383-922; fax: +44-1225-826-473.
E-mail address: B.Rayton@bath.ac.uk (B.A. Rayton).

0929-1199/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0929-1199(02)00017-2

that the form of contracts is irrelevant. Firms will continue to write contracts so long as gains from exchange exist, and most (if not all) of the observed "contracts" will be simple market exchanges. However, transaction costs are not zero. Otherwise-profitable exchanges will not occur when transaction costs are sufficiently high, and some exchanges will now take place outside the market. Resources will not all flow to their most valued uses, and economic performance suffers relative to the first-best solution.[1]

A well-written contract can produce significant savings in a world of positive transaction costs, and so we expect firms to write contracts that achieve desired results at minimum cost. Examination of the contracts of successful firms may reveal common features of those contracts. Ichniowski et al. (1997) vividly demonstrate the importance of labor contract design. These authors examine productivity data from 36 steel finishing lines, and conclude that the lines using "innovative" work practices have higher productivity when compared to lines using "traditional" work practices.[2] Blinder (1990) considers the issue from a more aggregated perspective. He argues that a 1% increase in the productivity of labor can have a large impact on performance because labor accounts for over 70% of manufacturing costs. It seems that the magnitude of labor costs makes these costs a likely place to search for common features of the contracts of successful firms.

Many kinds of labor contracts exist in the economy. Some of these contracts are formal while others are informal agreements. Most employment relationships involve a combination of these formal and informal arrangements, and documenting the arrangements even in one firm is a difficult task. Baker et al. (1994a,b) undertake this task using personnel data from a single firm, but accomplishing this type of analysis for any group of firms is currently impractical. Instead of examining data from one firm, this paper uses the approach of Rayton (1997, 1999, in press) to examine the stream of annual payments directed to the employees of a panel of firms. This method leaves the source of the link between pay and performance unexamined, and simply measures the magnitude of the pay–performance link.

Although documenting the sources of incentive pay is a valuable project, such documentation is not required to measure the collective result of incentive pay arrangements. It is possible to examine the link between the overall flow of payments to employees and the overall performance of the firm without detailed knowledge of the payment mechanism. When attempting to determine the degree of incentive alignment in firms, it does not matter if the payments to employees are based directly on firm performance, or if these payments are based on individual performance measures. The net effect of either regime is an observable link between payments to employees and firm performance. Furthermore, the incentives for employees to work in ways that maximize firm value will be an increasing function of the magnitude of this link regardless of the regime.

The results of this paper indicate that high performance firms link pay and performance, while there is no evidence of such a link in low performance firms. These results are

[1] Coase (1960), North (1990), Williamson (1985) and Eggertson (1990, chap. 5) all emphasize this point.

[2] Ichniowski et al. (1997) define traditional practices by the use of narrow job specifications, close supervision, hourly pay, and strict work rules. They treat the use of incentive pay, team structures, and flexible job assignments as examples of innovative work practices.

supported by an array of alternative specifications. High performance firms display performance elasticities of approximately 0.193. This implies that a doubling of firm value will eventually result in a 19.3% increase in the pay of the average employee. Clearly, the best performing manufacturing firms are characterized by relatively "sharp" incentive structures.

1. Why link pay to performance?

The average employee of the firm faces decisions on the job: Should I work hard or not work hard? Should I spend the extra time and effort required to use resources efficiently, or should I waste resources (paid for by my employer) to make my work easier? These agency problems are only important in a world of positive transaction costs. There are many dimensions of the employment relationship that are difficult to define in such a world, and this creates opportunities for wealth capture.[3]

The decisions employees ultimately make depend on the relative costs and benefits of available alternatives, and firms can change the structure of the decision problem by rewarding correct behavior. Employees will maximize firm value if they receive net increases in utility for such behavior, and the magnitude of the link between pay and performance is commonly interpreted as a measure of these incentives. All else equal, increases in the magnitude of the pay–performance link imply decreased agency costs.

Financial "carrots" are not the only way to motivate employees. For example, firms may choose to monitor employees, and threaten to dismiss employees who are not performing to expectations. Ultimately, we should expect firms to find the lowest cost methods of aligning incentives. Nalbantian and Schotter (1997) use an experimental approach to examine productivity levels under various group incentive schemes. They find that monitoring can be effective at maintaining employee effort levels, but only when the probability of monitoring is high. This implies that monitoring may be a costly option when compared with financial reward systems like profit-sharing, gain-sharing, and competitive teams.

Other authors examine the link between pay and performance for average employees. The profit-sharing literature finds significant links between firm performance and the structure of average employee compensation.[4] For example, Bhargava (1994) examines the impact of a profit-sharing dummy variable on firm performance. This is interesting, but Bhargava only knows of the existence of one particular type of incentive pay. Bhargava has data on the level of profit-sharing pay in each firm, but he has no indication of other incentive measures which may be in effect.

The executive compensation literature reverses the causation of the profit-sharing work, and this literature finds robust empirical indications of a positive link between the pay of CEOs and firm performance. These studies of executives assume that performance causes

[3] Barzel (1989).

[4] See Kruse (1993) for an excellent review of the profit-sharing literature.

pay, and they measure CEO claims to changes in firm value.[5] Pay may come in many forms, and some of these payments may link pay and performance explicitly, while others only make the connection in an implicit manner. Regardless of the source of the changes in pay, the empirical approach of the executive compensation literature allows the measurement of the pay–performance link.

Rayton (in press) examines the link between pay and market performance, and finds that average employee hourly compensation in U.S. manufacturing firms is characterized by a performance elasticity of approximately 0.1. This result implies that a doubling of firm value eventually leads to a 10% increase in pay. Put differently, average employees eventually receive increases in wages and salaries equal to approximately 4.1% of increases in firm value. Other authors find similar results. For example, Ang and Fatemi (1995) find a performance elasticity of 0.1 for firms in the United Kingdom, and Blanchflower et al. (1996) find a performance elasticity for U.S. manufacturing firms of 0.08.

Blanchflower et al. (1996) is an example of the rent-sharing literature. This literature focuses on the relative bargaining power of employees and firms. The rent-sharing studies typically employ accounting measures of profit, and these studies produce performance elasticity estimates that are similar to the results of Rayton (1995, 1999, in press). While the rent-sharing literature does predict a link between pay and performance, Rayton (1997) finds that the observed elasticities in large and small firms are inconsistent with the exclusive truth of the rent-sharing hypothesis. While it may be true that rent-sharing behavior is present in the economy, rent-sharing theories alone cannot explain the inverse relationship between firm size and the link between pay and performance reported in Rayton (1995, 1997, 1999, in press). Agency theories can explain this difference, and Rayton concludes that some portion of the link is the result of attempts to get the incentives of employees aligned with value maximization.

Smith and Watts (1992) find a positive relationship between firm growth and the use of incentive pay. They start from the premise that principal–agent models predict some sort of incentive pay arrangement for managers if managerial actions are not perfectly observable. Smith and Watts then argue that managerial actions are less readily observable if the firm has more investment opportunities due to increased difficulties of communicating information about investment options. Taken together, these ideas predict greater use of incentive pay as growth increases. By extension, increases in the complexity of the relationship between the work of non-managerial employees and firm performance should also be matched by an increased incidence of incentive pay arrangements.

Agency costs are smaller when employee and shareholder objectives are closely linked, and firms can foster such links through the use of explicit contracts, informal arrangements and social customs to tie the pay of employees to firm performance. Whatever the mechanisms chosen, these links between pay and performance foster employee effort. Controlling for all other factors, firms that do a better job of linking pay and performance will experience lower agency costs and higher performance than otherwise identical firms with mismatched incentives. This paper uses the approach and data from Rayton (1999) to determine if there are significant differences between performance elasticities in low and

[5] See Jensen and Murphy (1990), Coughlan and Schmidt (1985), and others.

high performance firms. Specifically: do high performance firms link pay more closely to performance than low performance firms?

2. Empirical model

This paper examines the link between growth in firm value and growth in pay. Coefficients in this simple model are elasticities, and the sum of the return coefficients ($\beta_0 + \beta_1$) is a measure of the alignment of shareholder and employee objectives. The regression equation is

$$\ln\left(\frac{w_{i,t}}{w_{i,t-1}}\right) = \psi_{\mathrm{SIC},t} D_{\mathrm{SIC},t} + \beta_0 r_{i,t} + \beta_1 r_{i,t-1} + \omega_{i,t}. \tag{1}$$

The dependent variable is the growth rate of average hourly pay, w, for the ith firm in year t. The primary dependent variables are historical growth rates of firm value. These growth rates, r, are defined as the rate of return to common stock. Jensen and Murphy (1990) multiply this measure with beginning-of-period firm value to obtain the level change in firm value. This paper uses growth rates because rates of change are preferable to level changes when firms differ in size.[6] The regression coefficients are directly interpretable as elasticities, and the lag structure allows 2 years for changes in firm value to alter employee pay.

The use of firm-level panel data allows the imposition of disaggregated fixed time effects. These fixed effects, imposed at the two-digit industry level, control for industry-specific changes in omitted variables over time. Examples of such changes include cost shocks, industry-specific changes in product demand, and industry-specific changes in market value. Joskow et al. (1993) and Kruse (1993), among others, find significant cross-industry and time series variation in pay–performance sensitivities. The use of disaggregated fixed time effects in this model means that only time-varying shocks entering the model below the two-digit industry level of disaggregation can bias parameter estimates. Regressions using four-digit industry-year fixed effects yield similar results.

2.1. Effects of output demand shocks

Output demand shocks can influence wages, and it is important that these effects do not drive the results of this paper. Increases in the demand for output generate increases in the derived demand for labor, and this can sometimes lead to increases in wages. Suppose the demand shock in question is a firm-specific increase in output demand. This shock will have no effect on wages if the firm operates in a perfectly competitive labor market. The labor supply curve facing each individual firm is horizontal, and so the firm can alter the number of workers it employs without having any effect on the market wage.

Suppose instead that the demand shock is not firm-specific, but instead manifests itself at the industry level. In other words, suppose there is a general increase in the demand for the

[6] There are greater opportunity costs associated with owning large firms. See Rayton (1995) for further explanation.

output of a particular industry. Now all firms in the industry experience an increase in the demand for labor, and the net effect of these changes is movement along the existing labor supply curve. The market reaches a new equilibrium at a higher wage level. The industry-year fixed effects in the empirical model control for just this kind of industry-specific wage change.

Lastly, what happens if we move away from a perfectly competitive labor market? Suppose, for example, that there are significant costs of hiring and firing employees. Now firms have an incentive to retain employees in the face of negative demand shocks and to work employees harder when faced with positive demand shocks. One symptom of this situation is the use of overtime. Employees are routinely paid 1.5 times their normal wage for overtime hours, and this generates increases in average hourly pay that may be unrelated to incentives. The observed increases in average pay are simply an optimal response to transaction costs in the labor market. The industry-year fixed effects help alleviate the influence of overtime, but only imperfectly. This paper uses industry-level data from the Bureau of Labor Statistics (BLS), detailed in Section 3.2, to control for the effects of overtime. The BLS controls for industry-specific fluctuations in hours and overtime. Although this control is imperfect, there are good reasons to believe that the use of overtime has strong industry-specific components.

2.2. Endogeneity of contemporaneous stock returns

Abowd (1989) documents the endogeneity of current labor costs and current value. He finds a dollar-for-dollar tradeoff between unexpected changes in collectively bargained labor costs and changes in the value of common stock. The salient observations here are that the fiscal-year-end stock market valuation represents firm value after payments to employees, and that rational stockholders will account for the existing contractual structure in the valuation of shares. The market will downwardly adjust firm valuation in response to unexpected performance increases if employment contracts use performance incentives. Consider an example of a firm that pays performance-based bonuses to employees that amount to the entire increase in firm performance. This is an example of “giving away the firm.” The stock market valuation of this firm would remain constant through time if there were no changes in any firm characteristics besides employee performance. Even so, there are significant incentives created by the bonus system. The market valuation of the firm would never reflect the increases in employee performance in any year because these increases would be matched dollar for dollar by increases in current labor costs.

This endogeneity is present in CEO pay regressions, but it can be safely ignored since CEO pay is such a small portion of total costs. The same cannot be said for total labor costs, and so this paper uses a two-stage least-squares model to correct for this endogeneity. The instruments are taken from an empirical model employed in Bhargava (1994). Bhargava uses changes in sales, historical firm performance, and industry-specific fixed time effects as the independent variables in a regression on changes in accounting profit.[7] This paper uses 3 years of sales growth, two lags of the returns to common stock,

[7] Bhargava (1994) also includes a binary variable indicating the presence or absence of a profit-sharing system. Analogous data is not available at this time for the firms in this sample.

and industry-specific fixed time effects to instrument current period common stock returns. Thus, the first stage regression equation is given by

$$r_{i,t} = \xi_{\mathrm{SIC},t} D_{\mathrm{SIC},t} + \alpha_0 s_{i,t} + \alpha_1 s_{i,t-1} + \alpha_2 s_{i,t-2} + \alpha_3 r_{i,t-1} + \alpha_4 r_{i,t-2} + \varepsilon_{i,t} \tag{2}$$

Lagged firm performance is included in Eq. (2) to allow for persistence in firm performance. Mueller (1990) discusses the nature of this persistence at length, but here it is sufficient to recognize that firm performance may exhibit autocorrelation. Sales growth is used as an instrument for common stock returns because it captures the revenue dimension of profitability without serious corruption by labor costs. Models of profitability routinely incorporate other variables, such as market share, industry concentration, R&D expenditure, and advertising intensity. Consistent with the approach of Bhargava, and with certain limitations of the data, these variables are omitted here. To the extent that some of these effects are stable over time, estimation of the model in changes controls for these influences on performance. Furthermore, the industry-specific fixed time effects will control for the influence of the industry-level factors.

Any paper using an instrumental variables technique is vulnerable to criticisms regarding the validity of the instruments selected. This paper is no exception, however, it is hoped that the selection of a vetted model for the first stage regression will blunt any speculation that these instruments have been creatively selected. The two-stage least-squares estimator has been shown in Monte Carlo studies to have excellent small-sample properties, and the estimates tend to be quite robust to the presence of multicolinearity and specification errors.[8]

3. Data

Data for this paper is drawn from the Standard and Poor's (S&P) Compustat database and the Bureau of Labor Statistics Current Employment Statistics program. Compustat reports a 20-year window of annual data for publicly traded firms, and the compilation of three editions of this data provides 22 years of annual data for over 6000 firms. Limitations on data availability within Compustat and the removal of outliers limit the final dataset to 2133 firm-years of data from 194 firms.

3.1. Construction of dependent variable

The dependent variable is constructed from the Compustat and BLS data. The construction is governed by

$$w_{i,t} = \frac{L_{i,t}}{N_{i,t}\left(H_{\mathrm{SIC},t} + \frac{O_{\mathrm{SIC},t}}{2}\right)} \tag{3}$$

where w, the average hourly labor cost, is based on the labor and related expenses variable, denoted here by L. The number of employees is given by N, and BLS measurements of average weekly total hours (H) and average weekly overtime hours (O) for production and non-supervisory workers to control compensation data for non-numeric changes in labor

[8] Kennedy (1992, p. 160).

use. Eq. (3) assumes that overtime hours are paid 1.5 times the base wage. The labor variable and the BLS data on hours and overtime are discussed in some detail below.

3.1.1. Compensation data

I use Compustat's labor and related expenditures variable to measure total compensation. Compustat also reports the number of employees for each firm, and this allows construction of per-employee compensation.

The labor variable represents the costs of employees' compensation and benefits allocated to continuing operations. This variable contains information about the magnitude of annual firm expenditures on wages, salaries, incentive compensation, profit sharing, payroll taxes, pension costs, and some other benefit plans. This variable represents a gross accounting cost of labor. The absence of detailed compensation information precludes the analysis of the incentives generated by different forms of compensation, but the labor variable allows the measurement of the overall link between firm performance and average employee pay.

Point estimates based on the labor variable will understate the degree of incentive alignment because firms have many other incentive alignment tools at their discretion. Most components of the labor variable are usually not tied directly to firm performance, and an insignificant estimate of the relationship between Compustat's narrow definition of compensation and firm performance does not rule out the conclusion that firms tie the pay of average employees to performance. However, any significant relation between Compustat's narrow definition of compensation and firm performance strengthens the argument that compensation contracts throughout the firm reward employees for good performance.

Reporting of the labor variable is purely voluntary, and only 12% of Compustat firms choose to report this variable during the sample period. This represents the most severe constraint on the size of the final dataset.[9] Firms that report the labor variable tend to report it each year, and I have discerned only one obvious criterion differentiating reporting and non-reporting firms. Reporting firms tend to be larger than non-reporting firms. For example, the median value for real total assets in the full Compustat database is approximately \$46 million.[10] The median value of total assets among those firms who report the labor variable is \$1.674 billion. Not only are firms in the Compustat database larger than average firms in the economy, but the firms admitted to the regression data are large in comparison with the average firm in Compustat. Selection bias is a potential problem, but Rayton (1995) finds no evidence of selection biases introduced by the decision to report the labor and related expenses variable.[11]

3.1.2. Total hours and overtime hours worked

The Current Employment Statistics Program provides data on employment, hours, and earnings from a broad sample of firms in considerable industrial detail. The Bureau of

[9] Twenty-nine percent of the observations otherwise eligible for admission to the regression sample do not report the labor and related expenses variable.

[10] All dollar values in this paper are reported in 1987 dollars.

[11] Rayton investigates selection bias using Heckman's two-stage method. This method uses the inverse Mill's ratio (calculated from a probit on the reporting status of the firms) as an explanatory variable in the two-stage least-squares procedure used to generate the baseline results. The inverse Mill's ratio is *never* found to be significant. This approach is suggested by Hsiao (1986, pp. 198–200) for use with panel data.

Labor Statistics (BLS) cooperates with state agencies to collect monthly data from a sample of establishments involved in all non-farm activities. This data is publicly available for most industries at the four-digit level, and is available for all industries at the two-digit level of disaggregation. The data is constructed from a mail survey of approximately 375,000 employer units with over 40% of total payroll employment. The sample contains about 300,000 employer units for the construction of hours and earnings data for production and non-supervisory workers in private, non-agricultural industries.

I use the measurements of average weekly hours and average weekly overtime hours for production and non-supervisory workers to control compensation data for non-numeric changes in the workforce. It is important to correct for changes in hours because firms change the workload of current employees over the business cycle. Failure to correct for changes in hours and overtime over the business cycle could induce spurious correlation between compensation expenses and firm value, both of which vary procyclically. An expanding economy increases both firm value and the overtime of existing employees. This can raise compensation expenditures for reasons unrelated to incentives.

I impute BLS industry numbers to all firms in that industry. When four-digit industry data is available I use it, and when it is not available I use the least-aggregated of the available data. The BLS data is available for all industries at the two-digit level, and so all corrections are made at or below this level of aggregation. This approach to controlling for changes in the number of hours worked is obviously imperfect, but the confidentiality of BLS data prohibits more detailed analysis. The BLS data serves as a proxy for firm-level data on annual changes in employee hours, and this proxy will be particularly effective if changes in hours and overtime have industry-specific determinants.

3.1.3. Unions

The BLS data cannot control for all changes in workforce composition. These other changes in the composition of the workforce will not bias the parameter estimates as long as the effects can be treated as mean zero measurement error, but there are reasons to think that this may not be the case. For example, when unionized firms lay off workers in recessions they often begin with the least experienced, and thus the lowest paid, employees. This induces a countercyclical component to average hourly pay. Countercyclical average pay introduces a negative bias to the point estimates because increases in average hourly pay accompany recessionary decreases in firm value. The data is insufficient to control for these kinds of non-numeric fluctuations in the labor force, but the use of industry-specific fixed time effects helps to alleviate these data problems. In any event, the biases and inefficiencies introduced by this countercyclical component of average hourly pay only make it harder to accept the hypothesis that pay is linked to performance in these firms.

There is some worry that heterogeneity in unionization patterns could influence the results of this paper. Unionized firms traditionally exhibit lower performance than non-union firms, and thus may be over-represented in the low performance class. If seniority-based layoff practices depress measured elasticities then we may observe a difference between the performance classes that is driven by the omission of unionization from the estimation. That said, it is unlikely that the extent of unionization differs substantially across performance classes. All of the firms in this sample are quite large. The median firm has 20,000 employees in the low performance group and 20,900 in the high performance group. Brown et al. (1990)

report that 66% of establishments with 1000 or more workers have unionized non-office workers. Also, 30% of employees are unionized in companies with 500 or more employees. Taking the first year of data for each company in the sample as the point of comparison, only one firm has fewer than 100 employees, and this firm is in the low performance group. The next-smallest firm in the low performance group has just under 1000 employees. There are five firms in the high performance group with fewer than 1000 employees.

Certainly not all of the firms in this sample are unionized, but it is safe to say that each of these firms is a visible target for union organizers, and it is likely that most firms in both performance groups are subject to union influence. It is unlikely that there is a significant difference between the extent of unionization in the low and high performance groups in this sample. There are undoubtedly differences in unionization by industry, but remember that the performance criterion is industry-specific. Half of the firms in each industry (roughly) are placed into the high performance group and the other half are placed in the low performance group. While the presence of unions may tend to depress measured elasticities overall, I feel that a union-based difference in the measured elasticities across the performance classes is unlikely.

3.2. Returns to common stock

The primary independent variables are growth rates of firm value. I define these growth rates, r, as the rate of return to common stock. Common stock returns are calculated as

$$r_{i,t} = \frac{p_{i,t} + d_{i,t}}{p_{i,t-1}} \tag{4}$$

where $p_{i,t}$ is the fiscal-year-end price of common stock and $d_{i,t}$ represents dividends per share as reported by Compustat. This represents the change in the value of a share of common stock during year t.

I specify the model in growth rates because the influences of time-invariant fixed effects vanish, and because rates of change are preferable to level changes when firms differ in size. The regression equation includes contemporaneous and lagged values of this growth rate. This lag structure allows current changes in compensation to depend on past information. Current pay may depend on past performance if there is some delay in the observation of performance, or if there is some structural factor requiring a delay. For example, salary levels set by an annual review mechanism cannot reflect performance changes until the following year. The lag structure used in this paper is consistent with Jensen and Murphy (1990), but Section 6.1 of this paper examines the effect of an alternate lag structure on the estimated performance elasticity.

The lag structure allows 2 years for returns to alter employee pay. Lagged performance is important because some commonly used forms of pay, e.g., wages and salaries, are not directly linked to performance. These forms of pay can only be adjusted ex post, and annual adjustments of these forms of pay would lead to a link between current pay and lagged performance. The regression coefficients are elasticities, and the sum of the coefficients on firm returns ($\beta_0 + \beta_1$) is a measure of the alignment between shareholder and employee objectives.

3.3. Other data issues

This paper uses an unbalanced panel because Compustat increased its coverage in the 1980s. Balancing requires either rejecting all firms not covered (or not reporting the relevant variables) in previous years, or restricting the length of the panel. I admit firms to the dataset that contribute at least five contiguous observations. This allows at least 2 years of data to work completely through the regression equations for each firm. This approach is consistent with Bhargava (1994). I remove the observations from the 1% tails of each regression variable. This protects against results driven by a few extreme observations. All regressions have been run with outlier cutoffs ranging from 0.5% to 3.0%. The results are not significantly changed by these variations in the severity of the outlier removal algorithm.[12] After the creation of lags and removal of outliers there are 198 surviving firms contributing over 2100 firm-years from 1974 to 1992.

4. Separation criterion

The separation into performance classes is obviously central to the results of this paper. Put simply, a firm is labeled as "high performance" if it exhibits returns to common stock in excess of the industry average more often than other firms in the same industry. The separation criterion, G_i, is constructed from an underlying indicator variable, $g_{i,t}$, where

$$g_{i,t} = \begin{cases} 1 & \text{if } r_{i,t} > r_{\mathrm{SIC},t}. \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The indicator equals one if the common stock returns of the ith firm exceed the median industry returns for a given year. I calculate these industry returns from the entire Compustat data file, and not only from those firms that report the labor and related expenses variable. This gives a better indication of average industry performance than an average calculated from the restricted sample because, as discussed in Section 3, the restricted sample contains larger firms than the general Compustat population. Even so, the calculation of the performance benchmarks from the restricted sample does not alter the results of this paper.

The criterion governing the final separation is based on the frequency with which individual firms trip the indicator ($g_{i,t}$), or

$$G_i = \frac{1}{T}\sum_{t=0}^{T} g_{i,t}. \tag{6}$$

I define firm i as, "high performance" if G_i is greater than the analogous indicators of at least half of all other firms. I use this particular dividing line because it separates the

[12] Estimates of the performance elasticity range from 0.016 to 0.054 for low performance firms (all insignificant), and the estimates for high performance firms range from 0.193 to 0.254 (all significant).

sample into two groups containing roughly equal numbers of firms, but the selection of alternative division points does not affect the results.

Basing the separation criterion on industry-specific returns generates a symmetric representation of each industry across the performance classes. There are reasons to expect differences in returns across industries, and this approach avoids the separation of the sample based exclusively on these industry-specific characteristics. Industry-specific differences in the returns to common stock cannot explain the differences in coefficient magnitudes because of the symmetry of the performance classes. This symmetry is particularly important when considering the influence of omitted variables on the estimates. These omitted variables cannot influence the relative magnitudes of the coefficient estimates unless these variables exert different influences on low and high performance firms within the same industry.

5. Descriptive statistics

Tables 1 and 2 summarize the means and medians of relevant variables. Analysis of these descriptive statistics illustrates some interesting features of the data. Application of the separation criterion generates 102 low performance firms and 96 high performance firms. Each firm contributes an average of almost 11 observations to the regressions. By construction, there is a difference in common stock returns across classes. Low performance firms have a mean return of only 2.91%, while high performance firms average 5.49%. The median returns of the two groups display an even wider gap. Examination of Figs. 1 and 2 indicates that common stock returns are subject to a high degree of variability. The distributions of common stock returns for the two performance groups display considerable overlap. This implies that high performance firms have bad years and low performance firms have good years. Analysis of Table 3 also demonstrates this point. The correlation of common stock returns with the performance separation criterion is significantly positive, but it is only equal to 0.053. The correlation of returns with the underlying indicator variable that determines the eventual assignment of the separation criterion is also significantly positive, but this correlation is only 0.080. All of these pieces of evidence demonstrate that the use of the performance separation criteria is considerably different than a simple separation based on the level of returns in any given year.

The distribution of firm sizes is, not surprisingly, skewed regardless of the metric chosen. More surprising is that the relative magnitude of average firm size in the low and

Table 1
Low performance firms: selected firm characteristics, 1974–1992. Dollar figures reported in constant 1987 dollars (102 firms)

Characteristic	Mean	Median
Number of employees	37,959.99	20,000
Total assets (in millions)	$4866.44	$2007.38
Annual per-employee compensation	$32,936.42	$32,790.90
Hourly compensation	$14.67	$14.58
Change in logarithm of hourly pay	0.011	0.014
Firm value (in millions)	$3103.48	$1372.65
Returns to common stock (percent)	2.91	0.84

Table 2
High performance firms: selected firm characteristics, 1974–1992. Dollar figures reported in constant 1987 dollars (96 firms)

Characteristic	Mean	Median
Number of employees	35,825.67	20,900
Total assets (in millions)	$6758.26	$2193.96
Annual per-employee compensation	$31,399.38	$31,880.23
Hourly compensation	$13.82	$13.73
Change in logarithm of hourly pay	0.011	0.014
Firm value (in millions)	$4298.39	$1391.83
Returns to common stock (percent)	5.49	4.84

high performance classes is not invariant to the metric. Specifically, the mean number of employees indicates that low performance firms are larger than high performance firms. Conversely, the median number of employees, the mean and median of total assets, and the mean and median of firm value are all larger for high performance firms than for low performance firms. Another interesting fact is that employees in low performance firms receive approximately 4.9% more pay per year (6.2% more per hour) than employees in high performance firms. The levels of pay across the two classes are growing at approximately the same rate.

6. Empirical results

This section presents the empirical findings of the paper. Section 6.1 presents the primary results of the paper. These results are based on the performance separation

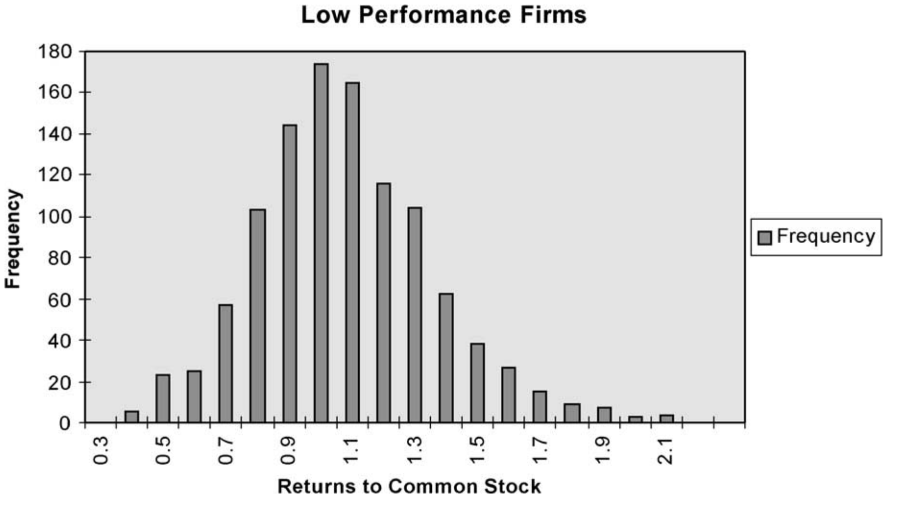

Fig. 1. Histogram of the returns to common stock for those firms whose returns exceed industry returns less than at least 50% of the other firms in the sample.

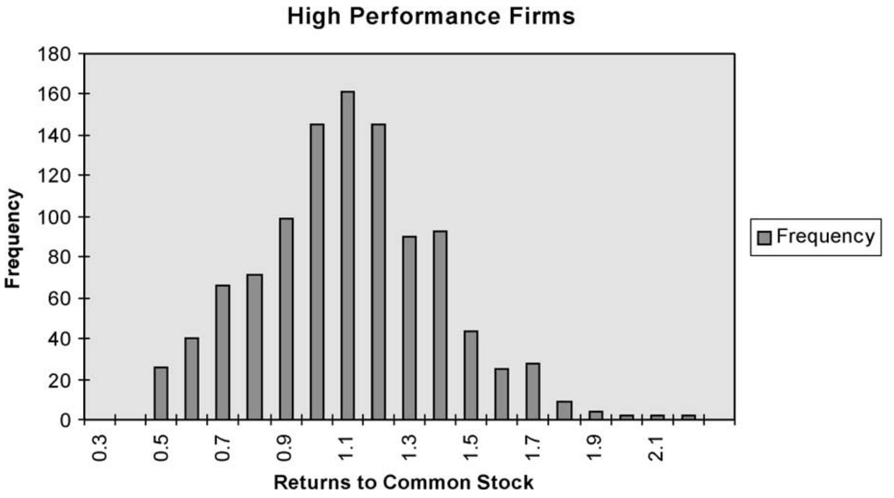

Fig. 2. Histogram of the returns to common stock for those firms whose returns exceed industry returns more than at least 50% of the other firms in the sample.

described in Section 4 of this paper. However, other definitions of success may yield different results. Analysis of data from groups classed by bond rating, an alternative definition of firm quality, appears in Section 6.2. Section 6.3 demonstrates the robustness of the empirical findings in Section 6.1 to many sensible changes in the specification of the regression equation.

6.1. Performance class

Table 4 presents regression results for the performance-separated regressions, and the results show that the compensation arrangements of high performance firms are characterized by considerably higher performance elasticities than the compensation arrangements of low performance firms. There are 102 firms defined as low performance, and these firms contribute 1081 observations to the estimation routine. These firms are compared with 96 high performance firms contributing 1052 observations.

Table 3
Correlations of performance dummy, D_i, the underlying performance indicator, G_i and the returns to common stock, $r_{i,t}$. P-values in parentheses

	D_i	G_i	$r_{i,t}$
D_i	1.0000 (0.0000)		
G_i	0.77782 (0.0001)	1.00000 (0.0000)	
$r_{i,t}$	0.05283 (0.0092)	0.07953 (0.0001)	1.00000 (0.0000)

$r_{i,t}$ = the return to common stock of firm i in year t.
G_i = the percentage of the years in which firm i exceeds industry returns.
D_i = the performance dummy variable. D_i equals 1 if G_i is greater than the median of G_i.

Table 4

Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	102	96
Sample size	1081	1052
R-squared	0.3160	0.2556
Return to common stock (t)	0.033920 (0.859)	0.150036 (2.629)
Return to common stock ($t-1$)	− 0.011029 (− 1.063)	0.042608 (2.906)
Estimated performance elasticity	0.022891 (0.540)	0.192644 (2.847)

T-statistic for difference of performance elasticities = 2.138.

Application of the empirical model yields some provocative results. The estimated performance-elasticity of per-employee compensation for low performance firms is equal to 0.023, and this estimate is not significantly different from zero. This contrasts with an elasticity estimate of 0.193 in high performance firms. This elasticity estimate is significant at the 1% level. Taken together, the two point estimates are significantly different from each other at the 5% level. The bulk of the difference between the two samples lies in the point estimate for the impact of contemporaneous returns on wage growth. Rayton (1995) reports performance elasticities for low and high performance firms of 0.08 and 0.141. These results are obtained by the inclusion of two lags of market value growth instead of the approach used here of including a single lag of the returns to common stock. Table 5 reports the results of the addition of an additional lag of returns to the current model, and this allows a more direct comparison with the results of Rayton (1995). The inclusion of the additional lag also demonstrates the robustness of the primary estimates to the addition of another lag of returns. The elasticity estimate for the low performance firms more than doubles, but it remains insignificant. The estimated elasticity for the high performance firms also increases, but only by 36% of the standard error of the elasticity from the original specification.

Table 5

Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	102	96
Sample size	1081	1052
R-squared	0.3135	0.2480
Return to common stock (t)	0.047262 (1.126)	0.163550 (2.598)
Return to common stock ($t-1$)	− 0.010340 (− 0.987)	0.046322 (2.835)
Return to common stock ($t-2$)	0.011162 (0.996)	0.007118 (0.569)
Estimated performance elasticity	0.048084 (0.968)	0.216990 (2.669)

T-statistic for difference of performance elasticities = 1.784.

Table 6
Separation by bond rating: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	S&P < A	S&P ≥ A
Number of firms	39	65
Sample size	177	345
R-squared	0.5167	0.2176
Return to common stock (*t*)	0.001329 (0.013)	0.199561 (2.754)
Return to common stock (*t* − 1)	− 0.029931 (− 0.688)	0.025630 (1.095)
Estimated performance elasticity	− 0.028602 (− 0.208)	0.225191 (2.596)

T-statistic for difference of performance elasticities = 1.626.

6.2. Bond rating

Bond ratings are another measure of the success of a firm. Good firms will be better able to repay bonds, and thus their bond ratings should be relatively high. Only 104 of the firms in the regression sample above solicited S&P bond ratings during the sample period. I define firms as having a low bond rating if their S&P rating ever falls below A during the sample period. I define those firms whose S&P bond ratings remain at least equal to an A rating as highly rated firms. The resulting sample includes only 39 low rated firms and 65 high rated firms from 1985 through 1992.[13]

Table 6 demonstrates that there is no significant link between pay and performance in the low rated firms. The point estimate is − 0.029. Conversely, the elasticity estimate for the high rated firms is 0.225, and it is significantly different from zero at the 1% level. The *t*-statistic for the difference between the point estimates of the two sub-samples is 1.626. While the test statistic is not significant at conventional levels, these results certainly do not refute the earlier findings. The key information imparted by Table 6 is the similarity of the coefficient estimates to the findings in Section 6.1.

6.3. Specification changes

The empirical specification used to generate the results in Sections 6.1 and 6.2 is quite simple. This simplicity is a conscious decision to emulate the models used to examine the incentives created by CEO pay;[14] however, this simplicity belies a high degree of stability in the coefficient estimates. This section demonstrates the robustness of the findings to the inclusion of lagged dependent variables in the regressions, as well as to the inclusion of the growth in the value of the capital stock.

Table 7 summarizes the results of including a lagged dependent variable. Inclusion of this variable allows past pay increases to provide information about future pay increases. The results indicate that an increase in the growth rate of pay this year implies a relatively

[13] Compustat does not begin to report the S&P bond rating until 1985.
[14] See Jensen and Murphy (1990), Coughlan and Schmidt (1985), and others.

Table 7
Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	100	96
Sample size	1056	1016
R-squared	0.3558	0.2766
Return to common stock (*t*)	− 0.027006 (− 0.783)	0.156085 (2.350)
Return to common stock (*t* − 1)	− 0.021086 (− 2.104)	0.041427 (2.643)
Δln{Average Hourly Pay} (*t* − 1)	− 0.108322 (− 4.890)	− 0.129861 (− 4.224)
Estimated performance elasticity	− 0.048092 (− 1.291)	0.192644 (2.528)

T-statistic for difference of performance elasticities = 2.872.

small change in the growth rate of pay in the following year. This relationship holds true for both the low and high performance firms, and it is not a surprise given the arithmetic relationship between current and lagged growth in pay. The addition of the lagged dependent variable to the analysis only strengthens the previously reported empirical findings. The link between pay and performance in high performance firms is virtually unchanged at 0.193, and this is significantly different from zero at the 5% level. The elasticity estimate for the low performance firms is negative and still insignificant. The two elasticity estimates are significantly different from each other at the 1% level. Table 8 examines the effect of adding a second lag of the dependent variable on the results. Again, the performance elasticity estimate is virtually unchanged.

Even a simple Cobb–Douglas production function features an interaction between the market wage and the capital stock, and so Table 9 examines the impact of including the growth in the capital stock as a regressor. The capital stock is measured using Compustat's measure of the value (net of depreciation) of property, plant, and equipment. The point estimate for the performance elasticity in high performance firms is higher than in previous specifications, but this change to the specification of the regression equation does not alter

Table 8
Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	96	93
Sample size	1023	977
R-squared	0.3776	0.2620
Return to common stock (*t*)	0.004339 (0.120)	0.173660 (2.175)
Return to common stock (*t* − 1)	− 0.019913 (− 1.938)	0.037817 (2.209)
Δln{Average Hourly Pay} (*t* − 1)	− 0.142171 (− 5.632)	− 0.125322 (− 3.791)
Δln{Average Hourly Pay} (*t* − 2)	− 0.066330 (− 3.895)	− 0.038214 (− 1.430)
Estimated performance elasticity	− 0.015574 (0.398)	0.211477 (2.265)

T-statistic for difference of performance elasticities = 2.278.

Table 9
Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	100	96
Sample size	1056	1016
R-squared	0.3357	0.2238
Return to common stock (*t*)	0.025254 (0.755)	0.230179 (2.936)
Return to common stock (*t* − 1)	− 0.015914 (− 1.543)	0.054639 (3.004)
Δln{Capital Stock} (*t*)	− 0.019210 (− 0.817)	− 0.061632 (− 2.204)
Estimated performance elasticity	0.009340 (0.259)	0.284818 (3.090)

T-statistic for difference of performance elasticities = 2.822.

the qualitative results reported earlier. The new specification increases the estimate of the performance elasticity for high performance firms to 0.286. The elasticity estimates are again significantly different from each other at the 1% level.

Tables 10 and 11 show the combined effects of the specification changes discussed in this section. The impact of the lagged dependent variables and the growth of the capital stock across the two sets of regressions is virtually identical. The elasticity estimate for low performance firms is not significantly different from zero, and the elasticity estimate for high performance firms is equal to approximately 0.218 in both regressions. The only substantial difference between the two tables is the level of significance for the difference of the performance elasticities. The model including both lags of the dependent variable as well as the growth of the capital stock produces a *t*-statistic of 2.086. The model including only one lag of the dependent variable and the growth of the capital stock produces a *t*-statistic of 2.769.

Taken together, these results support the hypothesis that high performance firms link pay more closely to performance than low performance firms. Although the estimated performance elasticity in high performance firms fluctuates with the specification of the model, the

Table 10
Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	100	96
Sample size	1056	1016
R-squared	0.3570	0.2701
Return to common stock (*t*)	− 0.005720 (− 0.171)	0.172974 (2.519)
Return to common stock (*t* − 1)	− 0.018433 (− 1.822)	0.044791 (2.772)
Δln{Average Hourly Pay} (*t* − 1)	− 0.106396 (− 4.811)	− 0.126001 − (3.993)
Δln{Capital Stock} (*t*)	− 0.024551 (− 0.127)	− 0.052690 − (2.468)
Estimated performance elasticity	− 0.024153 (− 0.670)	0.217765 (2.695)

T-statistic for difference of performance elasticities = 2.769.

Table 11
Separation by performance class: estimates of performance elasticity of per-employee hourly pay. Two-stage least-squares regressions. The dependent variable is the change in the natural logarithm of average employee hourly pay. All regressions include fixed time effects at the two-digit SIC level of disaggregation. *T*-statistics in parentheses

Independent variable	Low performance	High performance
Number of firms	96	93
Sample size	1023	977
R-squared	0.3750	0.2586
Return to common stock (t)	0.028273 (0.813)	0.179684 (2.238)
Return to common stock ($t-1$)	− 0.016987 (− 1.614)	0.039310 (2.281)
Δln{Average Hourly Pay} ($t-1$)	− 0.141429 (− 5.502)	− 0.123053 (− 3.646)
Δln{Average Hourly Pay} ($t-2$)	− 0.067081 (− 3.885)	− 0.036607 (− 1.353)
Δln{Capital Stock} (t)	− 0.022335 (− 1.109)	− 0.021586 (− 0.913)
Estimated performance elasticity	0.011286 (0.299)	0.218994 (2.332)

T-statistic for difference of performance elasticities = 2.086.

elasticity estimates for high performance firms are consistently greater than those for low performance firms.

7. Conclusion

Firms with a propensity to outperform their industry are also likely to link the pay of employees to performance. The magnitude of this link implies that the incentive systems in these high performance firms are relatively "sharp" by comparison with low performance firms in the same industry. Does this relationship imply causation? Could low performance firms become high performance firms by adopting different contracting schemes? An affirmative answer to this question would seem to imply some sort of systematic irrationality of firms, and I do not believe these results cannot support such a claim.

The results of this paper do demonstrate that the best performing firms have sharper links between pay and performance than their lower performance counterparts. These results are robust to changes in the lag structure of the regression equation, as well as to other sensible changes in the specification of the empirical model. Whether or not this performance differential is the result of the pay structure is a question best addressed with a significant amount of currently unavailable institutional detail, and such an effort is beyond the scope of this paper. The results presented here provide evidence that agency costs are a significant cost of production, and that firms that do a better job of minimizing these costs exhibit better performance. The precise sources of this performance differential should be an important focal point for future research.

Acknowledgements

This paper benefited from the helpful comments and suggestions of Bob Parks, Lee Benham, Fred Raines, Randy Nielsen, Andrew Oswald, Andrew Cooke and David Rose. I also received valuable suggestions from seminar participants at Warwick University, and the paper has benefited from the comments of an anonymous referee. Remaining errors are my responsibility.

References

Abowd, J.M., 1989. The effect of wage bargains on the stock market value of the firm. American Economic Review 79, 774–800.
Ang, J.S., Fatemi, A., 1995. The alignment of management and employee compensations: the case of U.K. firms. Journal of Multinational Financial Management 5, 73–85.
Baker, G.P., Gibbs, M., Holmstrom, B., 1994a. The internal economics of the firm: evidence from personnel data. Quarterly Journal of Economics 109, 881–919.
Baker, G.P., Gibbs, M., Holmstrom, B., 1994b. The wage policy of a firm. Quarterly Journal of Economics 109, 921–955.
Barzel, Y., 1989. Economic Analysis of Property Rights. Cambridge Univ. Press, Cambridge.
Bhargava, S., 1994. Profit sharing and the financial performance of companies: evidence from U.K. panel data. The Economic Journal 104, 1044–1056.
Blanchflower, A., Oswald, A., Sanfey, P., 1996. Quarterly Journal of Economics III, 227–252.
Blinder, A.S., 1990. Paying for Productivity: A Look at the Evidence. The Brookings Institution, Washington, DC.
Brown, C., Hamilton, J., Medoff, J., 1990. Employers Large and Small. Harvard University Press, Cambridge, MA.
Coase, R.H., 1960. The problem of social cost. Journal of Law and Economics 3, 1–44.
Coughlan, A.T., Schmidt, R.M., 1985. Executive compensation, management turnover, and firm performance. Journal of Accounting and Economics 7, 43–66.
Eggertsson, Thráinn, 1990. Economic Behavior and Institutions. Cambridge Univ. Press, New York.
Hsiao, C., 1986. Analysis of Panel Data. Cambridge Univ. Press, Cambridge.
Ichniowski, C., Shaw, K., Prennushi, G., 1997. The effects of human resource management practices on productivity: a study of steel finishing lines. American Economic Review 87, 291–313.
Jensen, M.C., Murphy, K.J., 1990. Performance pay and top management incentives. Journal of Political Economy 98, 225–264.
Joskow, P.L., Rose, N.L., Shepard, A., 1993. Regulatory constraints on CEO compensation. Brookings Papers on Economic Activity 1, 1–58.
Kennedy, P., 1992. A Guide to Econometrics, 3rd edn. MIT Press, Cambridge.
Kruse, D.L., 1993. Profit Sharing: Does it Make a Difference, W.E. Upjohn Institute for Employment Research, Kalamazoo.
Mueller, D., 1990. The Dynamics of Company Profits: An International Comparison. Cambridge Univ. Press, Cambridge.
Nalbantian, H.R., Schotter, A., 1997. Productivity under group incentives: an experimental study. American Economic Review 87, 314–341.
North, D.C., 1990. Institutions, Institutional Change and Economic Performance. Cambridge Univ. Press, Cambridge, MA.
Rayton, B.A., 1995. The CEO is Not the Only Employee of the Firm: An Empirical Investigation of Firm-Level Pay–Performance Relationships. Ph.D. dissertation. Washington University, St. Louis.
Rayton, B.A., 1997. Rent-sharing or incentives? Estimating the residual claim of average employees. Applied Economics Letters 4, 725–728.
Rayton, B.A., 1999. The size of employee stakeholding in large UK corporations. Managerial and Decision Economics 20, 259–266.
Rayton, B.A., 2001. The residual claim of rank and file workers. Journal of Corporate Finance, in press.
Smith Jr., C.W., Watts, R.L., 1992. The investment opportunity set and corporate financing, dividend, and compensation policies. Journal of Financial Economics 32, 263–292.
Williamson, O.E., 1985. The Economic Institutions of Capitalism. The Free Press, New York.