

Has Moral Hazard Become a More Important Factor in Managerial Compensation?

By GEORGE-LEVI GAYLE AND ROBERT A. MILLER*

We estimate a principal-agent model of moral hazard with longitudinal data on firms and managerial compensation over two disjoint periods spanning 60 years to investigate increased value and variability in managerial compensation. We find exogenous growth in firm size largely explains these secular trends in compensation. In our framework, exogenous firm size works through two channels. First, conflicts of interest between shareholders and managers are magnified in large firms, so optimal compensation plans are now more closely linked to insider wealth. Second, the market for managers has become more differentiated, increasing the premium paid to managers of large versus small firms. (JEL D82, L25, M12, M52)

Managers are paid to organize human resources in creative ways that add value to their firm. Because their activities are hard to monitor, managers are rarely paid for their inputs. Called moral hazard, this is the main reason why managers are not paid like most other professionals, at a rate more or less equalized across a large market for similarly skilled workers after adjusting for cost-of-living and amenity indices.¹ Executive compensation is tied instead to various indicators of managerial effort, such as the firm's performance. Linking a manager's compensation to the firm's performance requires him (or, in rare cases, her) to hold a substantial amount of insider wealth, assets that are sensitive to the firm's performance, such as its stocks and options.

The dramatic increase in both the level of CEO compensation and its sensitivity to firm performance over the last quarter century is widely documented by Brian Hall and Jeffrey Liebman (1998) and Kevin Murphy (1999). These studies show that, of all the components making up executive pay—including cash, bonuses, stock grants, and retirement benefits—the biggest increases have been in option grants. Thus, much of the increase in managerial compensation is attributable to increases in asset grants whose value is explicitly tied to the value of the firm. Since moral hazard explains why managerial compensation and firm performance should be connected, it is tempting to suggest that changes in the nature of moral hazard might have triggered these trends.

The theory of moral hazard provides a plausible transmission mechanism for connecting the compensation paid to a firm's executives with the returns on their firm's assets. There are two channels for inducing secular change in managerial compensation within the principal-agent

* Gayle: Tepper School of Business, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213 (e-mail: ggayle@andrew.cmu.edu); Miller: Tepper School of Business, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213 (e-mail: ramiller@andrew.cmu.edu). We would like to thank Martin Burda for his excellent research assistance and Vesna Prasnikar for her comments. We also like to thank seminar participants at SOLE 2006, Econometric Society summer meetings 2006, Carnegie Mellon University, Cornell University, Duke University, University of Essex, Georgia State University, London School of Economics, University of Maryland, New York University, University of Pennsylvania, and Princeton University. Gayle acknowledges the support of the Berkman Faculty Development Grant and the Andrew Mellon Research Fellowship, while Miller was supported by National Science Foundation grant SES0721098.

¹ The surveys by Canice Prendergast (1999), John Abowd and David Kaplan (1999), and Pierre Chiappori and Bernard Salanié (2000) review a growing empirical literature that analyzes executive compensation as a tool for regulating managerial decisions that are not directly monitored by shareholders.

paradigm. First, contracts reflect heterogeneity across firms, such as their size, capital-labor ratios, the sectors they belong to, and the dispersion of their financial returns. Consequently, changing the heterogeneity across firms induces changes in the aggregate level and variability of compensation. Second, the optimal contract is a function of the preferences and risk attitudes of managers. Changing those preferences also affects the probability distribution of compensation across executives. The purpose of our study is to estimate a model of moral hazard with data spanning a 60-year period in order to investigate how well these two channels explain secular changes in managerial compensation and to assess their relative importance.

Various metrics have been used to assess the importance of moral hazard. In their widely cited paper, Michael Jensen and Kevin Murphy (1990) define the importance of moral hazard by the ratio of absolute changes in managers' compensation to changes in firm equity value. In the case of large corporations, the ratio is a very small but increasing positive number. However, questions about the economic costs of moral hazard, such as those posed by Abowd and Kaplan (1999), cannot be answered with this metric. The three metrics we adopt from Mary Margiotta and Robert Miller (2000) are defined in Section I: the loss in firm value from not contracting with a manager to overcome the moral-hazard problem; the gain to managers from putting their personal goals ahead of their firm's; and a shadow value of a firm's willingness to pay for monitoring technology that would eliminate moral hazard. They are modeled within the principal-agent framework of Margiotta and Miller (2000) laid out in Section II, which is a direct descendent of Sanford Grossman and Oliver Hart (1983) and Drew Fudenberg, Bengt Holmström, and Paul Milgrom (1990). The optimal contract is given in Section III.

The data for our empirical analysis, described in Section IV, are drawn from two samples that collectively span approximately 60 years from 1944, with a 15-year break at 1978.² Previous research in accounting guided us on how to measure managerial compensation and firm returns. Appealing to a notion of current income equivalent or opportunity cost, Antle and Smith (1985, 1986) have persuasively argued that measures of the manager's income should include returns from insider wealth, rather than only salary, bonus, additions to retirement benefits, and (stock and option) grants. They also found that comparable results are obtained whether financial returns or an accounting measure, such as the firm's net income, is used. For these reasons, our definitions of compensation and returns use the inclusive measure of managerial compensation advocated by Antle and Smith, and focus on the financial returns of firms.

There is less agreement on how to benchmark firms' returns and managerial compensation, that is, relative to the performance of other assets and payments made to comparably placed executives. With regard to firm assets, since no single manager can affect the return on the market portfolio, returns to each firm should be purged of factor components that depend on returns from holding the market portfolio. Additionally, it is debatable whether managers can affect returns to all firms in the industry, depending on industry concentration and equilibrium advertising strategies. Our study shows that firm returns imputed net of the market portfolio factor are highly correlated with firm returns imputed net of both the market portfolio factor and an industry component factor. Consequently, our empirical results on the importance of moral hazard are robust whether industry components are subtracted to obtain a measure of abnormal returns or not. With regard to managerial compensation, we included a set of time-varying firm and industry characteristics to capture compensating differentials between executive positions.

² The first dataset, originally constructed by Robert Masson (1971) and later extended by Rick Antle and Abbie Smith (1985, 1986), covers the period 1944 to 1978. The second dataset, constructed from US Securities and Exchange Commission (SEC) records, covers the period 1993 to 2003.

The firms and their managers are selected from three industrial sectors—*aerospace, chemicals, and electronics*—broadly representative of all publicly traded corporations. The summary data on managerial compensation also match previous findings from other datasets; CEO average compensation has risen more rapidly than average wages, but less than its standard deviation. There are three noteworthy features about the firm data we investigate. First is the increase in size, as measured by assets, equity, and sales. Second, the capital-labor ratio has also increased; the firms in our samples experienced relatively moderate changes in employment, with declines in two out of the three sectors. A third feature is that, overall, the variance of financial returns net of the returns on the market portfolio has trended upward, but not uniformly across sectors, actually falling in the aerospace industry. Our empirical framework accommodates changes in the processes determining firm size and returns by separately estimating models of managerial compensation for the two samples and by parametrically allowing for the effects of changes on the contracts within each sample.

Identification of the moral-hazard model is discussed in Section V, the nonlinear structural estimator for a parameterization of our model is described in Section VI and explained in the Web Appendix (available at <http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.5.1740>), while our main results are presented in Section VII. We find that there are compensating differentials among the three industries we focus on that achieve both statistical and economic significance. However, the certainty equivalent of executive compensation for a firm with a given set of characteristics has increased at the same rate as national income, not faster. Our estimates show that managers have become a little less risk averse, slightly reducing the mean compensation required to accept risk. Consequently, we do not attribute the increased mean and variance of managerial compensation to changes in attitudes toward risk, or to increased rent from human capital in the management profession relative to other occupational skills.

We find that exogenous growth in firm size largely explains the secular trends in managerial compensation. Firm size has both direct and indirect effects on managerial compensation. The direct effect, the compensating differential paid to manage a larger firm, is unrelated to the moral hazard problem. The bigger the firm's workforce, the greater must be compensation to attract and retain more talented executives by persuading them to postpone retirement and decline employment elsewhere. Furthermore, although changes in the distribution of the firm workforce have been modest, the responsibility for managing them, as estimated by the compensating differential, have increased. This may reflect higher wages paid to workers and greater protection offered by labor law. The coefficient determining the marginal cost of effort of managing firm assets has also increased. In the old sample, managers were willing to take a reduction in their baseline pay to manage a larger firm, but in the new sample the CEO requires a premium to manage a firm with more assets. Overall, we find that the direct effects have not increased managerial compensation, but that the compensating differential for managing big firms has increased. This means that managers of small firms have experienced a decline relative to the increase in national income per capita.

The indirect effect of firm size on compensation stems from our empirical finding that the larger the firm, whether measured by assets or number of employees, the greater the conflict of interest between shareholders and managers. To align the incentives of managers with shareholders, managers tie their compensation to the firm's performance, exposing themselves to risk. Since managers are risk averse, this raises the expected compensation to offset the greater uncertainty. These changes in the mean and variance of compensation from the indirect effect are attributable to moral hazard. Overall, our estimates, therefore, attribute a much greater portion of average increases in executive compensation to the indirect effects of higher moral hazard costs from managing larger firms rather than to the direct effects of increased demand from larger firms for managerial services. Comparing the two samples, the marginal effect of adding

workers is lower in the more recent sample when firms in two of the three sectors had lower firm size on average. Similarly, the marginal effect of increasing assets on the conflict of interest between chief executive officers and shareholders is greater in the more recent sample where firm assets are on average higher. Both results corroborate the hypothesis that this conflict is a convex increasing cost of firm size, whether measured by the number of employees or by total firm assets. Our empirical results on firm size, both within each dataset and between sets, supports early work by Adolf Berle and Gardiner Means (1932) and Oliver Williamson (1967) about the size of firms and control of managers, as well as the argument that the size of organizations is limited by the scope of management.

The cost of moral hazard is also affected by the quality of abnormal financial returns to shareholders as a signal of managerial effort. In theory, increasing the volatility of abnormal returns does not necessarily weaken this signal, but in the three sectors we investigate, our results indicate that the signal about managerial effort deteriorates when abnormal returns become more dispersed. Thus, in chemicals and electronics, where the firms' returns become more variable in the more recent dataset, the effects of increased firm size on managerial compensation are reinforced by the weaker signal shareholders receive about managerial performance. In aerospace, where returns have become less dispersed, the stronger signal that shareholders now receive reduces the variance in compensation for a given firm size. In contrast to the other two sectors, the average firm in aerospace has experienced growth in both assets and employment. These growth effects have more than offset the effects of receiving a clearer signal. Thus, changes in signal quality have played a smaller role than exogenous firm growth has played in explaining trends in managerial compensation.

Armed with our parameter estimates, the last parts of the paper answer the question posed in the paper's title: has moral hazard become a more important factor in managerial compensation? Our estimates imply that the loss from ignoring the moral hazard issue altogether and not providing the appropriate incentives to management has increased substantially, largely because more assets are at stake in each firm. The gains to managers from deviating from shareholder interests when they are not given incentives to align their actions with shareholder interests have not increased; there is little to support the notion that managers employed in a firm with a given set of characteristics now require more incentives than before to act in the interests of their shareholders. The composition of firms in the three sectors we investigate has changed, however, and the growth in average firm assets has been accompanied by more dissonance between managerial and firm objectives. This change is reflected in the sharply increased welfare cost of moral hazard, which is the reservation value of shareholders to rid the firm of moral hazard at the executive level and pay risk-averse managers their certainty equivalent wage. The higher welfare cost has risen more rapidly than managerial compensation, and is the main force driving increased average payouts as well as increased volatility.

I. Measuring Moral Hazard

Arguing that "agency theory remains the only viable candidate for the answer to the question 'How well does executive compensation work?'" Abowd and Kaplan (1999) recently posed six questions that need answering. Our measures, borrowed from Margiotta and Miller (2000), directly relate to three of them: "How much does executive compensation cost the firm?"; "How much is executive compensation worth to the recipient?"; and "What are the effects of executive compensation?" We characterize the importance of moral hazard three ways: the gross loss shareholders would incur (before accounting for managerial compensation) from the manager tending his own interests; the benefits accruing to the manager from tending his own interests instead of those of his shareholders; and how much the shareholders are willing to pay to eliminate the problem of moral hazard altogether.

The first measure, denoted τ_1 , is the expected gross output loss to the firm for switching from the distribution of abnormal returns for diligent work to the distribution for shirking, that is, the difference between the expected output to the plant from the manager pursuing the firm's goals versus his own, before netting out expected managerial compensation. Let v denote the value of the firm at the beginning of the period, and let x denote the firm's abnormal return realized at the end of the period. Following literary convention, we describe a manager who pursues the interests of the firm as working, and a manager who pursues his own interests, when compensation is independent of firm performance, as shirking:

$$(1) \quad \begin{aligned} \tau_1 &= E[x \mid \text{manager works}]v - E[x \mid \text{manager shirks}]v \\ &= -E[x \mid \text{manager shirks}]v, \end{aligned}$$

where the second equality exploits the identity that the expected value of abnormal returns is zero when the manager is pursuing the interests of the firm.

The second measure, τ_2 , is the nonpecuniary benefits to the manager from shirking, that is, pursuing his own goals within the firm. Let w_2 denote the manager's reservation wage to work under perfect monitoring or if there were no moral hazard problem, and let w_1 denote the manager's reservation wage to shirk. Then τ_2 , the compensating differential for these two activities, can be expressed as the difference:

$$(2) \quad \tau_2 = w_2 - w_1.$$

We also estimate the maximum amount shareholders are willing pay to eliminate the moral-hazard problem, the value of a perfect monitor. If managerial effort is observed by the shareholders, then the firm would pay the manager the fixed wage w_2 . However, if the managerial effort is not observed by the shareholders and shareholders want the manager to work diligently, then the manager is paid according to the optimal compensation schedule $w(x)$. The firm's willingness to pay to eliminate the moral-hazard problem, denoted τ_3 , is accordingly defined as

$$(3) \quad \tau_3 = E[w(x)] - w_2.$$

Against the output reduction from shirking, τ_1 , is the savings in managerial compensation coming from two terms, the shadow value of a perfect monitor, and the cost of inducing the manager to work diligently when a perfect monitor is removed. Subtracting from τ_1 the sum of τ_2 and τ_3 , we obtain the net income loss a firm would sustain from signing a shirking contract with a manager. This net amount represents the value of preventing the manager from undoing contracts that align his incentives with the firm, by dealing with a lender who does not recognize the folly of allowing the manager to insure himself against poor firm performance, and is unaware of public disclosure laws that require the manager to report his holdings of firm-related securities.

II. A Model

This section lays out the theoretical principal-agent framework on which our empirical analysis is based.³ At each time period t , there are three activities in which a person can be engaged:

³ For an introduction to the vast literature on hidden actions and moral hazard, see the recently published texts of Jean-Jacques Laffont and David Martimont (2002), Salanié (2005), or Patrick Bolton and Mathias Dewatripont (2005).

not employed by the firm, employed as a manager at the firm but pursuing interests different from those of the shareholders, and working as the firm manager in the shareholders' interests. Let $l_t \equiv (l_{0t}, l_{1t}, l_{2t})$ denote the three possible activities, where $l_{jt} \in \{0, 1\}$ is an indicator for choice $j \in \{1, 2, 3\}$ and

$$(4) \quad \sum_{j=0}^{j=2} l_{jt} = 1.$$

If $l_{0t} = 1$, we say that the manager is not engaged by the firm and this activity is publicly observed, $l_{1t} = 1$ denotes shirking, and $l_{2t} = 1$ denotes working diligently. While l_{0t} is common knowledge, the values of (l_{1t}, l_{2t}) are hidden from the shareholders. Apart from choosing his activity, the manager also chooses his consumption for the period. Let c_t denote the manager's consumption in period t .

We assume that preferences over consumption and work are parameterized by a utility function exhibiting absolute risk aversion that is additively separable over periods and multiplicatively separable with respect to consumption and work activity within periods. In the model we estimate, lifetime utility can be expressed as

$$(5) \quad -\sum_{t=0}^{\infty} \sum_{j=0}^3 \alpha_j \beta^t l_{jt} \exp(-\rho c_t),$$

where β is the constant subjective discount factor, α_j are utility parameters associated with setting $l_{jnt} = 1$, and ρ is the constant absolute risk aversion (CARA). We set $\alpha_0 = 1$ as a normalization, since behavior is invariant to linear transformation of the utility function under the independence axiom. We assume that $\alpha_2 > \alpha_1$, or that diligence is more distasteful than shirking. This assumption is the vehicle by which the manager's preferences are not aligned with shareholder interests. We are not suggesting that managers are inherently lazy, merely that their personal goals do not motivate them to maximize the value of the firm if their compensation is independent of the firm's performance. Finally, we require $\alpha_1 > 0$ to ensure utility is increasing in consumption.

An optimal contract, with shareholder-created incentives to induce diligent work, compels the manager to bear risk on only that part of the return whose probability distribution is affected by his actions. Assuming the manager is risk averse, his certainty equivalent for a risk-bearing security is less than the expected value of the security, so shareholders would diversify among themselves the firm securities whose returns are independent of the manager's activities, rather than use them to pay the manager. We define the abnormal returns of the firm as the residual component of returns that cannot be priced by aggregate factors the manager does not control. In an optimal contract, compensation to the manager might depend on this residual in order to provide him with appropriate incentives, but it should not depend on changes in stochastic factors that originate outside the firm, which in any event can be neutralized by adjustments within his wealth portfolio through the other stocks and bonds he holds.

More specifically, let v_t be the value of the firm at that point in time. Then the abnormal return attributable to the manager's actions is the residual

$$(6) \quad x_t \equiv \frac{w_t}{v_{t-1}} - \pi_t - z_t \gamma,$$

where π_t is the difference between the return on the market portfolio in period t and the return on the firm's stock, and $z_t \gamma$ is a linear combination of some risk factors, denoted z_t , that lead to systematic deviations between the expected return on the firm's shares and that on the market portfolio. This study assumes that x_t is a random variable that depends on the manager's effort

activity choice in the previous period but, conditional on (l_1, l_2) , is independently and identically distributed across both firms and periods. Given $l_{jt} = 1$, for $j \in \{1, 2\}$, we denote the probability density function of x_t by $f_j(x_t)$, where $j \in \{1, 2\}$ denotes shirking and working, respectively. So, when managers are motivated to work diligently, differences in firm returns are attributed to exogenous forces modeled as risk-adjustment factors, and the excess return is a random variable drawn from a sector-specific probability distribution that depends on firm characteristics.

III. Contracting

Within this model, there are five parameters that might account for differences in executive compensation apart from the firm's abnormal return. They are the probability distribution of abnormal returns conditional on working, $f_2(x)$; the probability distribution of abnormal returns conditional on shirking, $f_1(x)$; the risk aversion parameter, ρ ; the nonpecuniary benefit of shirking versus working, captured by parameter ratio α_2/α_1 ; and the nonpecuniary benefit of working versus retiring or accepting employment outside the firm, α_2/α_0 . The shirking density, $f_1(x)$, determines τ_1 ; the three parameters, ρ and α_2/α_1 , are used to define τ_2 ; and τ_3 is fully determined by the two densities $f_1(x)$ and $f_2(x)$ plus ρ and α_2/α_1 , as we show below. All three measures of moral hazard require us to compute a counterfactual. In the case of τ_1 , we must impute the firm's value before compensation is paid if the manager shirks. The manager's utility from shirking is required for τ_2 and what the firm would have paid if there were no moral-hazard problem for τ_3 .

To compute these counterfactuals, we make the identifying behavioral assumption that shareholders contract with the manager to maximize the expected value of their firm. One possibility is that shareholders pay the manager enough to remain with the firm, but do not attempt to influence his activities on the job. In this case, shareholders minimize his expected compensation subject to a weak inequality constraint that induces the manager not to quit the firm (participation), a constraint that is satisfied with equality at the optimal contract for shirking,

$$(7) \quad (\alpha_0/\alpha_j)^{1/(b_t-1)} = E \left[\exp \left(\frac{-\rho w_t}{b_{t+1}} \right) \right],$$

where $j = 1$ and b_t is the price of a bond at t paying a unit of consumption per period forever. The solution to this problem is pay the manager a flat wage, which exactly compensates him for not taking a market or nonmarket alternative after allowing for the nonpecuniary attributes associated with managing the firm,

$$(8) \quad w_{1t} = \frac{b_{t+1}}{\rho(b_t-1)} \ln(\alpha_1/\alpha_0).$$

The other possibility is that shareholders not only pay the manager enough to deter him from quitting, but also create incentives for the manager to pursue the shareholders' interests rather than his own (incentive compatibility).⁴ When the manager is induced to work diligently, the participation constraint (by setting $j = 2$ in equation (7)) is satisfied and the incentive compatibility constraint is satisfied with the strict equality

⁴ This is a standard assumption in principal-agent models. See, for example, Grossman and Hart (1983). In conducting an empirical analysis of executive compensation that is explicitly based on a principal-agent model, our approach follows the work of John Garen (1994), Joseph Haubrich (1994), and Margiotta and Miller (2000), where the derivation of the optimal contract in our model can be found.

$$(9) \quad E \left\{ \exp \left(\frac{-\rho w_t}{b_{t+1}} \right) \left[\frac{f_1(x_t)}{f_2(x_t)} - (\alpha_2/\alpha_1)^{1/(b_t-1)} \right] \right\} = 0.$$

The optimal cost-minimizing contract that implements diligence behavior, solved by minimizing direct compensation subject to the participation and incentive-compatibility constraints, can then be written as

$$(10) \quad w_t = \frac{b_{t+1}}{\rho(b_t-1)} \ln(\alpha_2/\alpha_0) + \frac{b_{t+1}}{\rho} \ln [1 + \eta_t(\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta_t g(x_t)],$$

where $g(x_t) \equiv f_1(x_t)/f_2(x_t)$ is the likelihood ratio of the two probability density functions for shirking versus working, and η_t , the shadow value of relaxing the incentive-compatibility constraint, is the unique, strictly positive solution to the equation

$$(11) \quad \int [\eta(\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta g(x_t) + 1]^{-1} f_2(x) dx = 1.$$

In words, optimal compensation for diligent work is the sum of two pieces. If moral hazard was not a factor because managerial effort could be monitored, then a manager would be paid the flat rate:

$$(12) \quad w_{2t} = \frac{b_{t+1}}{\rho(b_t-1)} \ln(\alpha_2/\alpha_0).$$

Changes in the demand for the services of managers, from increased globalization and a greater reliance on general versus specific capital, affect α_0 , while changes in the nonpecuniary aspects of providing managerial services, such as more travel and more demanding staff, affect α_2 . In the absence of variables that affect only demand or supply, we cannot identify the relative importance of these factors. From data on the level of managerial compensation and its covariation with abnormal returns to firms, we can, however, empirically distinguish between the forces that determine the certainty equivalent that managers command in equilibrium, and the risk premium they are paid to hold insider wealth.

The second piece in the compensation equation under diligent work determines how it varies with abnormal returns through the slope of the function $g(x_t)$, a nonnegative function with $E[g(x_t)] = 1$ under $f_2(x_t)$. We interpret $g(x_t)$ as the signal shareholders receive about the manager's effort choice. If $g(x_t) = 0$, then shareholders conclude that the manager must have worked diligently. The greater the realized value of the signal, the less confident they are. If $g(x_t) = \infty$, they are sure the manager shirked. If $g(x_t) = 1$ for all x_t , the signal is useless, so compensation does not vary with returns in that case.

Having derived the two optimal contracts for shirking and diligence, shareholders select the most profitable, creating incentives for the manager to work diligently if and only if $\tau_1 > \tau_2$.

IV. Data

We used two sources of compensation and returns data to construct three samples for our empirical study. The first source of compensation was originally collected by Masson (1971) and later extended by Antle and Smith (1985, 1986). They contain compensation data on the

top three executives of 37 firms for the period 1944 through 1978. A detailed description of the first dataset can be obtained from Antle and Smith (1985). The primary source for the other two samples is the June 2004 version of the S&P ExecuComp database. This database follows the 2,610 firms in the S&P 500, Midcap, and Smallcap indices and contains information on at least the five highest-paid executives. We supplemented these data with firm-level data obtained from the S&P COMPUSTAT North America database and monthly stock-price data from the Center for Research in Security Prices (CRSP) database.

The first dataset contains compensation data for three industrial sectors, namely, aerospace, chemicals, and electronics. To ensure comparability of our results across the two time periods, we constructed two separate samples from the second dataset. The first sample includes only firms that belong to the three sectors in the first dataset classified according to the Global Industrial Classification Standard code (GICS). The second sample includes all firms in the S&P ExecuComp database. The first two samples allow us to directly compare the behavior of executive compensation across the two time periods controlling for aggregate conditions in the economy, and measures of the size and capital structure of firms. The third sample allows us to discern whether our restricted sample is broadly representative of the whole population of firms in the economy.

To facilitate comparisons across the three samples, we deleted observations on female executives in the two more recent samples and retained only the top three executives because there are no female executives in the older sample and no information on executives below the top three. We deleted observations with missing information, such as where compensation is reported for executives who had held the office for fewer than 50 weeks of a given year, and we also eliminated observations where the same executive is simultaneously listed with more than one company. This left us with 151 firms and 4,150 observations in the second sample and 1,517 firms and 82,578 observations in the third sample.

A. Abnormal Returns

We imputed x_t , gross abnormal returns to the firm in period t , as follows. First we computed the difference between the financial return on the individual stock and the return on the market portfolio. We then regressed this difference on a sector-specific constant and the time-varying factors, including GDP.

The sample means of the residual and their standard deviations are displayed in Table 1. The inclusion of a constant in the regression guarantees that the sample mean of the residual is numerically zero. All the estimated coefficients in the regression used to form this measure of abnormal returns proved significant, but the factors affect their dispersion by only a trivial amount. Table 1 shows that dispersion has increased in the chemicals and electronics sectors, but has declined in the aerospace sector. Dispersion in the unrestricted sample is higher than in the old sample of three sectors, but lower than in the new restricted sample.

B. Bond Prices

By definition, the price of a bond, b_{jt} , purchased at period t and maturing at date $t + j$, can be expressed as

$$(13) \quad b_{jt} = \sum_{s=0}^{t+j} \prod_{i=1}^s (1 + r_{it})^{-1},$$

where r_{it} is the (marginal) yield from lengthening the bond one period by extending the maturity date from $t + j$ to $t + j + 1$. Our bond price series comes from the Federal Reserve's Economic

TABLE 1—ABNORMAL RETURNS IN PERCENTAGE POINTS
(Standard deviations in parentheses)

Variable	Sector	Old	New restricted	New all
Abnormal returns	All	1.70E-8 (31.48)	4.17E-7 (53.43)	8.63E-8 (45.90)
	Aerospace	-2.50E-8 (42.24)	8.01E-8 (30.26)	
	Chemicals	3.82E-8 (23.89)	6.51E-8 (48.61)	
	Electronics	-1.68E-8 (41.16)	3.84E-7 (59.41)	

Research Data Base and is based on Treasury Bills with maturities of 1, 2, 3, 5, 7, 10, 20, and 30 years. We assumed the marginal annuitized yield rate for any bond maturing over 30 years is the same as the 30-year rate:

$$\begin{aligned}
 (14) \quad b_t &\equiv \sum_{s=1}^{\infty} \prod_{i=1}^s (1 + r_{it})^{-1} \\
 &= \sum_{s=1}^{t+30} \prod_{i=1}^s (1 + r_{it})^{-1} + \left[\prod_{i=1}^{30} (1 + r_{it})^{-1} \right] \sum_{s=31}^{\infty} (1 + r_{30,t})^{-s} \\
 &= \sum_{s=1}^{t+30} \prod_{i=1}^s (1 + r_{it})^{-1} + \frac{1}{r_{30,t}} \prod_{i=1}^{30} (1 + r_{it})^{-1}.
 \end{aligned}$$

For each date t , we imputed a yield curve using the data on newly issued bonds for various maturities, using a cubic spline for each date-maturity combination in the data, to obtain imputations \tilde{r}_{it} for each date t and for all $i \in \{1, \dots, 30\}$. The imputed series falls steeply from approximately \$120 to \$15 during the periods covered by our old dataset and is comparatively stable at the lower level during the periods covered by our new dataset; the steep decline in latter period is due to inflation.

C. Firm Characteristics

The characteristics of the manager's firm affect the nature of his responsibilities and the satisfaction he derives from his job. These characteristics are also relevant to the nonpecuniary satisfaction derived from pursuing his own goals within the firm.⁵ Table 2 is a cross-sectional summary of the characteristics we used in the estimation.

We partitioned each dataset by sector and focused on four indicators of size: sales, value of equity, total assets, and number of employees. These indicators convey some idea of the scope of managerial responsibilities. Sales have almost tripled in the three sectors, rising by less than a factor of two in chemicals, but by more than five in the other two sectors. Average sales per firm in these three sectors is about three-quarters of average sales in the unrestricted sample of

⁵ Several researchers have explored how differences in firms affect managerial compensation. For example, Peter Kostiuk (1990) and Scott Schaefer (1998) analyzed firm size and managerial compensation. Teresa John and Kose John (1993) investigated the capital structure of a firm and managerial compensation. Evidence provided by Rajeesh Aggarwal and Andrew Samwick (1999) shows that the volatility of abnormal returns is inversely related to the performance component of executive pay, as the theory of compensating differentials would predict.

TABLE 2—CROSS-SECTIONAL INFORMATION ON SECTORS
(All currency in millions of US\$ (2000); standard deviations in parentheses)

Variable	Sector	Old	New restricted	New all
Sales	All	1,243 (2,250)	3,028 (6,830)	4,168 (109,000)
	Aerospace	1,886 (3,236)	11,500 (14,900)	
	Chemicals	1,246 (2,018)	2,252 (2,091)	
	Electronics	319 (536)	2,469 (6,223)	
Value of equity	All	589 (1,034)	1,273 (2,863)	1,868 (4,648)
	Aerospace	391 (680)	3,132 (3,826)	
	Chemicals	677 (1,107)	800 (869)	
	Electronics	159 (365)	1,283 (3,096)	
Number of firms	All	37	151	1,517
	Aerospace	5	11	
	Chemicals	25	40	
	Electronics	7	100	
Number of employees	All	27,370 (28,850)	12,208 (26,676)	18,341 (46,960)
	Aerospace	49,920 (34,335)	58,139 (69,452)	
	Chemicals	23,537 (25,268)	8,351 (9,323)	
	Electronics	10,485 (7,664)	9,195 (18,266)	
Total assets	All	525 (924)	3,035 (6,550)	9,926 (40,300)
	Aerospace	726 (130)	10,600 (12,900)	
	Chemicals	548 (851)	2,385 (2,380)	
	Electronics	146 (233)	2,551 (6,311)	
Observations	All	1,797	3,260	82,578
	Aerospace	355	233	
	Chemicals	1,092	935	
	Electronics	252	2,092	

all listed firms. Similar changes in magnitude apply to equity value. However, all sectors have become much more capital intensive, as gauged by changes in total assets and employment. Assets have increased more than tenfold in aerospace and electronics, and by more than a factor of four in chemicals. Employment has declined in two out of the three sectors, most markedly in chemicals, where the average firm employs less than half the number of workers in the new

dataset compared to the old one. The size of the two datasets for the three sectors are very close in two of the sectors, but we have many more observations in the electronics sector, reflecting the growth of this sector over the last 30 years. Finally, although data on the three sectors is not a microcosm of the publicly listed corporations, it is quite representative: all the measures of size fall within a standard deviation of the sample mean for all the sectors.

D. Definition of Compensation

The cost to shareholders of employing a manager, called direct compensation, is the sum of salary and bonus, the value of restricted stocks and options granted, and value of retirement and long-term compensation schemes. The discounted sum of these direct compensation items measures the reduction in the firm's value from outlays to management. Total compensation to a manager is defined as direct compensation plus changes in wealth from holding firm options, and changes in wealth from holding firm stock. In order to compute the remaining two components in total compensation, one must take a stand on how managers would dispose of this wealth if it were not held in their firms' financial securities. We assume that the manager would hold a well-diversified portfolio instead, an implication of our model. When forming their portfolio of real and financial assets, managers recognize that part of the return from their firm-denominated securities should be attributed to aggregate factors, so they reduce their holdings of other stocks to neutralize those factors. Hence, the change in wealth from holding their firms' stock is the value of the stock at the beginning of the period multiplied by the abnormal return.

Our model implies that changes in wealth from holding firm options and changes in wealth from holding firm stock both have mean zero. Hence, direct and total compensation have the same expected value. Therefore, whether risk-neutral shareholders minimize expected total compensation or expected direct compensation is moot. However, changes in wealth from holding firm stock and options reflect the costs a manager incurs from not being able to fully diversify his wealth portfolio because of restrictions on stock and option sales. Consequently, managers care about total compensation, not direct compensation, because the former determines how their wealth changes from period to period when they optimally smooth their consumption over the life cycle and make optimal portfolio choices.

A third measure of compensation, called constrained compensation, is the sum of cash, bonus, the value of restricted stock and option grants, plus the change in the value of restricted stock and grant holdings. Constrained compensation exposes the manager to aggregate risk to the degree the firm's share price fluctuates with the market. Rational managers would neutralize their market risk by reducing their holdings of the market portfolio to compensate for the additional market risk that holding restricted stock entails. Suppose managers held no diversified stock after receiving their compensation, and were prevented from selling futures in the market portfolio (maturing when their firm-specific securities can be redeemed through sales). Then, we might conclude compensation is based on market returns if cash and bonus payments were not sufficiently countercyclical to offset the aggregate risk impounded within the manager's portfolio of his firm's financial securities. We are unaware of any evidence showing that the wealth portfolio of a manager is constrained by his own shareholders to hold more market risk than he voluntarily chooses. This explains why we followed Antle and Smith (1985, 1986), Hall and Liebman (1998), and Margiotta and Miller (2000) by using total compensation rather than constrained compensation in our study.

E. Total Compensation

Table 3 provides a cross-sectional summary of total compensation in the three samples by sector and rank. As can be seen from the top three rows, the means and standard deviations of

TABLE 3—CROSS-SECTIONAL INFORMATION ON TOTAL COMPENSATION
(In thousands of US\$ (2000); standard deviations in parentheses)

Rank	Sector	Old	New restricted	New all
All	All	528 (1,243)	4,121 (19,283)	2,319 (12,121)
CEO	All	729 (1,472)	6,109 (24,250)	5,320 (19,369)
Non-CEO	All	400 (1,026)	2,256 (12,729)	1,562 (9,303)
All	Aerospace	744 (1,140)	6,407 (20,689)	
CEO	Aerospace	950 (1,292)	11,664 (19,416)	
Non-CEO	Aerospace	624 (695)	1,997 (18,563)	
All	Chemicals	543 (1,348)	2,802 (9,555)	
CEO	Chemicals	718 (1,527)	3,673 (7,072)	
Non-CEO	Chemicals	401 (241)	477 (23,390)	
All	Electronics	370 (1,057)	4,501 (22,118)	
CEO	Electronics	457 (1,407)	5,325 (24,576)	
Non-CEO	Electronics	108 (61)	1,635 (18,810)	

the unrestricted sample lie between the corresponding numbers for the other two samples. The means and standard deviations are higher for the restricted samples, but there is large overlap between the empirical distributions characterizing the restricted and unrestricted samples.

Average total compensation for the three sectors in the new restricted sample is more than seven times larger than in the old one, with payments to the CEO rising by more than a factor of eight and payments to the other executives by less than six. The sector differences in the average growth rates are most pronounced for other executives, whose average compensation grew by less than 20 percent in the chemicals sector, but by fifteenfold in electronics. The sector differences are less pronounced at the CEO level, ranging between 5 times (in chemicals) to about 12 times (in aerospace and chemicals).

Changes in the average levels of compensation are considerably less significant than changes in their dispersion. The standard deviation increases by more than a factor of ten across all the subsamples, except in the chemicals sector.

F. Components of Compensation

Table 4 breaks out total compensation into its main components: salary and bonus, the value of restricted options granted, the value of restricted stock granted, changes in wealth from holding firm options, and changes in wealth from holding firm stock. The first three components collectively account for about 80 percent of the sum. While not contributing significantly to mean of total compensation, the latter two components account for much of its variability. The remaining unlisted components come from retirement and long-term compensation schemes.

TABLE 4—CROSS-SECTIONAL INFORMATION ON COMPONENTS OF COMPENSATION
(In thousands of US\$ (2000); standard deviations in parentheses)

Variable	Rank	Old	New restricted	New all
Salary and bonus	All	219 (114)	838 (1,066)	667 (905)
	CEO	261 (115)	1,037 (1,365)	1,127 (1,282)
	Non-CEO	179 (97)	640 (576)	552 (738)
Value of options granted	All	79 (338)	2,401 (13,225)	903 (3,753)
	CEO	111 (439)	3,402 (18,172)	1,782 (7,169)
	Non-CEO	51 (198)	1,401 (4,237)	681 (2,106)
Value of restricted stock granted	All	11 (95)	187 (1,633)	152 (936)
	CEO	8 (72)	242 (2,021)	298 (1,464)
	Non-CEO	13 (112)	133 (1,118)	115 (743)
Change in wealth from options held	All	5 (134)	785 (14,636)	281 (8,710)
	CEO	7 (167)	1,667 (17,078)	1,474 (13,567)
	Non-CEO	3 (94)	-76 (11,706)	-18 (6,939)
Change in wealth from stock held	All	-3 (439)	-40 (5,681)	125 (4,350)
	CEO	0.434 (479)	-14 (6,712)	264 (6,791)
	Non-CEO	-7 (398)	-64 (4,496)	90 (3,473)

The table shows that salary and bonus increased almost fourfold in the three sectors, and the sample mean in the three sectors is about 25 percent higher than the average salary and bonus in all sectors. Comparing this table with the previous one, we see that salary and bonus accounted for almost half of total compensation in the old sample, but these components account for less than one-quarter in the new three-sector sample. Thus, total compensation has increased much faster than salary and bonus in the three sectors. In fact, CEOs in the restricted sample received a lower salary and bonus, on average, than CEOs in the unrestricted sample, whereas Table 3 shows the inequality is reversed for average total compensation.

The component contributing the most to this dramatic shift is the options granted to managers, valued using the Black-Scholes formula. Options granted have increased on average more than thirtyfold in the three sectors. More than half of the total compensation comes from options granted in the restricted sample, or about three times the amount for salary and bonus. Comparing the two new samples, the value of stock options granted figures less prominently in the unrestricted sample than in the restricted, but is still over one-third. In both the restricted

and the unrestricted samples, the value of options grants is the biggest component of managerial compensation. In the old sample, the value of options granted is seven times the value of stock granted, greater than the ratio in the new unrestricted sample (six) but much less than the ratio in the new restricted sample (fourteen). Thus, stock grants in the three sectors, a relatively small component of managerial compensation, has diminished in importance.

Holding financial securities in their own firms rather than a well-diversified market portfolio exposes managers to considerable uncertainty. Table 4 shows that changes in wealth from holding options are more dispersed than any other component. This point is all the more noteworthy considering that much of cash, bonus, and grants is not contingent on firm performance; indeed, our analysis below shows that these components are partly explained by sector, firm size, and general affluence as captured by GDP. Changes in wealth from holding firm stock also add considerable volatility to compensation; the standard deviation is higher than for cash and bonus, option grants, and stock grants. Note that the standard deviation of both these components has dramatically increased—changes in stocks and options by more than one hundredfold. The two components underlie the increased variation in managerial compensation.

To summarize the discussion of our dataset, managerial compensation has substantially increased in real terms and become more dispersed. This has been accomplished by dramatically increasing stock option grants. These trends compellingly confirm previous evidence based on shorter time periods.⁶ How managerial compensation has changed relative to firm size depends on the measure used. Comparing the old sample with the new, average managerial compensation has increased relative to employment, fallen in two sectors relative to the value of assets, and increased roughly proportionately with output. Although the three sectors have notable differences, they are comparable to the unrestricted sample in many respects. Given the similarities we have noted between our data and those used in previous research and the comparability between restricted and unrestricted samples, we believe conclusions reached about the importance of moral hazard in the restricted sectors have broader applicability.

V. Identification

We estimate our model from panel data on financial returns, managerial compensation, and auxiliary information describing the firm and the executive position. Our measure of excess returns is constructed by first regressing financial returns on a vector of risk adjustment factors to estimate γ in equation (6). Since the estimator of γ is \sqrt{N} -consistent for a sample of size N , our discussion of identification is unaffected by assuming we can construct the true series of excess returns. We also assume that our measure of compensation, denoted by \tilde{w}_t , is an error-ridden measure of true compensation, w_t , where the error $\varepsilon_t \equiv \tilde{w}_t - w_t$ is orthogonal to the other variables of interest with standard deviation $2b_{t+1}\rho^{-2}\zeta$.

In our model, all firm-manager matches of the same type choose whether to implement a shirking or a diligent work contract, depending on which is more profitable. If shareholders implement a shirking contract, the manager receives a constant wage given by equation (8). In principle, we can test which types of firm-manager matches are governed by contracts that depend on excess returns and then partition the firm-manager match types by the contract they implement. This would allow us to estimate $f_1(x)$ and $f_2(x)$ nonparametrically, and also exploit equation (8) as a moment condition which jointly restricts the values of ρ and (α_1/α_0) . In practice, 99.98 percent of the managers in the new sample and 99.88 percent of the managers in the old sample hold

⁶ See Hall and Liebman (1998) and Murphy (1999).

firm-specific assets in any given period. Since this channel is the largest source of variation in compensation attributable to excess returns, $f_1(x)$ cannot be estimated this way in practice.

Instead, we assume that all managers are offered incentives to work diligently, undertaking tasks they would not choose if they were paid a flat wage. Consequently, the contract depends on the industry, economic conditions, and the firm's characteristics. Hence, our empirical analysis includes factors to account for heterogeneity across the sample of firms when estimating the structural parameters. As noted above, $f_2(x)$ is nonparametrically identified from excess returns alone for the set of managers whose compensation varies with excess returns (in our sample all of them), and can be consistently estimated from a univariate nonparametric regression. Similarly the optimal compensation contract, $w_t(x_t)$ for diligent work, is identified from a panel on (\tilde{w}_t, x_t) , and consistent estimators of the wage contract for each type of firm-manager match for any given bond price can be obtained from successive cross sections on (\tilde{w}_t, x_t) .

Having demonstrated that $f_2(x)$ and $w_t(x_t)$ are identified, we can, without loss of generality, proceed as if the true $w_t(x_t)$ and $f_2(x)$ are known for the purpose of identifying the other parameters, namely: the risk-aversion parameter ρ ; tastes for shirking over diligence α_2/α_1 ; tastes for diligence over the value of quitting α_2/α_0 ; and the signaling function $g(x)$. A detailed proof of this result for this particular framework can be found in the Web Appendix, while similar results for a more general class of models of which this is a special case can be found in Gayle and Miller (2009).

The first step is to prove that if the risk-aversion parameter, ρ , is known, then the remaining parameters are nonparametrically identified. Intuitively, we can nonparametrically recover the signal function $g(x_t)$ up to a normalization by tracing out the slope of the contract as a function of abnormal returns x_t . A boundary condition, such as that $g(\infty) = 0$, which is satisfied in our framework, then fully determines $g(x_t)$. Note, also, that $f_2(x_t)$ can be nonparametrically estimated from data on abnormal returns x_t . Consequently, $f_1(x_t) = g(x_t)f_2(x_t)$, the probability density function for shirking, is identified. Given $f_1(x_t)$, $f_2(x_t)$, and ρ , the participation constraint can be used to identify the ratio α_2/α_0 , and the incentive-compatibility constraint can then be used to identify the ratio α_2/α_1 .

A regards ρ , an observer with cross-sectional data on a homogeneous set of firms and managerial compensation paid out in that period cannot distinguish between a sample of managers with a high risk tolerance and unpleasant working conditions and a sample with lower tolerance but more nonpecuniary benefit. Thus, to identify ρ , we assume there are data on at least two states $r \in S$ and $s \in S$, that is, dates with distinct bond prices, or sectors where the nonpecuniary benefits of the job and the alternative opportunities for work are the same. More formally, the two states have different compensation plans, $w_r(x)$ and $w_s(x)$, but the same nonpecuniary benefits from diligent work α_2 . In this case, $w_r(x) \neq w_s(x)$ because the probability density function of abnormal returns from working diligently differs by state, that is, $f_{2r}(x) \neq f_{2s}(x)$; or the density from shirking differs, that is, $f_{1r}(x) \neq f_{1s}(x)$.

The existence of multiple states provides a means of identifying ρ . Since the participation condition holds state by state, we can, in principle, solve moment conditions of the form

$$\left[\int \exp[-\rho b_{r+1}^{-1} w_r(x)] f_{2r}(x) dx \right]^{\varkappa(r)} = \left[\int \exp[-\rho b_{s+1}^{-1} w_s(x)] f_{2s}(x) dx \right]^{\varkappa(s)}$$

in ρ , where $\varkappa(r) = \varkappa(s) = 1$ when they are sectors and $\varkappa(r) = 1 - b_r$ and $\varkappa(s) = 1 - b_s$ when they are dates. If there is a unique root common to all possible pairs of separate states, then ρ is identified, and consequently the other parameters are too.

In our study, bond prices vary significantly over the 60-year period, and we also impose exclusion restrictions, because the model is not saturated with a full set of interactions for every

different firm size and sector. In the old dataset, only 37 firms are tracked over 29 years, and the bond price declines steeply; in the new dataset there are 151 firms tracked over only 10 years, and bond prices remain relatively stable. So, in contrast to the old dataset, the sources of variation in the new dataset come predominantly from the cross section rather than the time series characteristics of the panel.

VI. Estimation

Nonparametric techniques for estimating this class of models are developed in Gayle and Miller (2009). Mainly to avoid the intractability of undertaking nonparametric estimation in the presence of many covariates, several of them continuous, we settled on a parametric specification defined below and used nonlinear parametric estimation techniques. Our empirical analysis assumes $f_j(x_t)$ is truncated normal with support bounded below by ψ ,

$$(15) \quad f_j(x) = \left[\Phi \left(\frac{\mu_j - \psi}{\sigma} \right) \sigma \sqrt{2\pi} \right]^{-1} \exp \left[-\frac{(x - \mu_j)^2}{2\sigma^2} \right],$$

where Φ is the standard normal distribution function and (μ_j, σ^2) denotes the mean and variance of the parent normal distribution. As indicated in the previous section, we cannot reject the null hypothesis of restricting the mean of abnormal returns to zero conditional on working in the data. We impose this restriction in the estimation of the parameter μ_2 . This restriction implies that μ_2 is determined as an implicit function of the parameters the truncated normal distribution under diligence. Denoting by φ the standard normal probability density function, the implicit function for μ_2 is given by⁷

$$(16) \quad 0 = E(x_t | I_{2t} = 1) = \mu_2 + \frac{\sigma \varphi[(\psi - \mu_2)/\sigma]}{1 - \Phi[(\psi - \mu_2)/\sigma]}.$$

This leaves the truncation point ψ , the mean of the parent normal distribution under shirking μ_1 , the common variance of the parent normal σ , the risk-aversion parameter ρ , the ratio of nonpecuniary benefits from working and shirking α_2/α_1 , and the ratio of nonpecuniary benefits from working and quitting α_2/α_0 to be estimated.

For each sector, the parameters of the distribution of returns are estimated separately, and the production parameters μ_1 and σ^2 are specified as functions of the number of employees in the firm, the firm's asset-to-equity ratio, and an aggregate economic condition, annual gross domestic product. Denoting the controls for observed heterogeneity by z_{1t} , we assume

$$(17) \quad \mu_1 = u'_1 z_{1t}$$

and

$$(18) \quad \sigma^2 = \exp(s'z_{1t}).$$

⁷ This equation is derived in G. S. Maddala (1983, 365).

The taste parameters α_2/α_1 and α_2/α_0 were specified as linear mappings of executive rank, firm sector, the number of employees in the firm, and the total assets of the firm. Denoting this vector of controls by z_{2t} , we assume

$$(19) \quad \alpha_2/\alpha_1 = a_1' z_{2t}$$

and

$$(20) \quad \alpha_2/\alpha_0 = a_2' z_{2t}.$$

The parameter estimates and their asymptotic standards were obtained in three steps. First, maximum likelihood estimates of the parameter vector determining the distribution of abnormal returns, (ψ, s) , were obtained using data on abnormal returns over time and across companies. In the second step, we used data on the abnormal returns and managerial compensation to form a generalized method-of-moments estimator from the participation constraint, the incentive compatibility constraint, and the managerial compensation schedule, and thus the remaining parameters (ρ, u_1, a_1, a_2) . In the third step, we corrected the estimated standard errors in the second step to account for the pre-estimation in the first step. Details of the estimation procedure are provided in the Web Appendix.

VII. Results

We now present the main results of this paper, including our measures of moral hazard and the structural parameter estimates that are used to derive them. First, we report our estimates of the distribution of abnormal returns, both when managers are diligent, $f_2(x)$, and when they shirk, $f_1(x)$. This yields an estimator for $g(x)$, the signal function, by sample and sector. Estimates of these probability distributions also yield, for each observation, a consistent estimator of τ_1 , the expected gross loss to a firm from ignoring the moral-hazard problem and paying managers a fixed wage. We investigate why the average loss has increased over the two sample periods, decomposing the change into three factors: namely, changes in firm size denoted by Δv , changes in the signal quality $\Delta g(x)$, and changes in $f_2(x)$, the returns density with diligence.

The latter parts of this section report on our estimates of the remaining parameters, including the other two measures of moral hazard, which depend on managerial preferences. First, we present our estimates of the parameters that determine whether the manager works for the firm or not, α_2/α_0 , and the divergence between managerial and shareholder interests, α_2/α_1 , along with the estimated coefficient of absolute risk aversion, ρ . The nonpecuniary benefits to the manager from pursuing his own objectives within the firm rather than profit maximizing, τ_2 , is a function of these three parameters only, and our results on the distribution of τ_2 are then discussed. We conclude the section with our findings on the welfare cost of moral hazard, τ_3 , which depends on both managerial preferences and the distribution of abnormal returns.

A. Abnormal Returns from Working Diligently

The first step in estimation, estimating the probability distribution of abnormal returns from working, provides further evidence on potential sources of change in managerial contracting. If the estimates of (ψ, s) vary between the two datasets, we might infer that the technology of production has changed in ways that might rationalize the trends observed in compensation plans. The parameter estimates and their standard errors are displayed in Table 5.

The table shows that there are significant differences between the two time periods, and that these differences are, for the most part, sector specific. The only common trend is that the effect

TABLE 5—PARAMETER ESTIMATES OF THE DILIGENT-RETURNS DISTRIBUTION
(Standard deviations in parentheses)

Parameter	Sector	Variable	Old	New
σ^2	Aerospace	Constant	-1.42 (0.375)	4.184 (1.492)
		Asset to equity ratio	-354 (135)	33.57 (49.94)
		Number of employees	-7.08 (1.38)	-0.106 (0.135)
		GDP	0.379 (3.14)	-8.23 (1.64)
σ^2	Chemicals	Constant	-3.08 (0.097)	-4.16 (0.703)
		Asset to equity ratio	77.3 (8.28)	-6.92 (7.69)
		Number of employees	-0.352 (0.222)	0.533 (0.559)
		GDP	-5.53 (1.15)	1.97 (0.777)
σ^2	Electronics	Constant	-2.07 (0.326)	-7.12 (0.46)
		Asset to equity ratio	-1.119 (139)	8.926 (13.5)
		Number of employees	0.355 (0.28)	0.877 (0.22)
		GDP	-16.6 (2.13)	5.44 (0.47)
ψ	Aerospace		-0.71 (0.04)	-0.79 (0.08)
	Chemicals		-0.47 (0.11)	-1.26 (0.32)
	Electronics		-0.61 (0.15)	-1.6 (0.48)
$\sqrt{\text{Var}(x_{nt} l_{2nt} = 1)}$	Aerospace	Standard deviation	26.72 (6.00)	20.61 (5.98)
	Chemicals		17.42 (3.28)	32.40 (3.96)
	Electronics		22.09 (5.52)	37.90 (5.98)

on abnormal returns of adding workers to a firm has increased the dispersion. In the case of GDP, we cannot tell whether the different coefficient estimates are attributable to changes that have taken place, or to a nonlinear effect on the variance, because the level of GDP in every period of the new restricted sample is more than twice its level in the old sample. What we can conclude, however, is that not only have the covariates that determine the higher-order moments of the abnormal returns changed, as we saw in Table 2, but their marginal impact has also changed, potentially confounding attempts to explain why managerial compensation has increased and become more sensitive to firm performance. Finally, reported at the bottom of the table is the average variance for the returns of firms in each sector. Imposing the truncated-normal assumption on the distribution of abnormal

TABLE 6—PARAMETER ESTIMATES OF THE SHIRKING-RETURNS DISTRIBUTION
(Standard deviations in parentheses)

Parameter	Sector	Variable	Old	New
μ_1	Aerospace	Constant	-0.051 (0.011)	-0.085 (0.001)
		Asset to equity ratio	0.042 (1.71)	-0.003 (4.1E-5)
		Number of employees	-0.015 (9.52)	0.019 (3.5E-4)
		GDP	-0.056 (11.0)	-0.014 (1.9E-4)
	Chemicals	Constant	-0.015 (0.003)	-0.021 (2.1E-3)
		Asset to equity ratio	-0.063 (0.077)	-0.071 (0.005)
		Number of employees	-0.0428 (0.001)	-0.088 (0.005)
		GDP	-0.025 (0.003)	-0.031 (3.29E-3)
	Electronics	Constant	-2.0E-4 (1.1E-5)	-0.008 (4.2E-4)
		Asset to equity ratio	-0.025 (6.6E-4)	-0.034 (4.9E-4)
		Number of employees	-0.024 (6.8E-4)	-0.011 (3.8E-3)
		GDP	-0.057 (0.002)	-0.017 (4.0E-4)
$E(x_{it} I_{it} = 1)$	Aerospace	Mean	-5.523 (1.47)	-8.755 (1.22)
	Chemicals		-3.176 (1.25)	-1.771 (1.22)
	Electronics		-2.137 (1.56)	-2.456 (0.23)

returns does not have a significant effect on its estimated variance. The estimates in Table 5 are comparable to our consistent estimates of the unconditional standard deviations presented in Table 1.

B. Abnormal Returns from Shirking

The remaining parameter estimates were obtained from the second step. Table 6 presents our estimates of the coefficient vector u_1 —which determines μ_1 , the sample mean of the parent distribution of abnormal returns under shirking, for different values of the covariates.

Although there are significant differences between the coefficients, they are not as pronounced as those reported in Table 5. One summary measure of the effects of shirking is the expected decline in abnormal returns by sector. The estimates on the last three lines of the table show that not providing an incentive to the manager would have led, on average, to losses of between 1.77 percent and 8.75 percent of the equity value of the firm, depending on the sample and the sector. There is, however, no statistical evidence that the returns from shirking have fallen. In chemicals the estimated average returns to equity have risen, and in aerospace they have fallen, but, after accounting for asymptotic estimation error in the parameters and the standard deviations within the sector samples, none of the three differences is significant.

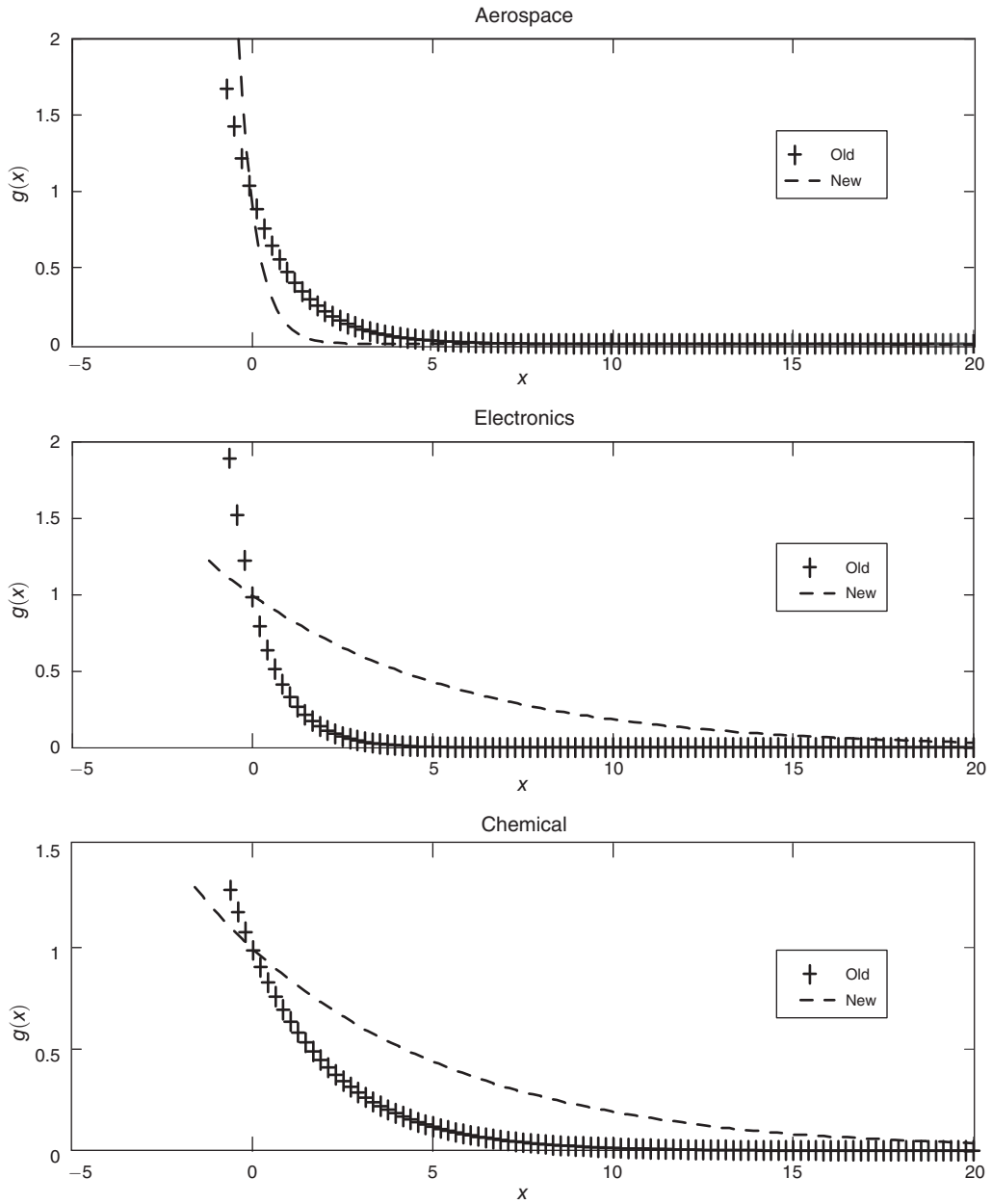


FIGURE 1. LIKELIHOOD RATIO

C. Signal Quality

Using the estimates of the probability density function for abnormal returns to firms conditional on managerial effort, Figure 1 plots the signal of managerial effort, $g(x)$, by sector and sample. To interpret the plots, note that signal quality can be measured by the distance between $g(x)$ and the constant one. Heterogeneity of signal quality among the three sectors has increased

TABLE 7—GROSS LOSSES TO FIRMS FROM SHIRKING IN MILLIONS OF US\$ (2000)
(Standard deviations in parentheses)

Parameter		Industry	Old	New
τ_1	Per year	Aerospace	13.751	180.212
			(29.522)	(261.294)
	Present value		81.065	1,261.484
			(177.132)	(1,829.058)
	Per year	Chemicals	33.392	160.038
			(73.537)	(240.970)
Present value		200.352	1,120.266	
		(441.222)	(1,686.79)	
Per year	Electronics	16.650	230.566	
		(49.182)	(600.607)	
Present value		99.907	1,613.962	
		(894.492)	(4,204.249)	

from the earlier to the later sample periods. Increased variation of abnormal returns in the chemical and engineering sectors presented in Table 1 translates to a weaker signal about managerial effort (because the distance has shrunk), while the decline in variance of abnormal returns has improved the precision of the signal in the aerospace sector.

Comparing the figure with Table 3, there is no discernible pattern linking signal quality to variance in managerial compensation. In aerospace, the increased variation in managerial compensation is comparable to that in electronics and greater than in chemicals, yet the signal became stronger in aerospace, weaker in chemicals, and much weaker in electronics.

D. Loss from Ignoring the Moral Hazard

The model implies that the expected loss per period to the firm from the manager pursuing his own interests rather than value maximization is

$$(21) \quad \tau_1 = -v \int x f_1(x) dx,$$

where v is the value of the firm in the previous period. Multiplying the expected loss for each firm by its size, we obtain estimates of τ_1 for each firm-year observation. Table 7 displays the estimated average over all firms (that is, before compensation), from inducing the manager to shirk, both per year and as a net present value calculation, by sector and for the two samples.

The implied average losses have increased more than tenfold in the aerospace and electronics sectors, and by a factor of about five in the chemicals sector. In aerospace and electronics, the mean return to firms from the manager shirking has fallen, and the size of the firms has increased. Both factors contribute to the larger expected losses. In the chemicals sector, the mean return from shirking, while negative, has increased and this partly offsets the greater loss due to the fact that chemical firms are larger. Comparing the present value of the losses as a ratio of the total assets and the equity value of the firm reported in Table 2, we see two measures of how much claimants on the firm, and in the latter case shareholders, would lose from not providing an incentive to managers. Controlling for sector, as a ratio of total assets, the implied losses are of the same order of magnitude in the two datasets, roughly one-ninth in aerospace, just under

one-half in chemicals, and about two-thirds in electronics. As a fraction of assets, the losses that would be incurred by not providing an incentive to managers appears relatively stable in these three sectors. Since firms are more leveraged than before, the loss has increased as a fraction of equity value. This is most noticeable in two of the sectors (electronics and chemicals), where the average estimated present value of losses exceeds the average equity value in the new data, but not the in old.

The dominant role of firm size in explaining the large increase in the cost of ignoring moral hazard is evident from expressing τ_1 as the negative of the product of firm size v and the expected value of the signal $g(x)$ when the manager works diligently. Differencing the estimates obtained for the two regimes, we obtain the decomposition

$$(22) \quad -\Delta\tau_1 = (\Delta v) \int xg(x)f_2(x) dx + v \int x[\Delta g(x)] f_2(x) dx + v \int xg(x)[\Delta f_2(x)] dx.$$

The first of the three expressions on the right side, the change in cost of moral hazard due to the increasing size of firms, is unambiguously positive. The second expression arises because of changes in $g(x)$. In two of the sectors, the signal has weakened, reducing the gap between $f_1(x)$ and $f_2(x)$ and thus mitigating the losses that would be incurred from encouraging the manager to pursue his own goals instead of expected-value maximization. The third expression captures the effects of the change in the distribution of abnormal returns. Noting that $f_2(x)$ has undergone a mean preserving spread in two sectors and that $g(x)$ is a convex decreasing function, it follows that the third expression is positive for these sectors, thus reducing the loss incurred. In summary, the growth of firms increased the losses from shirking so much that it dominates the other two effects.

E. Participation and Alternative Work

We now turn to our estimates of the preference parameters and the two other measures of moral hazard we described earlier. Given the attitude of managers toward risk, which as we show below has not changed much, changes in α_2/α_0 come from the shifting supply and demand for managerial services. Our estimates of α_2/α_0 for the three sectors are displayed in Table 8. The parameter is estimated for each sector, separately for CEOs and other executives, as a linear mapping of firm size by assets and employment. Comparing the results for CEOs from the old and new data, every sector coefficient has significantly increased, in chemicals the most and in aerospace the least. In addition the coefficient on total assets, negative but insignificant in the old dataset, acquires positive significance in the new dataset, while the coefficient on employment has significantly increased too.

The results for other executives are less noteworthy but similar. The sector coefficients in aerospace and electronics have significantly increased, while in chemicals we cannot reject the null hypothesis that the sector coefficient (the sum of the constant and the dummy) is unchanged. The coefficient on employment has also significantly increased, but the large estimated standard error on assets in the old sample implies that we cannot reject the hypothesis of no change for this measure of firm size.

If a freely available perfect monitor existed to eliminate moral hazard, managers would be paid their reservation wage, w_2 , in our framework. Consequently, our model predicts that, in the absence of moral hazard, with little change in preferences for risk, and roughly constant interest rates, compensation would have increased by the log of the ratios for the estimated α_2/α_0 parameters. Our estimates imply, that for an average size firm in the old sample, risk adjusted compensation would have risen by 2.3, which, up to two decimal places, is identical to the increase in national income per capita over the comparable period. Therefore, we should not attribute the rise in compensation relative to the rise in aggregate output and general living standards to a decline in the attractiveness

TABLE 8—NONPECUNIARY BENEFITS FROM DILIGENCE RELATIVE TO THE OUTSIDE OPTION
(Standard deviations in parentheses)

Parameter	Rank	Variable	Old	New	
α_2/α_0	CEO	Constant	0.985 (0.052)	3.91 (0.004)	
		Assets	-0.475 (0.102)	2.31 (0.314)	
		Employees	1.08 (0.104)	2.8189 (0.121)	
		Aerospace dummy	2.32 (1.03)	1.06 (0.082)	
		Chemicals dummy	0.403 (0.116)	1.75 (0.072)	
		Non-CEO	Constant	0.838 (0.198)	2.44 (0.025)
			Assets	1.77 (7.41)	0.605 (0.041)
	Employees		0.626 (1.96)	2.35 (0.06)	
	Aerospace dummy		1.29 (11.1)	1.46 (0.031)	
	Chemicals dummy		0.458 (1.2)	-1.42 (0.0043)	

of managerial work, or to demands for increased skills in managers. At the same time, our finding that the marginal rewards from heading a larger firm, whether measured by employment or assets, have increased corroborates findings cited in the introduction that the market for CEOs has become more competitive on both the supply and demand sides.

F. Conflict between Management and Shareholder Goals

Table 9 reports our estimates of α_2/α_1 , a measure of the divergence between managerial and shareholder interests characterizing how much worse off a manager is by working instead of shirking when compensation is determined independently of his choice. The higher this ratio, the less desirable is diligence compared to shirking, with incentives perfectly aligned at $\alpha_2/\alpha_1 = 1$. We did not constrain the parameters to satisfy the inequality $\alpha_2 > \alpha_1$. Thus, our findings that the estimates are all significantly greater than one validates our model on this dimension.

Comparing the results from the old and new datasets, we cannot, for the most part, reject the hypothesis of no change against the one-sided hypothesis that the ratio has risen. Many of the changes are insignificant. With regard to CEOs, the coefficients on the electronics sector and the number of employees have significantly fallen; only the coefficient on assets has significantly increased. The main evidence against the hypothesis of no change is found in the chemicals sector for the non-CEOs, where the interests of non-CEOs and shareholders appear to have diverged.

In both samples, increasing a firm's assets or its workforce exacerbates the conflict between management and shareholders, a finding that broadly conforms to the predictions of theories about corporate governance and the degree of separation between firm ownership and control. Although the evidence with respect to non-CEOs is more ambiguous, the coefficient on assets for CEOs has increased, but the coefficient on employees has declined. Noting that average firm size has increased as measured by asset value, but has declined as measured number of employees, the

TABLE 9—NONPECUNIARY BENEFITS FROM DILIGENCE RELATIVE TO SHIRKING
(Standard deviations in parentheses)

Parameter	Rank	Variable	Old	New	
α_2/α_1	CEO	Constant	3.23 (0.317)	1.85 (0.022)	
		Assets	0.109 (0.015)	2.72 (0.02)	
		Employees	5.06 (0.251)	1.90 (0.07)	
		Aerospace dummy	10.224 (4.17)	13.123 (1.62)	
		Chemicals dummy	8.0 (2.09)	9.53 (2.4)	
		Non-CEO	Constant	3.08 (1.73)	3.02 (0.011)
			Assets	13.0 (4.07)	7.39 (0.125)
	Employees		3.06 (1.96)	2.35 (0.041)	
	Aerospace dummy		4.47 (9.22)	8.39 (0.123)	
	Chemical dummy		2.35 (1.86)	8.04 (1.45)	

change in coefficient values is broadly consistent with convex increasing organizational costs in these two dimensions of firm size. Williamson (1967), among others, noted that this cost structure can explain the limits to natural monopoly in the presence of strong technological returns to scale.

G. Risk Aversion and Fit

The estimate for the risk-aversion parameter ρ , calculated for million-dollar units, is 0.519 for the old sample with an estimated standard error of $4.9\text{E}-3$, and 0.501 for the new sample with a standard error of $6.7\text{E}-6$. Although they differ across the two samples in a statistical sense, the difference has a small economic impact on the optimal contract, and on the measures of moral hazard reported below. For example, a manager with risk-aversion parameter of 0.501 would be willing to pay \$240,670 to avoid a gamble that has an equal probability of losing or winning one million dollars, whereas the certainty equivalent loss for a manager with risk-aversion parameter of 0.519 is \$248,620.

Comparing the estimates obtained for the old and new samples, ξ , a parameter that is proportional to the variance of the measurement error ($2b_{t+1}\rho^{-2}\xi$) increases by a multiple of about 12 from 0.008 with an estimated standard error of $2.6\text{E}-4$, to 0.101 with an estimated standard error of $8.7\text{E}-5$. This increase is offset by the significant reduction in bond prices over the two samples (as we indicated earlier). Hence, the variance of the measurement error has increased by far less. The variation in compensation explained by the optimal contract, R^2 , is comparable in the two datasets—0.63 in old dataset and 0.46 in the new—rising to over two-thirds when the samples are trimmed by a 5 percent outlier tail.⁸ Future research on corporate governance

⁸ In this context, we define R^2 by $1 - \text{Var}(\varepsilon)/\text{Var}(\tilde{w})$, where ε is the measurement error from the final stage estimation of the equation $\tilde{w} = w(x) + \varepsilon$.

TABLE 10—NONPECUNIARY BENEFITS OF SHIRKING AND WELFARE COST
(In thousands of US\$ (2000); standard deviations in parentheses)

Parameter	Industry	Rank	Old	New
τ_2	Aerospace	CEO	2,380 (43)	4,000 (92)
		Non-CEO	1,500 (72)	3,400 (78)
	Chemicals	CEO	920 (274)	3,800 (209)
		Non-CEO	812 (321)	600 (451)
	Electronics	CEO	747 (432)	3,048 (387)
		Non-CEO	436 (515)	2,070 (366)
τ_3	Aerospace	CEO	500 (1,316)	10,350 (15,473)
		Non-CEO	330 (1,413)	1,280 (10,501)
	Chemicals	CEO	490 (1,437)	2,973 (5,087)
		Non-CEO	299 (206)	301 (1,678)
	Electronics	CEO	278 (1,257)	4,873 (17,285)
		Non-CEO	67 (188)	1,206 (11,159)

might explain the role of directors and management within a noncooperative game, and thereby improve the model fit to the data, especially in the tails.⁹

H. Benefits for Managers from Shirking

The two remaining measures of moral hazard, τ_2 and τ_3 , can now be computed from the estimated parameters. The nonpecuniary value of the deviating from the incentive-based contract depends only on the preferences of the manager, not the distribution of the abnormal returns. For each observation, we computed a consistent estimator for τ_2 :

$$(23) \quad \tau_2 = [\rho(b_t - 1)]^{-1} b_{t+1} \ln(\alpha_2/\alpha_1).$$

The top panel of Table 10 reports, by sector and executive position, the average of the consistent estimators, and consistent estimates of their respective standard deviations. The firm averages for each executive type by sector have increased in five out of the six categories, by a factor of more than three for CEOs in two sectors. As a proportion of total compensation averaged over

⁹ Commenting on the sensitivity of previous results to measures of firm size in the analysis, Chiappori and Salanié (2003) noted that unless heterogeneity between firms is treated within the analysis, interpreting the findings and attributing causality is problematic. Our results confirm their conclusions by showing the large extent to which including a set of control variates explains trends in executive compensation through firms' changing composition.

observations for each executive type by sector, the compensating differential to managers for pursuing their own interests has fallen in all six categories. A key factor contributing to this measure of importance, τ_2/w , is that changes in the supply and demand for managerial services has roughly doubled the compensation of managers of a firm with any given set of characteristics.

In both samples, the average τ_2 is tiny compared to the expected losses a firm would incur; our model predicts there are enormous gains from having managers act in the interest of shareholders. From the manager's perspective, however, τ_2 is quite substantial, and for a sizeable proportion of the sample population, exceeds actual and even expected compensation. This paradox is resolved by noting that the manager would be harshly penalized if the firm does poorly, which is of course more likely if he shirks.

Perhaps the most striking feature of these results is how they compare with those reported in Table 9. Recall that Table 9 exhibits estimates of α_2/α_1 for each firm and executive type, whereas Table 10 averages over firms within sector after taking logarithms and scaling by ρ^{-1} . Since ρ has not changed much, and the estimated changes in α_2/α_1 are for the most part insignificant or negative, we attribute the dramatic differences between the tables to the changing composition of firms within each sector. More specifically, the effects of the average growth in firm assets dominates the decline in employment, and is largely responsible for the increased compensating differential to work for shareholders versus pursuing some other agenda.

I. Cost of the Moral Hazard to Shareholders

The last measure of moral hazard, τ_3 , is the welfare cost of the moral hazard—the willingness of a firm to pay for a perfect monitor—thus eliminating moral hazard. From the definition of τ_3 and the solution for the optimal contract, it follows that the welfare cost may be expressed as

$$(24) \quad \tau_3 = b_{t+1} \rho^{-1} \int \ln [1 + \eta_t (\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta_t g(x_t)] f_2(x) dx.$$

The bottom panel of Table 10 presents consistent estimates of the average of τ_3 in the two samples and three sectors, along with the consistent estimates of the standard deviations. The table shows that the increase in managerial compensation presented in Table 3 is mirrored in the increased cost of moral hazard.

From the formula above and the formula for η_t , changes in τ_3 are ultimately attributable to changes in α_2/α_1 , $f_1(x)$, and $f_2(x)$ only. After adjusting for the general rise in living standards, the estimated model attributes practically all the increase in managerial compensation to moral hazard, and hardly any of it to changes in the supply and demand for managers, as reflected in the participation condition and hence α_2/α_0 .

To further investigate the sharply increased cost of moral hazard, we first note that changes in b_{t+1}/ρ between the two samples are minimal, and decompose $\Delta\tau_3$ into those stemming from changes in $f_2(x)$, and changes in the integrand. Since $g(x_t)$ is a convex decreasing function,

$$(25) \quad \ln [1 + \eta_t (\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta_t g(x_t)]$$

is a concave increasing function. Noting again that $\Delta f_2(x)$ is a mean-preserving spread in chemicals and engineering but not in aerospace, it therefore follows that

$$(26) \quad b_{t+1} \rho^{-1} \int \ln [1 + \eta_t (\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta_t g(x_t)] \Delta f_2(x) dx$$

is positive in chemicals and engineering but negative in aerospace. Thus, changes in the distribution of abnormal returns cannot explain why the welfare costs of the moral hazard increased in aerospace, the sector where the biggest increases occurred. The remaining component to explain $\Delta\tau_3$ is

$$(27) \quad b_{t+1}\rho^{-1} \int \Delta \ln [1 + \eta_t(\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta_t g(x_t)] f_2(x) dx.$$

The predominant change is due to a sharp increase in α_2/α_1 averaged over firms for the reasons demonstrated in Table 10. This component is the most important factor responsible for the increase in τ_3 . To recapitulate, increased firm assets exacerbated the conflict between managers and shareholders by creating new opportunities for managers to act against shareholder interests. These were resolved through the compensation schedule by placing greater weight on penalizing poor firm performance and rewarding superior abnormal firm returns, thus subjecting risk-averse managers to the vagaries of greater insider wealth and causing their expected compensation to rise at a rate much greater than that of national income per capita.

VIII. Conclusion

The welfare cost of moral hazard is a compensating differential paid to risk-averse managers to hold insider wealth and accept nondiversifiable risk that realigns their incentives to those of the stockholders, who do not price risk from an individual firm's abnormal returns because of their portfolio choices. Table 10 shows that the welfare cost of the moral hazard associated with employing CEOs has increased by an estimated factor of more than 20 times in the aerospace and electronics sectors and sixfold in the chemicals sector. Subtracting the welfare costs of the moral hazard displayed in Table 10 from the expected compensation paid to top executives reported in Table 3, we obtain, for each of the six categories, the average certainty equivalent wage, which equates the supply and demand for managerial services for a given firm. The overall increase in the 60-year period is 2.3, the same as the increase in national income per capita. Therefore, our results attribute all the difference between the rate of increase in managerial compensation and the rate of increase in national income per capita to the rising welfare cost of moral hazard.

The cost of moral hazard depends on the preferences of managers, what shareholders observe about their behavior, the distribution of abnormal returns accruing to firms, and the characteristics of the firms they manage. We do not attribute the steep increase in the welfare cost to changing tastes. Table 9 shows that, if anything, the conflict between a firm with a given set of characteristics and its executives has declined. As documented in Table 1, and our estimates in Tables 5 and 6, there have been changes in the probability distributions of abnormal returns, but not all in the same direction. Table 10 shows that managerial preferences for risk have remained stable in an economic sense, and the compensating differential of deviating from the goal of maximizing the expected value of the firm with a given set of characteristics has not increased. Nevertheless, Table 7 shows that if managers were paid a flat wage to prevent skimming, and our model of moral hazard was correctly specified, then conflict between managerial and shareholder objectives would remain unresolved, and the ensuing losses incurred by firms would be catastrophic and would have grown substantially over the last 60 years.

The exogenously changing distribution of firm size is the primary driver of the steep increase in the welfare costs of moral hazard and managerial compensation. In an average firm belonging to the three sectors we studied, Table 2 shows that sales tripled, employment halved, and assets increased sixfold. Our model does not explain firm size, but does distinguish between the direct effect of firm size through the market-clearing certainty-equivalent wage, and its indirect effect

through moral hazard and its welfare costs. The differences we previously noted between Tables 9 and 10 essentially establish that the indirect effect of firm size on managerial compensation is largely responsible for the increased costs of moral hazard. Our empirical results lend support to the view that the welfare cost of moral hazard has risen because large firms are more susceptible to governance problems than small firms, managers having greater opportunities to engage in opportunistic behavior at the expense of firm owners, a well-understood and common intuition that dates back to Berle and Means's (1932) early work on corporations.

The direct effect of firm size on managerial compensation is determined by the demand of firms of various sizes for managers of differing skill levels, the supply of such managers, and how the nonpecuniary benefits of diligent management are affected by firm size. We interpret Table 8 as empirical evidence showing that the market for managers has become more differentiated and the relative premium for managing a larger versus smaller firm, whether measured in terms of employees or assets, has increased, a result that corroborates recent work by Xavier Gabaix and Augustin Landier (2008), and Murphy and Jan Zabojnik (2004, 2006).¹⁰ Noting that the average certainty-equivalent wage for managers has risen only at the same rate as the price of other factor inputs, it immediately follows that the certainty equivalent of managers of small firms has not kept pace with aggregate growth. Consequently the reduction in the net demand for managerial services, and/or increased nonpecuniary benefits from being a manager (less likely in our opinion), was essentially offset by the first effect of increasing firm size along with an increased premium for managing larger firms.

The evidence we present in Tables 2 through 4 shows that the three sectors we investigated are quite representative of all sectors and exhibit trends that have been found in previous studies using other data. All three measures of moral hazard in the three sectors for the new dataset are comparable to values we found in a much broader range of sectors for almost the same years.¹¹ Our model explains about half of the observed variation of managerial compensation in both datasets we studied. This leads us to conclude that the steep increase in the welfare costs of moral hazard induced by the exogenous shifting composition of firms is the most important factor explaining the increased levels and volatility in managerial compensation.

REFERENCES

- Abowd, John M., and David S. Kaplan.** 1999. "Executive Compensation: Six Questions That Need Answering." *Journal of Economic Perspectives*, 13(4): 145–68.
- Aggarwal, Rajesh K., and Andrew A. Samwick.** 1999. "The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation." *Journal of Political Economy*, 107(1): 65–105.
- Antle, Rick, and Abbie Smith.** 1985. "Measuring Executive Compensation: Methods and an Application." *Journal of Accounting Research*, 23(1): 296–325.
- Antle, Rick, and Abbie Smith.** 1986. "An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives." *Journal of Accounting Research*, 24(1): 1–39.
- Berle, Adolf A., and Gardiner C. Means.** 1932. *The Modern Corporation and Private Property*. New York: Harcourt, Brace, & World.
- Bolton, Patrick, and Mathias Dewatripont.** 2005. *Contract Theory*. Cambridge, MA: MIT Press.
- Chiappori, Pierre-André and Bernard Salanié.** 2003. "Testing Contract Theory: A Survey of Some Recent Work." In *Advances in Economics and Econometrics—Theory and Applications, Eighth World*

¹⁰ See also Vicente Cuñat and Maria Guadalupe (2005) for an empirical study analyzing the impact of globalization on incentives.

¹¹ See Gayle and Miller (forthcoming). The sectors analyzed in that paper include energy, materials, industrials, consumer discretionary, consumer staples, health care, financial, information technology, telecommunication, and utilities. Our companion paper covers the years 1992 through to 2001, whereas the new dataset in this paper covers the years 1992 to 2003.

- Congress*, ed. M. Dewatripont, L. Hansen, and P. Turnovsky, 115–49. Cambridge: Cambridge University Press.
- Cuñat, Vicente, and Maria Guadalupe.** 2005. “Globalization and the Provision of Incentives Inside the Firm.” Centre for Economic Policy Research Discussion Paper 5950.
- Fudenberg, Drew, Bengt Holmström, and Paul Milgrom.** 1990. “Short-Term Contracts and Long-Term Agency Relationships.” *Journal of Economic Theory*, 51(1): 1–31.
- Gabaix, Xavier, and Augustin Landier.** 2008. “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics*, 123(1): 49–100.
- Garen, John E.** 1994. “Executive Compensation and Principal-Agent Theory.” *Journal of Political Economy*, 102(6): 1175–99.
- Gayle, George-Levi, and Robert A. Miller.** 2009. “Identifying and Testing Models of Managerial Compensation.” Tepper School of Business Working Paper 2009–E7.
- Gayle, George-Levi, and Robert A. Miller.** Forthcoming. “The Paradox of Insider Information and Performance Pay.” *CESifo Economic Studies*.
- Grossman, Sanford J., and Oliver D. Hart.** 1983. “An Analysis of the Principal-Agent Problem.” *Econometrica*, 51(1): 7–45.
- Hall, Brian J., and Jeffrey B. Liebman.** 1998. “Are CEOs Really Paid Like Bureaucrats?” *Quarterly Journal of Economics*, 113(3): 653–91.
- Haubrich, Joseph G.** 1994. “Risk Aversion, Performance Pay, and the Principal-Agent Problem.” *Journal of Political Economy*, 102(2): 258–76.
- Jensen, Michael C., and Kevin J. Murphy.** 1990. “Performance Pay and Top-Management Incentives.” *Journal of Political Economy*, 98(2): 225–64.
- John, Teresa A., and Kose John.** 1993. “Top-Management Compensation and Capital Structure.” *Journal of Finance*, 48(3): 949–74.
- Kostiuk, Peter F.** 1990. “Firm Size and Executive Compensation.” *Journal of Human Resources*, 25(1): 90–105.
- Laffont, Jean-Jacques, and David Martimont.** 2002. *The Theory of Incentives: The Principal-Agent Model*. Princeton, NJ: Princeton University Press.
- Maddala, G. S.** 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.
- Margiotta, Mary M., and Robert A. Miller.** 2000. “Managerial Compensation and the Cost of Moral Hazard.” *International Economic Review*, 41(3): 669–719.
- Masson, Robert Tempest.** 1971. “Executive Motivations, Earnings, and Consequent Equity Performance.” *Journal of Political Economy*, 79(6): 1278–92.
- Murphy, Kevin J.** 1999. “Executive Compensation.” In *Handbook of Labor Economics Volume 3b*, ed. Orley Ashenfelter and David Card, 2485–2563. Amsterdam: Elsevier Science, North-Holland.
- Murphy, Kevin J., and Jan Zabojnik.** 2004. “CEO Pay and Appointments: A Market-Based Explanation for Recent Trends.” *American Economic Review*, 94(2): 192–96.
- Murphy, Kevin J., and Jan Zabojnik.** 2006. “Managerial Capital and the Market for CEOs.” Queen’s University Department of Economics Working Paper 1110.
- Prendergast, Canice.** 1999. “The Provision of Incentives in Firms.” *Journal of Economic Literature*, 37(1): 7–63.
- Salanié, Bernard.** 2005. *The Economics of Contracts: A Primer*. 2nd ed. Cambridge, MA: MIT Press.
- Schaefer, Scott.** 1998. “The Dependence of Pay-Performance Sensitivity on the Size of the Firm.” *Review of Economics and Statistics*, 80(3): 436–43.
- Williamson, Oliver.** 1967. “Hierarchical Control and Optimum Firm Size.” *Journal of Political Economy*, 75: 123–38.

Online Appendix to “Has Moral Hazard Become a More Important Factor in Managerial Compensation?”

George-Levi Gayle and Robert A. Miller

Tepper School of Business, Carnegie Mellon University

August 2008

Abstract

In this online appendix we formally show that the model in the main text is identified, describe the empirical implementation of our estimation technique, and derive the asymptotic covariance of our estimator.

I. Identification

The identification of this model is an application of Gayle and Miller (2008). For the reasons given in the text, we proceed as if true compensation, w_t , and excess returns, x_t , are observed for the purposes of establishing identification of the other parameters. Identification of the remaining parameters, namely the risk-aversion parameter (ρ), tastes for shirking over diligence (α_2/α_1), tastes for diligence over the value of quitting (α_2/α_0), and the signalling function ($g(x)$) proceeds in two steps. First, we prove (α_2/α_1) , (α_2/α_0) , and $g(x)$ are identified if ρ is known. Then we give sufficient conditions for identifying ρ .

Defining $v_t(x, \rho)$ as

$$(1) \quad v_t(x, \rho) \equiv \left(\frac{\alpha_0}{\alpha_2} \right)^{1/(b_t-1)} \exp \left[\frac{\rho w_t(x)}{b_{t+1}} \right],$$

it follows from the optimal contract for diligent work,

$$(2) \quad w_t = \frac{b_{t+1}}{\rho(b_t - 1)} \ln \left(\frac{\alpha_2}{\alpha_0} \right) + \frac{b_{t+1}}{\rho} \ln \left[1 + \eta_t \left(\frac{\alpha_2}{\alpha_1} \right)^{1/(b_t-1)} - \eta_t g(x_t) \right],$$

that for a given value for ρ , a transformation of the optimal compensation, depending only on (observed) bond prices, is a linear mapping of $g(x)$. Namely,

$$(3) \quad v_t(x, \rho) = 1 + \eta_t \left[(\alpha_2/\alpha_1)^{1/(b_t-1)} - g(x) \right].$$

So, if the values of the intercept and the slope of the mapping could be found, and the value of ρ were known, then $g(x)$ could be simply determined. Taking the expectation with respect to x conditional on the price of bonds at time t yields

$$(4) \quad E[v_t(x, \rho)] = 1 + \eta_t (\alpha_2/\alpha_1)^{1/(b_t-1)} - \eta_t \equiv \underline{v}_t(\rho).$$

We now impose a regularity condition on $g(x)$, satisfied by our parameterization, that says $g(x) \rightarrow 0$ as $x \rightarrow \infty$. Intuitively this condition states that the shareholders attach negligible probability to a manager shirking if the firm's excess returns are extraordinarily high. The condition implies

$$(5) \quad \lim_{x \rightarrow \infty} v_t(x, \rho) = 1 + \eta_t (\alpha_2/\alpha_1)^{1/(b_t-1)} \equiv \bar{v}_t(\rho).$$

Solving for the signaling function $g(x)$, the nonpecuniary benefit ratio (α_2/α_1) , and the tastes for participation (α_2/α_0) given ρ , using equations (3), (4), and (5), proves the following.

Proposition 1. For any $\rho^* > 0$,

$$\begin{aligned}\frac{\alpha_2^*}{\alpha_0^*} &= \left[E_t \left\{ \exp \left[-\frac{\rho^* w_t(x)}{b_{t+1}} \right] \right\} \right]^{1-b_t} \\ \frac{\alpha_2^*}{\alpha_1^*} &= \left(\frac{\bar{v}_t(\rho^*) - 1}{\bar{v}_t(\rho^*) - \underline{v}_t(\rho^*)} \right)^{b_t-1} \\ g^*(x) &= \frac{\bar{v}_t(\rho^*) - v_t(x, \rho^*)}{\bar{v}_t(\rho^*) - \underline{v}_t(\rho^*)}.\end{aligned}$$

Proof of Proposition 1. The expression for α_2^*/α_0^* follows directly from rearranging the participation constraint (7). Subtracting equation (5) from (3), we obtain

$$\eta_t g(x) = \bar{v}_t(\rho^*) - v_t(x, \rho^*).$$

Subtracting equation (4) from (5) we obtain

$$(6) \quad \eta_t = \bar{v}_t(\rho^*) - \underline{v}_t(\rho^*).$$

Substituting for η_t using (6) in the previous equation and making $g(x)$ the subject of the resulting equation yields:

$$g(x) = \frac{\bar{v}_t(\rho^*) - v_t(x, \rho^*)}{\bar{v}_t(\rho^*) - \underline{v}_t(\rho^*)}.$$

Finally, making (α_2/α_1) the subject of equation (4) and then substituting for η_t using (6), we obtain

$$\frac{\alpha_2^*}{\alpha_1^*} = \left(\frac{\bar{v}_t(\rho^*) - 1}{\eta_t} \right)^{b_t-1} = \left(\frac{\bar{v}_t(\rho^*) - 1}{\bar{v}_t(\rho^*) - \underline{v}_t(\rho^*)} \right)^{b_t-1},$$

as required.

Proposition 1 establishes that if ρ^* is known then (α_2^*/α_1^*) , (α_2^*/α_0^*) , and $g^*(x)$ are identified since they can be written as a mapping of the data, because consistent estimates of the mappings $\bar{v}_t(\rho)$ and $\underline{v}_t(\rho)$ can be obtained from the data. See Gayle and Miller (2008) for details on constructing nonparametric consistent estimates of these quantities. A natural

place to begin investigating the identification of ρ^* is the participation constraint. When $(\alpha_2/\alpha_0) > 1$, meaning the nonpecuniary benefits of working do not fully compensate the manager for the total benefits of his alternative, and thus expected compensation is positive, the data imply a lower bound for the risk-aversion parameter, ρ . To picture this, define the mapping

$$\psi_t(\rho) \equiv E_t \left[\exp \left(-\frac{\rho w_t(x)}{b_{t+1}} \right) \right].$$

From its definition, $\psi_t(0) = 1$, while the assumption above implies

$$\psi_t'(0) = \frac{\partial}{\partial \rho} E \left[\exp \left(-\frac{\rho w_t(x)}{b_{t+1}} \right) \right]_{\rho=0} = -E \left[\frac{w_t(x)}{b_{t+1}} \right] < 0.$$

Also $\psi_t(\rho)$ is convex in ρ because

$$\frac{\partial^2}{\partial \rho^2} \left[\exp \left(-\frac{\rho w_t(x)}{b_{t+1}} \right) \right] = \left(\frac{w_t(x)}{b_{t+1}} \right)^2 \exp \left(-\frac{\rho w_t(x)}{b_{t+1}} \right) > 0$$

and the expectations operator preserves convexity. Assuming $\alpha_2/\alpha_0 > 1$, it now follows that $\psi_t(\rho)$ crosses the unit level from below just once at say ρ_t , which implies $\psi_t(\rho) > 1$ for all $\rho > \rho_t$. This rules out the possibility that $\rho^* \leq \rho_t$. Intuitively, the participation equation is satisfied by different combinations of ρ and α_2/α_0 satisfying $\rho > \rho_t$ and $\alpha_2/\alpha_0 = \psi_t(\rho)^{1-b_t}$ as we see in Figure 1.

Along this line, as ρ increases, the person becomes more risk averse, the expected utility from $w_t(x)$ declines along with its certainty equivalent, but this is just offset by nonpecuniary amenities from the job. Consequently an observer with cross-sectional data on a homogeneous set of firms and managerial compensation paid out in that period cannot distinguish between a sample of managers with a high risk tolerance and unpleasant working conditions, versus a sample with lower tolerance but more nonpecuniary benefits. The remaining parameters are then inferred from the value ascribed to ρ , the slope of the contract with respect to abnormal returns determining $g(x)$ and thence the probability distribution

of abnormal returns under shirking.

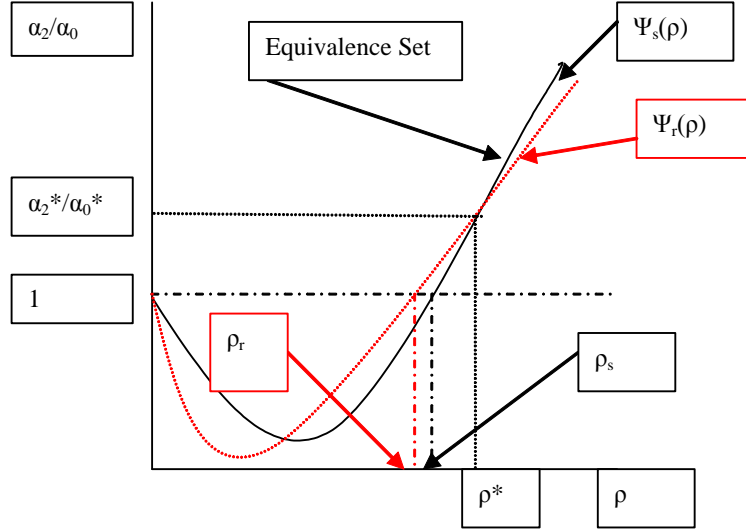


Figure 1: Equivalence Set

Accordingly, we now suppose there are data on at least two states $s \in S$, that is dates with distinct bond prices, or sectors where the nonpecuniary benefits of the job and the alternative opportunities for work are the same. More formally, the two states have different compensation plans $w_r(x)$ and $w_s(x)$ but the same nonpecuniary benefits from diligent work α_2 . In this case $w_r(x) \neq w_s(x)$ because the probability density function of abnormal returns from working diligently differs by state, that is $f_{2r}(x) \neq f_{2s}(x)$, or the density from shirking differs, that is $f_{1r}(x) \neq f_{1s}(x)$.

The existence of multiple states provides a means of identifying ρ . Since the participation condition holds for each state $s \in S$ separately, we can in principle solve moment conditions of the form

$$\left[\int \exp[-\rho b_{r+1}^{-1} w_r(x)] f_{2r}(x) dx \right]^{\varkappa(r)} = \left[\int \exp[-\rho b_{s+1}^{-1} w_r(x)] f_{2s}(x) dx \right]^{\varkappa(s)}$$

in ρ , where $\varkappa(r) = \varkappa(s) = 1$ when they are sectors and $\varkappa(r) = 1 - b_r$ and $\varkappa(s) = 1 - b_s$ when they are dates. Figure 1 illustrates how identification would be achieved with two states, ρ^* determined by a unique intersection of $\psi_s(\rho)$ with $\psi_t(\rho)$. Although there may be multiple

roots in ρ to the equations defined by the separate states $r \in S$ and $s \in S$, if there is a unique root common to all possible pairs, then ρ is identified.

II. Empirical Implementation and Standard Errors

In the old sample and the new restricted sample, the data are ordered by $n \in \{1, \dots, N\}$, where each observation refers to a firm-year vector of variables, including compensation paid to the three top executives, the abnormal return, the number of employees, the asset-to-equity ratio, GDP that year, the bond price in the current year (denoted b_n), the bond price the following year (denoted b_{1n}), and sector dummy variables.

A. Stage Zero

Recall that

$$x_n = \pi_n - \pi - z_n \gamma,$$

where γ is a 2×1 vector and z_{nt} is a 1×2 vector of sector-specific constants and GDP, and that x is estimate as the residual of the regression of z_n on $\pi_n - \pi$. Let $\gamma^{(N)}$ denote the estimate of γ from that regression. For each sector, we estimate the lower bound of the excess return distribution as

$$(7) \quad \psi^{(N)}(\gamma^{(N)}) = \min_{\{1, \dots, N\}} \{\pi_n - \pi - z_n \gamma^{(N)}\}.$$

Let γ_0 denote the true value of γ in the population. Note that if γ_0 were known then

$$\psi^{(N)}(\gamma_0) = \min_{\{1, \dots, N\}} \{\pi_n - \pi - z_n \gamma_0\}$$

would be a super-consistent estimate of $\psi_0(\gamma_0)$, the true value of ψ in the population. However, since we are using $\gamma^{(N)}$ instead of γ_0 , the following Lemma establishes that it is only \sqrt{N} consistent and gives its asymptotic variance.

Lemma 2. *Under standard regularity conditions,*

$$\sqrt{N} \left(\psi^{(N)}(\gamma^{(N)}) - \psi_0(\gamma_0) \right) \Rightarrow N(0, \text{var}(\psi_0)),$$

where $\text{var}(\psi_0) = \sigma^2(z) \underline{z}(z'z)^{-1} \underline{z}'$, $\sigma^2(z) = \text{var}(x|z)$ and \underline{z} is the value of z at the minimum of x .

Proof. Define

$$\psi^{(N)}(\gamma) = \min_{\{1, \dots, N\}} \{ \pi_n - \pi - z_n \gamma \} \text{ for any } \gamma \in \mathbb{R}$$

and

$$\psi_0(\gamma) = p \lim_{N \rightarrow \infty} \psi^{(N)}(\gamma) \text{ for any } \gamma \in \mathbb{R}.$$

Next note

$$(8) \quad \psi^{(N)}(\gamma^{(N)}) - \psi_0(\gamma_0) = \psi^{(N)}(\gamma^{(N)}) - \psi^{(N)}(\gamma_0) + \psi^{(N)}(\gamma_0) - \psi_0(\gamma_0).$$

Since $\gamma^{(N)}$ is a \sqrt{N} -consistent estimator of γ_0 ,

$$\psi^{(N)}(\gamma^{(N)}) - \psi^{(N)}(\gamma_0) = O_p(N^{-\frac{1}{2}}),$$

and since $\psi^{(N)}(\gamma_0)$ is a super-consistent estimator of $\psi_0(\gamma_0)$,

$$\psi^{(N)}(\gamma_0) - \psi_0(\gamma_0) = O_p(N^{-1}).$$

Therefore

$$(9) \quad \begin{aligned} \psi^{(N)}(\gamma^{(N)}) - \psi_0(\gamma_0) &= O_p(N^{-\frac{1}{2}}) + O_p(N^{-1}) \\ &= O_p(\max\{N^{-\frac{1}{2}}, N^{-1}\}) \\ &= O_p(N^{-\frac{1}{2}}). \end{aligned}$$

Hence $\psi^{(N)}(\gamma^{(N)})$ is \sqrt{N} consistent, which implies that $\sqrt{N}(\psi^{(N)}(\gamma^{(N)}) - \psi_0(\gamma_0)) \Rightarrow N(0, \text{var}(\psi_0))$. The variance formula for $\text{var}(\psi_0)$ follows from the asymptotic variance of $\psi_0(\gamma^{(N)})$.

B. Stage One

In order to take into account the pre-estimation in x , we now make its dependence on γ explicit by defining

$$x_n(\gamma) \equiv \pi_n - \pi - z_n \gamma.$$

For each sector, the log likelihood of observing $x_n(\gamma)$ is given by

$$(10) \quad l(\psi_0(\gamma_0), x(\gamma_0), \sigma) = \log \sigma + \ln \Phi \left[\frac{\mu(\psi_0(\gamma_0), \sigma) - \psi_0(\gamma_0)}{\sigma} \right] + \frac{[x(\gamma_0) - \mu(\psi_0(\gamma_0), \sigma)]^2}{2\sigma^2},$$

where $\mu(\psi_0(\gamma_0), \sigma)$ is defined as the implicit solution in μ_2 of the following equation.

$$(11) \quad \mu_2 + \sigma \frac{\phi \left[\frac{\mu_2 - \psi_0(\gamma_0)}{\sigma} \right]}{\Phi \left[\frac{\mu_2 - \psi_0(\gamma_0)}{\sigma} \right]} = 0$$

Let $S(\psi_0(\gamma_0), x(\gamma_0), \sigma)$, the score, be the derivative of $l(\psi_0(\gamma_0), x(\gamma_0), \sigma)$ with respect to σ and define

$$(12) \quad h_0(\psi_0(\gamma_0), x(\gamma_0), \sigma) = \begin{bmatrix} S(\psi_0(\gamma_0), x(\gamma_0), \sigma) \\ z'x(\gamma_0) \end{bmatrix}$$

to be the 3×1 vector of moment condition with

$$(13) \quad E[h_0(\psi_0(\gamma_0), x(\gamma_0), \sigma)] = 0.$$

Define

$$G_\sigma = E \left[\frac{\partial^2 l(\psi_0(\gamma_0), x(\gamma_0), \sigma)}{\partial \sigma \partial \sigma} \right],$$

$$G_\gamma = E \left[\frac{\partial S(\psi_0(\gamma_0), x(\gamma_0), \sigma)}{\partial \gamma} \right],$$

$$S(z) = S(\psi_0(\gamma_0), x(\gamma_0), \sigma),$$

$$D = -E[z'z],$$

and

$$\varphi(z) = E[z'z]^{-1} z'x(\gamma_0).$$

Under standard regularity conditions,*

$$\sqrt{N}(\sigma^N - \sigma) \implies N(0, V(\sigma)),$$

where

$$V(\sigma) = G_\sigma^{-1} E[\{S(z) + G_\gamma \varphi(z)\} \{S(z) + G_\gamma \varphi(z)\}'] G_\sigma^{-1'}$$

This follows directly from Theorem 6.1 of Whitney K. Newey and Daniel McFadden (1994).

C. Stage Two-Estimation and Standard Error

Having obtained estimates of the coefficients σ and $\psi_0(\gamma^0)$, which determine the probability density function for abnormal returns, $f_2(x)$, we estimated the remaining parameters $\theta \equiv (\rho, u_1, a_1, a_2, \xi)$ from orthogonality conditions derived from the participation and incentive-compatibility constraints, along with the score of the optimal contract's likelihood function in a generalized method-of-moments procedure, after substituting our estimate for σ , $\psi_0(\gamma^0)$, and $x(\gamma^0)$ obtained in the first step. Let the true value of θ be denoted by $\theta^o \equiv (\rho^o, u_1^o, a_1^o, a_2^o, \xi^o)$.

The first vector of orthogonality conditions is constructed from the participation constraints (a vector of three executives) of the form

$$(14) \quad h_{1n}(\boldsymbol{\theta}) = \exp[-b_{1n}^{-1}(\rho \tilde{w}_n + \xi)] - (a_2' z_n)^{1/(1-b_n)}.$$

The distributional assumptions on ε_n imply

$$(15) \quad E \{ \exp[-b_{1n}^{-1} (\rho^o \tilde{w}_n + \xi)] \mid w_n, b_{1n} \} = \exp[-b_{1n}^{-1} (\rho^o w_n)].$$

Because the participation equation is met with equality under the optimal contract, it follows that

$$(16) \quad E[h_{1n}(\theta^o)] = 0.$$

The second vector of orthogonality conditions is based on the incentive-compatibility constraint. Define the vector

$$(17) \quad h_{2n}(\boldsymbol{\theta}, x_n(\gamma), \sigma, \psi(\gamma)) = \exp[-b_{1n}^{-1} (\rho \tilde{w}_n + \xi)] \left[\frac{f_1(x_n(\gamma), \sigma, \psi(\gamma))}{f_2(x_n(\gamma), \boldsymbol{\theta}, \sigma, \psi(\gamma))} - (a'_1 z_n)^{1/(b_n-1)} \right].$$

The incentive-compatibility constraint is also met with equality under the optimal contract, when the parameters are set to their true values, implying

$$(18) \quad E [h_{2n}(\boldsymbol{\theta}^o, x_n(\gamma^o), \sigma^o, \psi(\gamma^o))] = 0,$$

where $(\sigma^o, \psi(\gamma^o), \gamma^o)$ are the true values of (σ, ψ, γ) .

The final set of orthogonality conditions comes from the properties of the optimal contract. According to definition of ε , the observed compensation can be written as

$$(19) \quad \tilde{w}_n = \frac{b_{1n}}{\rho(b_n - 1)} \ln(a'_2 z_n) + \frac{b_{1n}}{\rho} \ln \left[1 + \eta_n (a'_1 z_n)^{\frac{1}{(b_n-1)}} - \eta_n \frac{f_1(x_n(\gamma), \sigma, \psi(\gamma))}{f_2(x_n(\gamma), \boldsymbol{\theta}, \sigma, \psi(\gamma))} \right] + \varepsilon_n,$$

where η_n is the unique, strictly positive solution to the following equation in η .

$$(20) \quad \int [\eta (a'_1 z_n)^{1/(b_n-1)} - \eta \frac{f_1(x_n(\gamma), \sigma, \psi(\gamma))}{f_2(x_n(\gamma), \boldsymbol{\theta}, \sigma, \psi(\gamma))} + 1]^{-1} f_2(x_n(\gamma), \boldsymbol{\theta}, \sigma, \psi(\gamma)) dx = 1$$

Denoting the density of \tilde{w}_n conditional on z_n and x_n as $f_{\theta,\sigma,\psi,\gamma}(\tilde{w}_n | z_n, x_n)$, we can write the score with respect to θ for the likelihood of observing \tilde{w}_n as

$$(21) \quad h_{3n}(\boldsymbol{\theta}, x_n(\gamma), \sigma, \psi(\gamma)) = \nabla_{\theta} \ln f_{\theta,\sigma,\psi,\gamma}(\tilde{w}_n | z_n, x_n).$$

From the definition of a score,

$$(22) \quad E [h_{3n}(\boldsymbol{\theta}^0, x_n(\gamma^0), \sigma^0, \psi(\gamma^0))] = 0.$$

Our estimator for θ was found by forming a $q \times 1$ vector function $h_{4n}(\boldsymbol{\theta}, x_n(\gamma), \sigma, \psi(\gamma))$ from $h_{1n}(\boldsymbol{\theta})$, $h_{2n}(\boldsymbol{\theta}, x_n(\gamma), \sigma, \psi(\gamma))$ and $h_{3n}(\boldsymbol{\theta}, x_n(\gamma), \sigma, \psi(\gamma))$ and minimizing

$$(23) \quad \left[\frac{1}{N} \sum_{n=1} h_{4n}(\boldsymbol{\theta}, x_n(\gamma^N), \sigma^N, \psi(\gamma^N)) \right]' A_N \left[\frac{1}{N} \sum_{n=1} h_{4n}(\boldsymbol{\theta}, x_n(\gamma^N), \sigma^N, \psi(\gamma^N)) \right]$$

with respect to $\boldsymbol{\theta}$ subject to equation(20) which defines η_n , where A_N , which is a $q \times q$ matrix converging to some constant nonsingular matrix A , and the estimators $(\sigma^{(N)}, \psi(\gamma^{(N)}), x_n(\gamma^{(N)}))$ come from the first two steps.

Let

$$h_{4n}(\boldsymbol{\theta}, \boldsymbol{\theta}_1) = h_{4n}(\boldsymbol{\theta}, x_n(\gamma), \sigma, \psi(\gamma)),$$

where $\boldsymbol{\theta}_1 = (\gamma, \sigma)'$. Next, define

$$G_{\boldsymbol{\theta}} = E [\nabla_{\boldsymbol{\theta}} h_{4n}(\boldsymbol{\theta}, \boldsymbol{\theta}_1)],$$

$$G_{\boldsymbol{\theta}_1} = E [\nabla_{\boldsymbol{\theta}_1} h_{4n}(\boldsymbol{\theta}, \boldsymbol{\theta}_1)],$$

$$h_4(z) = h_{4n}(\boldsymbol{\theta}^0, \boldsymbol{\theta}_1^0),$$

$$M = E [\nabla_{\boldsymbol{\theta}_1} h_0(\boldsymbol{\theta}_1^0)],$$

and

$$\varphi_1(z) = -M^{-1}h_0(\boldsymbol{\theta}_1^0).$$

Under standard regularity conditions

$$\sqrt{N}(\boldsymbol{\theta}^N - \boldsymbol{\theta}^0) \implies N(0, V_1),$$

where

$$V_1 = (G'_{\boldsymbol{\theta}}AG_{\boldsymbol{\theta}})^{-1}E[\{G'_{\boldsymbol{\theta}}Ah_4(z) + G'_{\boldsymbol{\theta}}AG_{\boldsymbol{\theta}_1}\varphi_1(z)\}\{G'_{\boldsymbol{\theta}}Ah_4(z) + G'_{\boldsymbol{\theta}}AG_{\boldsymbol{\theta}_1}\varphi_1(z)\}'](G'_{\boldsymbol{\theta}}AG_{\boldsymbol{\theta}})^{-1}.$$

The result follows from applying Theorem 6.1 of Whitney Newey and Daniel McFadden (1994) to moments based on the first-order conditions of the minimization problem (23). In our application:

$$\text{plim}A_N = E\{h_4(z)h_4(z)'\}^{-1}.$$

REFERENCES

Gayle, George-Levi, and Robert A. Miller. 2008. “Identifying and Testing Generalized Moral Hazard Models of Managerial Compensation.” Tepper school of Business, Carnegie Mellon University.

Newey, Whitney K., and Daniel McFadden. 1994. “Large Sample Estimation and Hypothesis Testing,” in *Handbook of Econometrics*, ed. R. F. Engle and D. L. McFadden. 2111–245. Amsterdam: North-Holland.

Notes

*See Newey and McFadden (1994) for examples of these regularity conditions.