Incentive Pay and the Market for CEOs:

An Analysis of Pay-for-Performance Sensitivity

Charles P. Himmelberg and R. Glenn Hubbard[1]

March 6, 2000

Preliminary Draft

Comments Welcome

[1]Respectively, Columbia University, and Columbia University and the National Bureau of Economic Research. We have benefitted from very helpful discussions with Tim Baldenius, Marianne Bertrand, Ehud Houminer, Mike Kirschenheiter, Andy Newman, Mike Riordan, David Scharfstein, Nachum Sicherman, Paolo Siconolfi, Scott Stern, Julie Wulf, Amir Ziv, and seminar participants at the Federal Reserve Bank of New York and the Columbia Finance Free Lunch. We also thank Eric Engstrom for excellent research assistance.

Abstract

Significant increases in CEO compensation in recent years have coincided with an unprecedented bull market for stocks. They have also coincided with increasingly frequent arguments from academics and practitioners that CEO pay is excessive. Academics in particular point to departures from relative performance evaluation as a major puzzle. Observing that the sensitivity of CEO compensation to market returns is inconsistent with the predictions of principal-agent theory, they have increasingly been tempted to resort to political explanations to explain the compensation of CEOs. The sensitivity of CEO compensation to market returns is not so puzzling, however, when one recognizes the special feature of the labor market for CEOs. The combination talent sorting of CEOs among firms and aggregate shocks that affect the marginal value of CEO services to the firm predicts a positive relationship between CEO pay and aggregate stock market returns. We argue that the supply of highly skilled CEOs capable of running large, complex corporations is relatively inelastic, so shocks to aggregate demand for CEOs simultaneously raise firm value and raise the marginal value of CEO services to the firm. In equilibrium, such shocks bid up the value of their compensation packages. This phenomenon makes it appear as if (some) firms are violating relative performance evaluation.

Using Standard and Poor's Execucomp data for a large panel of U.S. firms, we show that the empirical "failure" of RPE is systematically restricted to the CEOs of certain types of firms. In our data, there is less of a puzzle for the firms in the bottom quartile of the size distribution. The behavior of these firms is in closer accordance with the predictions of agency theory. We interpret this as evidence that the supply of CEOs sufficiently skilled to manage such firms is relatively elastic. The compensation packages for the CEOs of larger, more complex firms, however, require highly skilled CEOs. Empirically, the compensation packages for the CEOs of these firms are more sensitive to aggregate shocks, which we interpret as evidence that the supply of highly skilled CEOs is relatively inelastic. These results are robust to robustness checks against alternative hypothesis such as "skimming" and strategic considerations.

Finally, the paper's findings suggest three extensions to i) analysis rent sharing between CEOs and firms, ii) comparisons of pay-performance sensitivities of CEOs and divisional managers, and iii) studies of cyclical patterns in earnings inequality.

1 Introduction

A cornerstone of the practitioner guidance regarding CEO compensation is the proposition that executives be paid for performance. Economists have used principal-agent theory to formalize this intuition, and have generally come to the same conclusion that optimal contracts for senior executives should display "relative performance evaluation," or RPE. Empirically, therefore, one would expect to find that CEOs are compensated based on firm-specific returns, filtering out observable aggregate shocks like movements in the S&P500 over which the CEO has no control.

In practice, however, evidence for RPE is mixed at best. Anecdotal evidence suggests that executive compensation has soared during the recent bull market in stocks. For example, in a recent survey of the literature on CEO pay, Abowd and Kaplan (1999, page 11) lament, "[w]hy shareholders allow CEOs to ride bull markets to huge increases in their wealth is an open question." Formal empirical studies support this view. Antle and Smith (1986) and more recently Bertrand and Mullainathan (1999) have argued that CEOs do, in fact, appear to be rewarded for "luck," and have interpreted the rise in CEO compensation during aggregate good times as evidence of pure rent extraction. Moreover, in recent reviews of the compensation literature, Murphy (1999), Abowd and Kaplan (1999), and Prendergast (1999) have identified the lack of RPE as a key unresolved puzzle. Indeed, the theoretical presumption in favor of relative performance contracts is sufficiently strong that many academics (e.g., Jensen and Murphy, 1990b) and public activists (e.g., Crystal, 1991) have taken the view that contemporary compensation practices are inefficient, and have consequently called on boards of directors to increase their reliance on relative performance evaluation for CEOs.[1] Critics have also questioned such practices as the repricing of stock options as further evidence that equity incentives are insufficiently indexed to the market

[1]Other research suggests that higher managerial ownership to align incentives is associated with higher corporate performance (see, e.g., Mørck, Shleifer, and Vishny, 1988; and McConnell and Servaes, 1990. For a contrary view, see Himmelberg, Hubbard, and Palia, 1999).

(see, e.g., Rappaport, 1999).

In this paper, we extend the standard agency view of CEO compensation and pay-performance sensitivities by incorporating insights from the literature on the distribution of earnings. We begin with the simple proposition that small increments in CEO talent can imply large increments in benefit to the firm due to the "scale of operations" under the CEO's control. This insight recognizes that large gains in wealth are generated by having more talented CEOs manage larger firms (see, *e.g.*, the review in Neal and Rosen, 1999). Under the general conditions identified by Rosen (1982), equilibrium is characterized by "positive assortative matching," meaning that high-skill CEOs match to large firms.[2]

While recognizing a scale-of-operations effect does not describe how match-specific rents should be allocated between the firm and the CEO, it does establish that some firms are, in principle, willing to pay substantial for incremental talent. To explain CEO pay, we need a model that explains the allocation of rents. To explain this allocation, we follow the principal-agent literature by making the standard assumption that CEOs (the agents) compete in the labor market to work for firms (the principals), and that the optimal contract in equilibrium is one that just meets the CEO's reservation utility (or outside opportunity). We incorporate these insights in our formulation of an empirical model with foundations in production theory and principal-agent theory to shed light on three questions. First, what explains the failure of RPE in the data? Second, what factors explain CEO rents (quasi-rents or pure rents)? Third, "How high is high" with regard to the estimated sensitivity of pay to performance?

With respect to these questions, we argue that empirical research on CEO compensation has given insufficient attention to the important role of CEO labor markets (despite the attention to labor-theoretic analysis by researchers like Rosen, 1992, for example). In the debate over RPE, for example, empirical investigations have not recognized and controlled

[2]The theoretical conditions under which positive assortive matching obtains have more recently been explored by McLaughlin (1994) and Shimer and Smith (1999), among others.

for the possibility that shocks to CEO labor demand might generate a positive correlation between aggregate stock returns and CEO compensation. The reason is that aggregate stock returns reflect, among other things, shocks to demand and productivity, and hence demand for managerial inputs. If managerial talent is scarce, then such shocks ought to increase the equilibrium wage paid to CEOs.[3] This effect could well offset the negative partial correlation (conditional on unfiltered performance measures) implied by relative performance evaluation.

Furthermore, aggregate shocks to the marginal value of managerial input have implications for the optimal compensation contract between stockholders and the CEO. If aggregate shocks raise demand for CEO "effort," and if the shocks can be contracted upon, then the contract will be designed to induce higher effort levels in response to such shocks. In the language of agency theory, the principal will demand more effort (or talent), and compensation levels have to rise to meet the agent's reservation utility; the magnitude of required increase depends on the shape of the CEO's utility function.

Optimal contracts also have to respond to the effect of aggregate productivity shocks on the level of agents' reservation utilities. In general equilibrium, a rise in CEO demand raises the value of outside opportunities for CEOs. This rise in the reservation utility raises the cost of to shareholders of eliciting managerial inputs, thus partially offsetting the initial increase that occurred in partial equilibrium, and further increasing equilibrium compensation levels.

We formalize these ideas in a simple principal-agent model of the optimal contract, which we then estimate using a large sample of CEO compensation data for U.S. firms. In addition to variables that measure the marginal productivity of the CEO's managerial input and relative performance, our specification introduces interactions between measures of the

[3]A related research program has addressed the extent to which workers share firm rents (Slichter, 1950; Dickens and Katz, 1987; and Katz and Summers, 1989). Analogous to the point we make in the context of executive compensation, Reder (1962) and Blanchflower, Oswald and Sanfrey (1996) argue that an upward-sloping short-run supply curve of labor can give rise to a positive correlation between wages and profitability. We return to the question of rent sharing in Section 4.

market demand for the CEO (the systematic component of the stock return) and measures of the supply elasticity of CEOs with similar skills. Consistent with the predictions of the model, we find that both firm-level, market-adjusted equity returns (a measure of relative firm performance) and the systematic component of the firm's return (a measure of the contractible portion of the firm's productivity shock) are important empirical determinants of executive compensation.[4] Moreover, we find that the sensitivity to the systematic component of the firm's return is higher for CEOs with skills that are inelastically supplied by the market, namely, CEOs of large firms. These results are consistent with the view that aggregate productivity shocks increase demand for CEO labor, forcing some firms to increase compensation levels to retain their CEOs.

We consider alternative explanations for these facts, but find little support for them. For example, we find little evidence for the view that the sensitivity of CEO compensation to market returns principally reflects pure rents, or "skimming," attributable to frictions in the market for corporate control. Second, we show that market sensitivity does not appear to reflect strategic considerations in the product market of the sort considered by Fershtmann and Judd (1987) and Aggarwal and Samwick (1999). We conclude that the evidence is consistent with a model of CEO compensation in which the optimal contract reflects not only the usual agency considerations, but also the firm's desire to retain high-quality CEOs in the face of aggregate demand shocks for CEO talent.

Our empirical specification for the compensation contract is derived from a formal principal-agent model, but for empirical reasons which we explain below, we depart from the standard specifications of utility and production. First, we model CEO preferences using power utility instead of the usual assumption of negative exponential function. This assumption endows CEOs with constant relative risk aversion (CRRA) rather than constant absolute risk aversion (CARA) preferences, thereby eliminating scale effects on risk aversion

[4]Murphy (1985), Gibbons and Murphy (1990), and Barro and Barro (1990) also use asset pricing models to identify the sensitivity of CEO compensation to the systematic component of stock returns.

and making it more plausible to assume that our empirical model can accommodate cross-sectional variation CEO wealth. Second, we model the production relationship between managerial input and firm value using Cobb-Douglas technology instead of the usual linear technology. This assumption is more natural for thinking about the effects of firm size and firm-level productivity on the marginal productivity of CEO inputs, and it also implies a downward-sloping demand curve for CEOs. This implication for CEO labor demand is an essential component of our equilibrium explanation for the response of reservation utilities to aggregate shocks.

In exchange for the desirable empirical properties of our functional form assumptions, we pay the analytic price of a typically nonlinear optimal contract which has no closed-form solution. However, there are good theoretical reasons for these nonlinearities, and our evidence suggests it is important to preserve these nonlinearities in the empirical specification. Hence instead of following the conventional approach of solving for the "best linear contract," we characterize the contract that solves a second-order Taylor series approximation to the shareholder's contract design program. Not surprisingly, the optimal contract to the approximate problem is nonlinear, too, but it has a closed-form solution, and with a log transformation, it turns out to be is approximately linear in total compensation and gross returns. This log-linear empirical specification fits the data well.

The paper is organized as follows. In Section 2, we derive our model of executive compensation. Section 3 reports our empirical results reporting the sensitivity of CEO compensation to firm performance. Section 4 concludes.

2 An Empirical Principal-Agent Model of CEO Compensation

To guide to our empirical specifications, we develop of principal-agent model of CEO compensation. The model setup is standard (only our functional form assumptions and solution method will differ).[5] The stockholders (principals) are risk-neutral, and they structure a compensation package $w(V, \hat{x})$ which is designed to elicit value-maximizing level of effort $\hat{x}$ from a risk-averse CEO who has utility over wealth and work effort given by $u(w, x)$, where $w(V, \hat{x})$ is CEO wealth tied to the value of the firm, V, and x is work effort. The optimal contract cannot depend on effort for the usual reason that effort is assumed to be unobservable. Thus the optimal contract is contingent on firm value, and must satisfy an incentive constraint, as well as a participation constraint that guarantees a minimum level of utility to the agent. Formally, the principal's problem is to

$$\max_{w,\hat{x}} E\left[V - w(V, \hat{x})\right] \tag{1}$$

subject to the constraints

$$E\left[u(w(V), \hat{x})\right] \geq \bar{u} \qquad \text{(Participation Constraint)} \tag{2}$$

$$\hat{x} \in \arg\max_{x} E\left[u(w(V), x)\right] \qquad \text{(Incentive Constraint)} \tag{3}$$

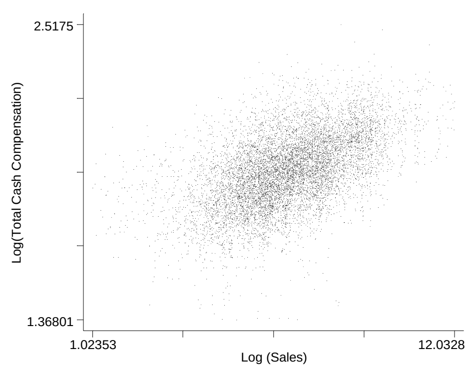
where the expectations are taken with respect to the distribution of V induced by $\hat{x}$, and $\bar{u}$ represents the agent's reservation utility.

To characterize the optimal contract, one must make assumptions about functional form for preferences and production. Most empirical studies follow the approach of Holmstrom and Milgrom (1983) by constraining the principal to use the best *linear* contract (see, for example, Garen, 1994; Haubrich, 1994; Baker and Hall, 1999; and the literature reviews by Murphy, 1999; and Prendergast, 1999).

Despite its intuitive appeal, the "best linear contract" approach suffers two drawbacks.

[5]Economic analysis of executive compensation has focused principally on implications of agency-theoretic problems in the separation of ownership and control in business corporations, highlighted initially by Berle and Means (1932) and formalized by Jensen and Meckling (1976), Mirrlees (1974, 1976), Holmstrom (1979, 1982), Fama (1980), and Grossman and Hart (1983), among others.

First, as is well known, optimal contracts are generally nonlinear.[6] Second, and more importantly, the empirical relationships observed in the data are frequently nonlinear. Indeed, much of the debate over the magnitude of the pay-for-performance sensitivities estimated by Jensen and Murphy (1990) and others reflects confusion over the appropriate functional form.[7] Murphy's (1999) recent review of the literature, for example, emphasizes a lack of consensus on the appropriate functional form to use. In our experience, the log specification works best. Figure 1 plots the log of total cash compensation against the log of firm sales (the log of CEO wealth has a very similar appearance). This plot reveals the well-documented relationship between size and compensation, and reveals the primary motive behind our desire to model log compensation rather than levels.[8] We also show that this functional form follows naturally from the optimal contract when utility is CRRA and production is Cobb Douglas.



Log CEO Comp vs. Log Sales

In the next three subsections, we motivate our functional form assumptions about preferences and technology, then discuss the approximately optimal contract which provides the structural justification for our empirical specification. We conclude with several short

[6]But see Holmstrom and Milgrom (1987).

[7]In addition, as Haubrich (1994) has correctly pointed out, it is difficult to interpret pay-for-performance sensitivities without a model. He addresses the analytical difficulties of the contracting problem by using numerical solutions. The strategy in our paper is similar in spirit, in that our approximations allow economic interpretation of parameter estimates.

[8]See Rosen (1992) and Murphy (1999).

discussion of the most important issues regarding empirical implementation, namely, identifying shocks to CEO labor market demand, finding proxies for CEO talent, and measuring relative performance evaluation.

2.1 Preferences

We assume that the utility of the agent is isoelastic with respect to wealth and effort, so that

$$u(w_{it}, x_{it}) = \frac{1}{1-\gamma} w_{it}^{1-\gamma} x_{it}^{-\delta(1-\gamma)}, \tag{4}$$

where w_{it} and x_{it} denote the agent's wealth and "work effort," as described early. With these preferences, agents display constant relative risk aversion (CRRA). To generate linear contracts, it is more common in the agency literature to assume constant absolute risk aversion (CARA). While both assumptions are plausible, CRRA is more widely used in models of consumption and saving (see, e.g., Hubbard, Skinner and Zeldes, 1995), as well as consumption-based empirical models of asset prices (see, e.g., Campbell, Lo, and MacKinlay, 1997). We follow the literature by assuming that utility is increasing concave in wealth, and decreasing concave in effort.[9]

Our functional-form assumption in equation (4) implies CRRA utility preferences, which have the attractive feature that, unlike CARA preferences, they do not imply large cross-sectional differences in risk aversion simply because of level effects of CEO wealth and salary differences. The resulting empirical specification of the optimal contracts is therefore better-suited for cross-section and panel data.[10]

[9]To avoid the technical complications associated with the case $\gamma = 1$, the derivations in the text assume $\gamma > 1$ (a common calibration choice is a value like $\gamma = 3$.) To impose the assumption that utility is decreasing and concave in effort, we assume $-\delta(1-\gamma) > 1$. A plausible set of parameters, for example, would be $\gamma = 3$ and $\delta = 1$.

[10]The assumption of power utility is popular in the consumption-based asset pricing literature for similar reasons, namely, that it avoids changes in risk aversion due to long-run economic growth, that is, the model scales as wealth increases. We also imitate this literature in our solution methods for the model by using

2.2 Technology

We assume the relationship between managerial input and firm value is given by

$$V_{it} = \theta_{it}(\tau_i x_{it})^{\phi} K_{it}, \tag{5}$$

where $0 < \phi < 1$ is the elasticity of firm value with respect to effort, τ_i and K_{it} are exogenous parameters indexing managerial talent and firm size, respectively, and θ_{it} is a lognormally distributed shock to firm value, the log of which follows a random walk with stochastic drift parameter $\beta_i' m_t$, where m_t represents a k-vector of observable, exogenous, aggregate shocks, and β_i represents a conformable vector of "factor loadings." That is, $\ln\theta_{it}$ evolves according to $\ln\theta_{it} = \ln\theta_{it-1} + \varepsilon_{it}$,where the distribution of ε_{it} (conditional on m_t) is $N(\beta_i' m_t, \sigma^2)$. The specification of the drift parameter introduces conditioning information (namely m_t) over which the agent has no control, but which can nevertheless be used by the principal to reduce the variance of the contract, hence reducing the cost of contract to the principal. We assume τ_i, K_{it}, β_i, and m_t are observable to the principal, and therefore contractible, but θ_{it} and x_{it} are not. Substituting for θ_{it} in equation (5) and log-differencing, we get

$$\Delta \ln V_{it} = \phi \Delta \ln x_{it} + \Delta \ln K_{it} + \beta_i' m_t + \varepsilon_{it}. \tag{6}$$

This specification linking the shocks to firm value allows for relative performance evaluation. We interpret $\beta_i' m_t$ as a factor-model representation of the systematic component of θ_{it}, the shock to firm value. Such factors will optimally be used by the principal to reduce the expected cost of the contract by reducing the CEO's exposure to risk. In theory, m_t could be any observable, exogenous index of industry or market performance that is correlated with θ_{it}, but not influenced by the CEO. In practice, however, we are constrained (as

second-order Taylor series approximations to derive analytic expressions for risk aversion. See, for example, the discussion of the equity premium puzzle in Campbell, Lo and MacKinlay (1997, ch. 8).

are, to some extent, boards of directors) by the need to work with empirical measures whose correlation with firm-level shocks can be estimated. This rules out low-frequency measures (e.g., GDP growth, or industry profitability) because we have seven or fewer years of annual, firm-level data with which to calculate the firm-specific factor loadings, β_i. Stock returns, by contrast, are readily available at the monthly frequency, which makes it possible to estimate firm-specific covariances of the aggregate shock with the firm-level shock. Assuming shocks to aggregate stock portfolios reflect real aggregate shocks that are correlated with the shock to firm value, this approach provides a credible measure of relative firm performance for CEO evaluation.

Our assumption that $0 < \phi < 1$ in equation (5) implies that firm value (gross of compensation costs) is an increasing and concave function of managerial input ("effort"). While a linear specification would simplify the analysis,[11] such a specification arguably fails to capture important empirical features. Perhaps most importantly, linear technology does not capture what Rosen (1982, 1992) has referred to as the "scale of operations" effect, in which the talents of the CEO are multiplied by the large scale of operations under his or her control. According to such models, even very small differences in CEO talent (or effort) can have very large effects on the absolute level of firm value. In the sorting equilibrium of talent across firms, high-τ_i CEOs would presumably sort into large firms (i.e., firms for which the value of talent is the highest). Thus the scale-of-operations model predicts the strong positive correlation between CEO compensation and firm size that we observe in the data.[12]

Our technology assumption incorporates the scale-of-operations effect by letting that

[11]It is more common to assume a linear technology of the form $V_{it} = a + \psi x_{it} + \varepsilon_{it}$ (see, e.g., the studies reviewed by Prendergast, 1999).

[12]Because of the "scale of operations" effect, gains in wealth are generated by having more talented CEOs manage larger firms (see, *e.g.*, the review in Neal and Rosen, 1999). While this assumption does not describe how match-specific rents should be allocated between the firm and the CEO, it does establish that some firms are, in principle, willing to pay substantial for incremental talent. Moreover, under the general conditions identified by Rosen (1982), we can assume that the equilibrium we analyze is characterized by positive assortative matching so that high-skill CEOs match to large firms, an idea extended to conditions of rent sharing by McLaughlin (1994).

the value of CEO inputs, $(\tau_i x_{it})^\phi$, interact multiplicatively with fixed assets, K_{it}, and the shock to firm value, θ_{it}.[13] That is, not only the total but also the *marginal* value of CEO inputs is increasing in firm size and the value shock. Therefore aggregate shocks m_t to the firm's value shock θ_{it} can shift the principal's demand for both CEO talent and effort. If the aggregate shock increases demand, and CEO talent is scarce, then the shock raises the equilibrium level of the CEO's reservation utility, causing CEO compensation to rise. This result is "surprising" only from the perspective of a "bare-bones" principal-agent model in which the marginal value of the CEO is independent of aggregate shocks. Our technology assumption, by contrast, generates a more realistic response to productivity shocks, namely the increased demand for talent and effort, and embeds this feature in the optimal contract.

2.3 The Optimal Compensation Contract

Under our assumptions about preferences and technology, the optimal contract is

$$\ln w_{it} = \ln A_{it} + \frac{\delta}{\phi} \ln (1 + r_{it}), \tag{7}$$

where r_{it} is the measure of the firm's relative performance. The relative performance measure is given by

$$r_{it} = \ln V_{it} - E_{it} \left(\ln V_{it} \,|\, \tau_i, K_{it}, \beta_i' m_t \right), \tag{8}$$

where the expectation on the right side of equation (8) is conditional on observable variables in the contracting environment.

We define "pay-for-performance sensitivity" as the slope coefficient on $\ln(1 + r_{it})$. This slope depends only on the parameters ϕ and δ, which we assume are constant across CEO's

[13] To bring size effects into the predicted compensation contract, Schaefer (1998) and Baker and Hall (1999) modify the standard linear technology (i.e., $V = a + \psi x + \varepsilon$) by letting the coefficient on effort, ψ, depend on size.

and over time. The intuition for the slope is that effort-averse CEOs (higher δ's) have to be given higher-powered incentives, while firms for which the marginal value of effort is high (higher ϕ's) can offer lower-powered incentives because the technology implies the marginal payoff to the CEO of a marginal unit of labor is high.

The contract parameter A_{it} generally depends on exogenous features of the contracting environment. In the empirical specification, we let A_{it} be depend on variables in the data because some of these features vary cross-sectionally and over time. One can show (see Appendix A) that the parameter A_{it} is given by

$$\ln A_{it} = c_0 + \left(\tfrac{1}{1-\gamma}\right)\left(1 - \tfrac{\delta}{\delta-\phi}\right)\ln \bar{u}_{it} + \left(\tfrac{\delta}{\delta-\phi}\right)\left(\beta_i' m_t + \phi \ln \tau_i + \ln K_i\right), \tag{9}$$

where c_0 is a constant which depends on other constant parameters in the contracting environment.

The aggregate component $\beta_i m_t$ of the firm's value shock enters the agent's contract three ways. First, it increases the mean of the conditional distribution of log performance, $\ln V_{it}$, but $\beta_i m_t$ itself is not affected by the CEOs effort. Hence in the optimal contract, it is optimal to contract on a performance measure that has been "benchmarked" (or "indexed") relative to $\beta_i m_t$. Second, it enters the intercept term $\ln A_{it}$ in equation (7). This is because the aggregate component of the firm's equity return has been filtered out. Other things equal, this filtering reduces the CEO compensation, but more work effort is desired when productivity is high, so compensation is increased to meet the participation constraint via the component of the log contract that is not contingent on the performance measure.

Third, the aggregate shocks m_t potentially change the contracting environment by driving up the CEO's reservation utility. In the next section, we provide an intuitive motivation for an empirical specification of the reservation utility using proxies for CEO supply elasticities and labor demand shocks.

2.4 Empirical Specification

Substituting equation (9) into equation (7) yields

$$\ln w_{it} = c_0 + c_1 \ln \bar{u}_{it} + c_2 \left(\beta_i' m_t + \phi \ln \tau_i + \ln K_i\right) + \frac{\delta}{\phi} \ln\left(1 + r_{it}\right) \tag{10}$$

where

$$\begin{aligned} c_1 &= \left(\tfrac{1}{1-\gamma}\right)\left(1 - \tfrac{\delta}{\delta-\phi}\right), \\ c_2 &= \left(\tfrac{\delta}{\delta-\phi}\right). \end{aligned}$$

In general, principal-agent models do not make strong predictions about the functional form of the optimal contract, and while it is common in the theoretical literature to derive results by assuming linear contracts,[14] linear specifications are not always used by empirical researchers. This gap between theoretical and empirical specifications is reflected in the variety of specifications (including linear and log-linear specifications) that have been used by empirical researchers in labor economics and corporate finance (see the reviews in Murphy, 1999; and Prendergast, 1999). Because our equation for the optimal contract is derived from empirically plausible utility and technology assumptions, we can interpret it more literally to guide our empirical specification. Like most studies of wages, our model suggests using a log-linear relationship between wages and their determinants. Compared with levels-on-levels specifications, the log-linear functional form has the added advantage that it tends tends to reduce heteroskedasticity, and produces convenient estimates of elasticities

[14] The model assumptions used to derive our log-linear contract specification are motivated primarily by empirical considerations, and are not essential to the intuition behind most of our theoretical predictions. For example, increases in reservation utility increase the intercept in the linear model, whereas they increase the log intercept in ours. The one dimension in which our model would differ from models with linear technology is in the predicted response to observable technology shocks. In our model, such shocks are used to measure relative performance, but they also increase the mean of (log) compensation because increase the marginal product of CEO effort. In the linear model, by contrast, technology shocks do not affect the marginal product effort.

and semi-elasticities rather than slopes.

2.4.1 The Specification of Reservation Utility

The expression for $\ln A_{it}$ in equation (10) includes terms which formalize the intuition that increases in reservation utility increase the CEO's compensation. In particular, inspection of equation (9) reveals that $\ln A_{it}$ depends linearly on $\ln \bar{u}_{it}$. Thus, changes in the reservation utility change the (log) mean of the CEO's compensation. We model these effects empirically by assuming that the equilibrium reservation utility (which is taken exogenously by the firm and the CEO) can be modelled by the specification

$$\bar{u}_{it} = \Lambda_i z_{it},$$

where z_{it} denotes "labor demand shocks," and Λ_i measures the sensitivity of $\bar{u}_{it}$ to these shocks.

We assume the labor demand shock z_{it} can be proxied by the systematic component of the firm's stock return, that is, $\beta_i' m_t$. A more precise measure of the labor demand shock would aggregate labor demand shocks (stock returns) over the subset of firms with similar demands for CEO talent, but this measure requires a metric of the distance between firms in "CEO talent space" which is difficult to identify a priori. (We are currently pursuing this idea.)

We model the firm-level sensitivity parameter, Λ_i, as a function of CEO talent proxies. With respect to our labor market story, the most important of these proxies is firm size. Our key assumption is that the elasticity of aggregate CEO talent and labor supply is less elastic for highly talented CEOs. Thus we expect CEO compensation to display a higher aggregate stock return elasticity of compensation for larger firms.

2.4.2 Proxies for CEO Talent

The term $\ln A_{it}$ also includes the log of CEO talent, $\ln \tau_{it}$, which is unobserved and therefore requires an empirical proxy. Our preferred measure is firm size. The justification for using firm size as a proxy for CEO talent comes from the theoretical assumption that there is a matching equilibrium that matches the most talented CEOs with the firms for we implicitly assume that in equilibrium, the market sorts high-talent CEOs into firms where there marginal value is the highest.[15] (Note that in contrast to reservation utility, firm size appears in the intercept rather than the interaction with stock market variables.)

To control for talent ($\ln \tau_i$) in the mean of the compensation regression, we also use with price-cost margins, R&D intensity, advertising intensity, and tenure. Price-cost margins are used on the grounds that high-margin firms likely earn quasi-rents attributable to intangible firm assets. If these assets represent unique firm assets that generate quasi-rents, then in the same way that large firms demand CEOs with superior talent, the sorting equilibrium will assign high-talent CEOs to firms with high margins.[16] The other proxies are standard.

2.4.3 Measures of CEO Compensation

Hall and Liebman (1998) have recently stressed that capital gains on stocks and options dominate the variance of CEO wealth; variation in salary and bonus is modest by comparison. Following Hall and Liebman (1998), we measure the CEO's wealth, w_{it}, in the firm at

[15] With panel data, one can use CEO fixed effects to control for the omitted variable bias introduced by τ (see, e.g., the discussion in Himmelberg, Hubbard, and Palia, 1999). In this paper, we rely on the sorting argument in the scale-of-control model to provide an empirical justification for using size as a proxy for CEO talent in compensation regressions.

[16] A possible exception to this assumption is that high margins could simply reflect pure rents, such as, for example, a government franchise to operate a monopoly. In this example, high margins might even be negatively correlated with CEO talent if monopolies require less talent to run profitably. On the other hand, it may require the CEO's (political) skill to maintain the monopoly franchise. In any event, we doubt that such cases are empirically important. In one of our robustness checks below, we use a Herfindahl index to investigate the exogenous effects of concentrated markets on the compensation contract, and find that this potential source of exogenous rents has little impact on the level or performance sensitivities of CEO contracts.

the end of time t using

$$w_{it} = VEQUITY_{it} + VOPTIONS_{it} + COMP_{it},$$

where $VEQUITY_{it}$, $VOPTIONS_{it}$, and $COMP_{it}$ respectively denote the (end-of-period) market value of equity, the (end-of-period) market value of options, and current year's cash compensation. Cash compensation includes the total current flow net of the value of stock and option grants, and thus includes salary, bonus, contributions to retirement benefits, and other misc flows. The value of equity and stock option granted during the current fiscal year is included in $VEQUITY_{it}$ and $VOPTIONS_{it}$. The present value of expected future cash compensation is missing from this measure, so we also construct a crude proxy by multiplying cash compensation by three.[17]

[17]This treatment of CEO wealth does not consider private wealth, and it does not consider the how private wealth holdings would be optimally structured to hedge systematic risk from wealth held in the firm. Private wealth might be handled as follows. Suppose CEO has private wealth w_{it}^P. Utility is be defined over total CEO wealth, which is private wealth plus wealth in the firm, w_{it}, or

$$w_{it}^T = w_{it}^P + w_{it}.$$

This presents a measurement issue for wealth which can be dealt with in the following way. First, there exists a c such that

$$\ln w_{it}^T \simeq c \ln w_{it}^P + (1-c) \ln w_{it}.$$

Suppose the true specification is $\ln w_{it}^T = bx_{it} + u_{it}$. Then using the approximation, we can re-specify the model in terms of $\ln w_{it}$:

$$\ln w_{it} = \frac{1}{1-c} \left(-c \ln w_{it}^P + bx_{it} + u_{it} \right).$$

Note that if personal wealth is negligible relative to the CEO's wealth in the firm (which seems likely), then $c \simeq 0$. If this is not the case, then we need to proxy for $\ln w_{it}^P$ on the right side of the equation. By appealing to the equilibrium sorting of high talent to large firms, private wealth should be correlated with talent, and hence firm size.

3 Empirical Results

3.1 Data Description

Our data is constructed by taking executive compensation data from Standard and Poor's ExecuComp and matching it with firm-level data from Compustat, monthly stock return data from CRSP, and industry data from the U.S. Department of the Census. The data cover the period 1992 to 1998, and the number of firms for which we could assemble a set of variables ranges 638 firms at the beginning of the sample to 1084 in 1996 and then tapering off some to 1045 at the end of the sample. The attrition at the end of the sample is caused by missing data for required Compustat and CRSP variables (most the latter). Table 1 defines the variables we use in our empirical analysis, and Table 2 presents summary statistics for the variables. Additional information on variables and their construction is contained in Appendix B.

In Table 3, we report median values of components of CEO compensation for each of the years in our sample, 1992-1998. The first column of medians presents firm sales. The second column presents median values of salary and bonus plus other cash compensation ("*S&B*"). Median salary and bonus rises only modestly in real terms over the period, but compensation rises more substantially when new grants of equity and of stock options ("*Grants*"), as shown in column 3. Given the strong performance of the stock market over this period, it is not surprising that the total value of equity held ("*Equity Value,*" in the fourth column) and especially options held ("*Option Value,*" in the fifth column) rise by even more. The last two columns of Table 3 present (median values of) two measures of CEO wealth in the firm. "*Total Wealth (A)*" is the sum of *S&B, Equity Value, and Option Value.* *"Total Wealth (B)"* attempts an annuity valuation of salary and bonus, and equals the sum of equity value, option value, and three times the current value of cash compensation.

3.2 Model Estimates

The model derived in section 2 emphasizes CEO wealth in the firm as the measure of w_{it}. Because empirical studies of executive compensation have used both flow and stock measures of CEO compensation, however, we present results using both definitions: "cash compensation" and "total wealth." Our baseline specification is the log specification in equation (10). Before examining the differential effect of relative and market returns on CEO compensation, we first present versions of equation (10) in which idiosyncratic and market returns are presented together as total returns. Tables 4 and 5 summarize our baseline results for the full sample, and also report year-by-year results to show that the basic patterns are robust over time. The determinants of the log of cash compensation and the log of total CEO wealth in the firm, respectively, as defined in Table 1. Both CEO compensation and wealth rise with proxies for talent requirements – log sales, tenure, price-cost margin, R&D intensity, advertising intensity, degree of firm diversification, and the existence of foreign operations are all associated with higher average CEO compensation and wealth. Moreover, compensation and wealth are both higher for firms with more volatile returns (as measured by $LOGSIGMA$).[18]

The last two regressors in Tables 4 and 5 report the impacts of relative (RPM) and market (MKT) components of firm returns on the log of cash compe nsation and total wealth, respectively. The results reported in these tables use rolling window of monthly stock returns for the five years prior to the current year to estimate an asset-pricing model for the firm. We tried three different model and chose the one that gave the highest average R^2, which was a four-factor model based on the following portfolios: the CRSP value-weighted market return (VWM), the return on a portfolio of 90-day Treasury bills (T90),

[18] In additional specifications (not reported), we also included a measure of accounting returns – return on assets – to investigate the possibility that firms' giving weight to accounting rather than market returns would lead to an omitted variable bias. However, after controlling for firm equity returns and CEO skill characteristics, this variable did not have a statistically significant impact on CEO cash compensation or wealth. We also do not report year dummies, which indicate that conditional on our independent variables, CEO cash compensation and total CEO wealth both rise on average about 10 percent per year.

an value-weighted industry portfolio made up of firms in the same 2-digit SIC, but excluding the own firm (IND), and a value-weighted size portfolio made up of firms in the same size decile which again excluded the own firm (SIZ). In Table 6 we report similar specifications using the runner-up models, which were the standard CAPM (using VWM and T90) and the model based on the portfolios advocated by Fama and French (1993).

Of particular interest in Tables 4, 5 and 6 is the influence of the relative performance measure versus the systematic component of the firm's stock return. Consistent with RPE, compensation increases with idiosyncratic (relative) performance. The full sample results reported in Tables 4 and 5 reveal that a one-percentage-point increase in relative performance raises CEO cash compensation by about 0.13 percent, and CEO wealth by 0.39 percent. Inconsistent with conventional formulations of RPE, however, CEO compensation is positively related to market returns; indeed, increases in market returns increase CEO cash compensation and wealth by 0.46 and 1.87 percent, respectively, which is substantially higher than the effect of relative performance.[19] These findings corroborate those of earlier studies which reject simple models of relative-performance-only evaluation.

3.2.1 Estimates of the Labor Demand Effect

Our model predicts market returns should affect CEO compensation to the extent that such returns increase the market value of his human capital. To test this prediction, we postulate that the sensitivity to observable demand shocks (e.g., value-weighted market returns) should be increasing in the value of the CEO's human capital. That is, we test the hypothesis that the supply of CEOs who can run large, complex organizations in the Fortune 50, for example, is less elastic than the supply of CEOs who can run small-capitalization firms traded on NASDAQ.

[19] The explanatory power of βm clearly does not reflect increased pay to compensate the CEO for holding more risk. Without dampening incentives, the firm could easily unload this risk using RPE, and reduce its expected compensation expense (holding CEO utility constant). In equilibrium, therefore, we would not expect to observe payments for holding market risk.

Using firm size as our preferred proxy for CEO talent, we estimate equation (14) for each quartile of the size distribution. Results are reported in Table 7 for the four-factor model (VWM/T90/IND/SIZ), estimated by pooled OLS. The four panels of Table 10 present the quartile estimates for the first through fourth quartiles. For both compensation measures, the pattern is striking: While sensitivity of compensation to idiosyncratic shareholder returns varies somewhat across the size groupings, variation in the sensitivity to market returns is much more pronounced for larger firms. As firm size increases, the sensitivity to market returns becomes large and statistically significant, suggesting that the supply of CEOs who are capable of running such firms is relatively less elastic. For CEO wealth, the coefficients on MKT across size quartiles are 1.74, 1.62, 2.04 and 2.68. Though we do not formally test the difference in estimates, the standard errors are on the order of 0.10 to 0.18, which strongly hints at the outcome of such a test.

The results in Table 8 provide a clearer picture of the statistical significance of the coefficient differences across size classes, and also report the results of additional robustness tests. Columns 1 of Table 8 reports estimates of a model for log cash compensation that includes interaction effects between MKT and dummy variables for the size quartiles. The coefficients on the size interactions for quartiles 2, 3 and 4 are 0.002, 0.610 and 1.170, respectively, with standard errors of 0.134, 0.151 and 0.178. Hence we can construct a simple t-test and see that there is substantial evidence against equality of the third and fourth quartile coefficients with the first. The fourth column reports similar results using log CEO wealth in place of cash compensation and finds similar results.

3.2.2 Are there Frictions in the Market for Corporate Control?

The empirical results presented thus far establish two clear findings. First, perhaps contrary to more impressionistic evidence, relative performance evaluation is an important feature of CEO compensation. Second, CEOs in firms in which managerial talent requirements are high are paid more on average and are paid more when market returns are high. These two sets of results are consistent with models of equilibrium sorting of CEOs across firms and with agency-theoretic models in which some executives reservation wage is affected by changes in market returns.

In a recent paper, Bertrand and Mullainathan (1999) have suggested that departures from relative performance evaluation represent evidence of pure rents going to CEOs, which they term "skimming." In their interpretation, beneficial aggregate shocks produce pure rents within the firm, and frictions in the market for corporate control imply that CEOs are in a position to "skim" a large fraction of these rents for themselves.[20]

We do not think this view can easily explain our results. First, it is not clear why skimming would be related to capital gains on the firm's stock price rather than free cash flow within the firm. Our specification of the mean compensation includes a measure of price-cost margins, which we use to proxy for the firm's talent demand, but which should also be closely related to the free cash flow available within the firm. Hence, if one reinterprets margins as a control variable for the degree of skimming, it is still the case that compensation rises with aggregate shocks.

Our view provides an alternative to the skimming view which is still consistent with the data, but interprets high levels of CEO compensation as *quasi* rents attributable to scarce CEO talent rather than as *pure* rents attributable to frictions in the market for corporate control. Bertrand and Mullainathan (1999) have suggested an interesting empirical strategy for distinguishing the two stories. Under the skimming view, the degree of skimming should be bounded by the magnitude of frictions in the market for corporate control. This has empirical implications if one can identify cross-sectional or time-series differences in the size of the friction in the market for corporate control. One can casually implement this suggestion by considering the evidence provided by the changes over the past 20 years, during which the market for corporate control has arguably become more liquid. Over the

[20]Similar assessments have been offered by practitioners and public activists like Crystal (1991).

same period, CEO compensation levels appear to have risen more rapidly than in the past. This observation poses a problem for the skimming view, but it works well for our labor market view because competition in the labor market for CEOs would tend to raise quasi-rents, much as the introduction free agency appeared to increase the compensation of star basketball players in the National Basketball Association.

We investigate whether our observed sensitivity of CEO compensation to market returns is driven by skimming by including in equation (10) proxies for the degree of frictions in the market for corporate control. Two logical proxies are (i) a dummy variable indicating whether the CEO was promoted from the inside or hired from the outside (OUT),[21] and (ii) the quartile dummies based on the CEO's tenure with the firm (TQ2-TQ4). In the former case, skimming would arguably be more likely for internal promotions by a captive board. Barro and Barro (1990), for example, have argued that if there is skimming, it should be less severe among newly hired CEOs. The second proxy is motivated by the idea that long a tenure gives the CEO more opportunity to co-opt the board of directors (or, perhaps more convincingly, a long tenure reveals ex post that the CEO has somehow become entrenched). Columns 2, 3, 6 and 7 of Table 8 report results for an expanded version of equation (10) to which the "skimming" proxies have been added. The estimated coefficients, however, reveal no significant difference in the market sensitivity to the OUTSIDER dummy.

The interaction of the market return with the tenure dummies TQ2, TQ3 and TQ4 are sometimes significantly different from zero, but the evidence is difficult to interpret as skimming. First, the interaction is negligible for RPM for both cash compensation and CEO wealth. It is not clear why the CEO would be able to skim only out of market returns. Second, two of the three coefficients on the interaction with MKT are not statistically different zero in the regression for cash compensation. The coefficients for log wealth are all positive and significant, but again, this is not the pattern we would expect if there were

[21] The dummy variable OUTSIDE is set equal to one if the CEO had been with the firm for fewer than four years at the time of his promotion to CEO, and equal to zero, otherwise.

skimming. Skimming would presumably be easier with cash transfers rather than market returns, but this is not what the results indicate. These findings suggest that *quasi*-rents are more important than *pure* rents in explaining the sensitivity of CEO compensation to shifts in market returns. A more likely explanation for the pattern of the tenure interactions with MKT is that CEOs are accumulating equity stakes and stock options the longer they work for the firm, so that their sensitivity to market returns is increasing.

The growth in the relative importance of equity-based over cash compensation over the term of the CEO's tenure is a robust feature of the data. Columns 2 and 5 of Table 8 show that as CEO tenure increases, both the mean level of CEO wealth and the sensitivity of this wealth to MKT appear to increase. One possible explanation for this pattern is that the relationship between the CEO and the firm involves idiosyncratic investment by the CEO. Since this creates a hold-up problem, the firm may try to commit to paying higher compensation when the CEO's outside opportunities rise by paying the CEO in equity rather than cash. As CEO investments in firm-specific human capital succeed in generating match-specific rents, increases in outside opportunities (proxied here by MKT) increase the CEO's bargaining position, and hence compensation. This provides another reason dimension to our emphasis on the CEO's outside options which helps to explain why firms could prefer not to index compensation granted in the form of stocks and options.

3.2.3 Incentives to Behave Strategically in Product Markets

Another alternative explanation for our findings is that the relative performance evaluation benchmark fails to recognize strategic interaction among firms in the product market (see, e.g., Aggarwal and Samwick, 1999). In concentrated markets, it may be optimal to soften relative performance incentives to maximize joint (industry) returns. Such stories have intuitive appeal, but they are difficult to test because large firms typically operate in many industries and lines of business.

Nevertheless, to test the robustness of our results to strategic considerations, we use the Herfindahl index for the firm's four-digit SIC industry as a measure of industry concentration. Columns 4 and 8 and Table 8 report results for an expanded version of equation (10) to which the Herfindahl index is added; for these tests, we use a smaller sample that includes manufacturing firms only. For these firms, the U.S. Census Bureau calculates the Herfindahl-Hirschmann index at the four-digit level.[22] Following on the interaction of the Herfindahl index of concentration, we find a weakly negative relationship between concentration and the sensitivity of CEO compensation to market returns to shareholders. But the inclusion of this variable has virtually no effect on our size interactions. While strategic interactions among firms may influence executive compensation, they do not appear to explain the pattern of market sensitivities across firm size.[23]

3.3 Is there Relative Performance Evaluation?

We have argued that the mere sensitivity of compensation to market returns or any other aggregate performance measure is not per se evidence that RPE is violated. Executive compensation should be sensitive to aggregate shocks if such shocks reflect changes in the marginal productivity of managerial services.[24] By contrast, Bertrand and Mullainathan (1999) use industry and market-level variables as instruments for firm-level returns, and thereby identify only the sensitivity to the systematic component of returns. Under the

[22]We use the Herfindahl-Hirschmann index provided in "Concentration Ratios in Manufacturing," 1992 Census of Manufactures, report MC92-S-2, available on the internet at http://www.census.gov/mcd/mancen/download/mc92cr.sum

[23]Many researchers have emphasized the special character of regulated industries, in particular that the talent requirements may be lower in regulated firms (see, e.g., Peltzman, 1993; and Hubbard and Palia, 1995). In regression not reported here, we ran a similar set of interaction effects for a dummy variable indicating whether the firm was in a regulated industry or not. We define the dummy as equal to one if $4900 \leq SIC \leq 4914$, and to zero otherwise. Although the effect on mean compensation was large and negative and statistically significant (-0.96 with a standard error of 0.10), the interaction effect was statistically insignificant.Our approach suggests the need to look further than political pressures as constraints on CEO compensation in regulated industries (see Joskow, Rose, and Wolfram, 1996). These CEOs presumably have different skills than the CEOs of large, unregulated industrial firms. Indeed, the greater CEO turnover and increased sensitivity to market returns for banking CEOs following deregulation in banking (Hubbard and Palia, 1995) and airlines (Kole and Lehn, 1996) is consistent with the labor market approach we take here.

[24]This possibility was also suggested by Barro and Barro (1990), who documented the failure of RPE for a sample of banks and, in their conclusion, conjecture this might reflect shocks to the marginal product of CEOs.

standard agency view, the coefficient on raw return instrumented with industry returns should be zero. As our discussion makes clear, the effect of aggregate shocks on compensation need not be zero, even when RPE is a feature of executive compensation. Hence finding sensitivity of executive compensation to aggregate variables is neither evidence against RPE nor persuasive evidence of skimming.

3.4 Is Pay-for-Performance Sensitivity Too Low?

The estimated strength of the link between compensation and firm performance is sometimes argued to be smaller in practice than would predicted by agency models. In their seminal paper, Jensen and Murphy (1990a) estimated an elasticity of CEO salary and bonus with respect to shareholder wealth of about 0.10, and concluded that this elasticity is "too small." More recently, Hall and Liebman (1998) estimated elasticities that were on the order of 30 times larger, in large part because they incorporate the entire portfolio of long-term compensation (the sum of salary, bonus, new awards of stocks and options, plus the change in the market value of stocks and options previously granted). Even so, Hall and Liebman (1998) conclude that CEOs' incentives are not as well aligned with shareholder interests.

In the absence of strong a priori beliefs about the model and parameter values, however, it is difficult to evaluate rigorously the magnitudes of estimated pay elasticities. For example, as Hall and Liebman (1998) have pointed out, the answer is sensitive to assumptions about functional form. For large corporations, even small sensitivities can imply large absolute wealth changes for the CEO, so the small coefficients in the linear models estimated by Jensen and Murphy (1990a) do not necessarily imply that incentives are weak. The most important difficulty in assessing the magnitude of performance sensitivities (or elasticities) is absence of a theoretical benchmark against which empirical magnitudes can be measured.

In our model, the semi-elasticity of CEO compensation with respect to returns is given by

$$\frac{\partial \ln w}{\partial r} = \frac{\delta}{\phi}.$$

We estimate semi-elasticities with respect to relative performance in the neighborhood of 0.5. It is difficult to convert this to a standard elasticity for comparison with Jensen and Murphy (1990a) because the mean of relative performance is around zero. Even without this comparison, we find it difficult to identify a standard by which we could claim that pay-for performance elasticities are "too small." To do so requires strong *a priori* knowledge about the relative magnitudes of the elasticity of value with respect to effort, ϕ, and the parameter governing the disutility of effort, δ. This is pointed out by Garen (1994), who also uses a structural model. Pursuing the structural approach to this question is a useful direction for future research.

4 Conclusions

Significant increases in CEO compensation in recent years have coincided with an unprecedented bull market for stocks. They have also coincided with increasingly frequent arguments from academics and practitioners that CEO pay is excessive. Academics in particular point to departures from relative performance evaluation as a major puzzle. Observing that the sensitivity of CEO compensation to market returns is inconsistent with the predictions of principal-agent theory, they have increasingly been tempted to resort to political explanations to explain the compensation of CEOs (e.g., Murphy, 1999). Such bold departures from standard economic theory may be premature, however.

The sensitivity of CEO compensation to market returns is not so puzzling when one recognizes the special feature of the labor market for CEOs. When added to a standard principal-agent model, the combination talent sorting of CEOs among firms and aggregate shocks that affect the marginal value of CEO services to the firm predicts a positive relationship between CEO pay and aggregate stock market returns. We argue that CEOs have scarce human capital, and that supply of highly skilled CEOs capable of running large, complex corporations is relatively inelastic. Under these assumptions, aggregate shocks to demand or productivity will simultaneously raise firm value and raise the marginal value of CEO services to the firm. In equilibrium, such shocks bid up the value of their compensation packages. This phenomenon makes it appear as if (some) firms are violating relative performance evaluation, a basic prediction of agency theory.

Using Standard and Poor's ExecuComp data for a large panel of U.S. firms, we show that the empirical "failure" of RPE is systematically restricted to the CEOs of certain types certain firms. In our data, there is less of a puzzle for the firms in the bottom quartile of the size distribution. The behavior of these firms is in closer accordance with the predictions of simple agency theory. We interpret this as evidence that the supply of CEOs sufficiently skilled to manage such firms is relatively elastic. The compensation packages for the CEOs of larger, more complex firms, however, require highly skilled CEOs. Empirically, the compensation packages for the CEOs of these firms are more sensitive to aggregate shocks, which we interpret as evidence that the supply of highly skilled CEOs is relatively inelastic. These results are robust to robustness checks against alternative hypothesis such as "skimming" and strategic considerations.

Four extensions are suggested by our findings. First, we plan to pursue more formally rent sharing between chief executive officers and firms. In general, the allocation of match-specific rents depends on the characteristics of competition among firms and among CEOs. One polar case is the implicit assumption in the principal-agent approach that firms (shareholders) receive all of the rents from superior management talent; that is, the compensation contract just meets the reservation utility of the CEO. This assumption, which assigns 100 percent of the pure rents to the firm, nevertheless leads to the surprising result that aggregate shocks can increase the CEO's compensation because they represent observable shocks to the marginal value of effort. That is, high-productivity states of the world call for more effort, which the optimal contract both elicits and rewards. Another polar case is the assumption that owners of capital and workers bid for the privilege of being managed by superior talent, in which case 100 percent of the rents go to the CEO. Rosen (1982) points out that this would help to explain of the skewed distribution of CEO compensation. This assumption may also explain the sensitivity of compensation to productivity shocks in general and the sensitivity of compensation to aggregate shocks in particular (because such shocks increase not only the match rents, but also the CEO's outside opportunities). An intermediate approach, in which the optimal contract divides rents between the CEO and the firm is analogous to the model of the endogenous evolution of boards of directors in Hermalin and Weisbach (1998), who take rents as given and address their allocation.

Second, the approach taken here could be extended to settings in which CEOs make firm-specific investments once a match is made. In such a setting, CEOs confront the hold-up problem, and the optimal contract will feature commitments by the firm to increase compensation with observable increases in the CEO's outside opportunities. One possible way to make this commitment would be to grant stock and stock options that are not indexed to the market. Third, empirical analysis of divisional managers' compensation in large firms may draw upon the labor market considerations outlined here for CEOs. Fourth, our model of CEO compensation predicts increases in earnings inequality within firms (and within nations) during periods of high market returns. This extension is related to recent work by Acemoglu and Newman (1997), in which the level of wages and the use of high-powered incentives are positively related.

A The Derivation of the Optimal Contract

Following Grossman and Hart (1983), one can express the principal's problem as a two-stage problem. First the principal chooses the optimal contract for any effort level $\hat{x}$. Second, given the optimal contract, the optimal effort level is chosen. The problem can be written

$$\max_{\hat{x}} \int \left[V - w(V,\hat{x})\right] f(V|\hat{x})dV,$$

where $w(V,\hat{x})$ is the cost of the optimal contract that implements $\hat{x}$. Assuming the first-order approach is valid (for sufficient conditions, see Jewitt, 1988, for example), this contract is defined by the solution to the sub-problem

$$w(V,\hat{x}) = \arg\min_{w(V)} \int w(V)f(V|\hat{x})dV$$

subject to

$$\int u(w(V),x)f(V|x)dV \geq \bar{u} \qquad \text{(Participation Constraint)} \tag{A1}$$

$$\frac{\partial}{\partial x}\int u(w(V),x)f(V|x)dV = 0 \qquad \text{(Incentive Constraint)} \tag{A2}$$

Associating lagrange multipliers λ and μ with the constraints (11) and (11), the first-order condition for the first problem is

$$\frac{1}{u_w(w(V),x)} = \lambda + \mu\left(\frac{f_x(V|x)}{f(V|x)} + \frac{u_{wx}(w(V),x)}{u_w(w(V),x)}\right). \tag{A3}$$

(Note: When utility is additively separable, observe that $u_{wx} = 0$, in which case Equation (A3) simplifies to the more familiar $\frac{1}{u_w} = \lambda + \mu\frac{f_x}{f}$.)

A.1 Model Assumptions

A.1.1 Utility

We assume the utility of the agent is described by the power utility function

$$u(w,x) = \frac{1}{1-\gamma}\left(wx^{-\delta}\right)^{1-\gamma} = \frac{1}{1-\gamma}w^{1-\gamma}x^{-\delta(1-\gamma)},$$

To avoid the technical complications associated with the case $\gamma = 1$, the following derivations assume $\gamma > 1$. To impose the assumption that utility is decreasing and concave in effort, we assume $-\delta(1-\gamma) > 1$. A plausible set of parameters, for example, would be $\gamma = 3$ and $\delta = 1$.

Note that $u_w = w^{-\gamma}x^{-\delta(1-\gamma)}$, which implies $1/u_w(w,x) = w^{\gamma}x^{\delta(1-\gamma)}$.Plugging this utility function into the first-order condition for the optimal contract in equation (A3) implies the optimal contract is of the form

$$w(V,\hat{x}) = \hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}} \left(\lambda + \mu h(V,\hat{x})\right)^{\frac{1}{\gamma}}, \tag{A4}$$

where $h(V,\hat{x}) \equiv \frac{u_{wx}(w(V),x)}{u_w(w(V),x)} + \frac{f_x(V,\hat{x})}{f(V,\hat{x})}$, and the $\hat{x}$, λ and μ are contract parameters to be chosen by the principal. Substituting equation (**??**) back into the utility function implies the agent's utility of the optimal contract is

$$u(w(V,\hat{x}),x) = \frac{1}{1-\gamma}\left(\lambda + \mu h(V,\hat{x})\right)^{\frac{1-\gamma}{\gamma}} \hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}} x^{-\delta(1-\gamma)}. \tag{A5}$$

We can substitute this equation into equations (A6) and (A7) to write the two constraint equations on the feasible contract as

$$\int \frac{1}{1-\gamma}\left(\lambda + \mu h(V,\hat{x})\right)^{\frac{1-\gamma}{\gamma}} \hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}} x^{-\delta(1-\gamma)} f(V|\hat{x})dV \;=\; \bar{u}, \tag{A6}$$

$$\frac{\partial}{\partial x}\left[\int \left(\lambda + \mu h(V,\hat{x})\right)^{\frac{1-\gamma}{\gamma}} \hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}} x^{-\delta(1-\gamma)} f(V|x)dV\right]_{x=\hat{x}} \;=\; 0, \tag{A7}$$

where $f(V|x)$ is defined above in equation (A8). Note the distinction between x and $\hat{x}$ in equation (11); $\hat{x}$ is the principal's desired effort level, whereas x is the agent's choice variable.

A.1.2 Production

We assume the production relationship between managerial "effort" and firm value is described by the Cobb-Douglas function

$$V = \theta(\tau x)^{\phi} K,$$

where $\ln\theta$ follows a random walk with drift parameter βm, so that $\Delta\ln\theta = \beta m + \varepsilon$,where $\varepsilon \sim N(0,\sigma^2)$. It follows that

$$\ln V = \ln V_{-1} + \phi\Delta\ln\tau + \phi\Delta\ln x + \Delta\ln K + \beta m + \varepsilon.$$

The distribution of V conditional on x is therefore lognormal, and the density of V (conditional on θ_{-1}, β, m, K and x) is

$$f(V,x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln V - a - \phi\ln x}{\sigma}\right)^2\right], \tag{A8}$$

where

$$a \equiv \theta_{-1} + \beta m + \ln K + \ln\tau.$$

Recall that $u_w = w^{-\gamma}x^{-\delta(1-\gamma)}$, so $u_{wx} = \delta(1-\gamma)w^{-\gamma}x^{-\delta(1-\gamma)-1}$. It therefore follows that $\frac{u_{wx}}{u_w} = \frac{\delta(1-\gamma)}{x}$, so $h \equiv \frac{f_x}{f} + \frac{u_{wx}}{u_w}$ is

$$h(V,x) = \frac{\phi}{\sigma x}\left(\frac{\ln V - a - \phi \ln x}{\sigma}\right) - \frac{\delta(1-\gamma)}{x}. \tag{A9}$$

Empirically, the ex-post performance measure is

$$\Delta \ln V - \phi \Delta \ln \tau - \phi \Delta \ln x - \Delta \ln K = \beta m + \varepsilon.$$

A.2 The Optimal Contract

Recall from equation (A4) that the optimal contract w solves

$$w(V,\hat{x}) = \hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}}\left[\lambda + \mu h(V,\hat{x})\right]^{\frac{1}{\gamma}}$$

Substituting from $h(V,\hat{x}) \equiv \frac{\phi}{\sigma \hat{x}}\left(\frac{\ln V - a - \phi \ln \hat{x}}{\sigma}\right) + \hat{h}$ yields

$$w = \hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}}\left[\left(\lambda + \mu\hat{h}\right) + \mu\frac{\phi}{\sigma\hat{x}}\left(\frac{\ln V - a - \phi \ln \hat{x}}{\sigma}\right)\right]^{\frac{1}{\gamma}},$$

which can be written more elegantly as

$$w = A\left[1 + \tfrac{B}{\sigma}r\right]^{\frac{1}{\gamma}},$$

where A and B are contract parameters, and r is the measure of relative performance, given by

$$\begin{aligned} A &= \left(\lambda + \mu\hat{h}\right)^{\frac{1}{\gamma}}\hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}}, \\ B &= \left(\tfrac{\mu}{\lambda + \mu\hat{h}}\right)\left(\tfrac{\phi}{\sigma\hat{x}}\right), \\ r &= \ln V - a - \phi \ln \hat{x}. \end{aligned}$$

This description of the optimal contract suggests esimating the parameters of the contract using the regression specification

$$\ln w = \ln A + \frac{1}{\gamma}\ln\left[1 + \tfrac{B}{\sigma}r\right],$$

which is approximated by

$$\ln w = \ln A + \frac{B}{\gamma\sigma}\ln\left[1 + r\right].$$

A.3 Approximation Arguments

Solving for the parameters A and B of the optimal contract is difficult because equations (A6) and (A7) do not have clean analytic solutions. We use second-order Taylor expansions of the integrands of these equations to obtain tractable analytic approximations. We then solve for the optimal contract which is the exact solution to the local approximation of the problem.

Under the assumption of lognormality for V, inspection of equation (A9) reveals that the function $h = h(V, \hat{x})$ defines a normally distributed random variable indexed by the parameter $\hat{x}$. (Note that $\hat{x}$ is not x, which indexes the distribution of V). Evaluation of equations (A6) and (A7), as well as the principal's objective, requires the calculation of expectations of the form $E\left[(\lambda + \mu h)^c \mid x\right]$, where $c \in \Re$. There are values of c (e.g., $c = 1/2$) for which $\lambda + \mu h$ would be negative, in which case real solutions to $(\lambda + \mu h)^c$ would not exist. We address this problem by implicitly assuming the contract satisfies the limited liability constraint $\lambda + \mu h \geq 0$ for all V. At the same time, we assume the constraint that binds with such low probability that it can be approximated by the unconstrained problem. (This will be the case when the variance of μh is small relative to λ.) Thus, we assume $E\left[(\lambda + \mu h)^c \mid \lambda + \mu h \geq 0,\, x\right] \simeq E\left[(\lambda + \mu h)^c \mid x\right]$.

To approximate expressions of the form $E\left[(\lambda + \mu h)^c |x\right]$, we use second-order Taylor series expansions of $(\lambda + \mu h)^c$ around $h = E\left[h|x\right]$:

$$(\lambda + \mu h)^c \simeq (\lambda + \mu E\left[h|x\right])^c \Big[1 + c\left(\tfrac{\mu}{\lambda + \mu E[h|x]}\right)(h - E\left[h|x\right]) + \tfrac{c(c-1)}{2}\left(\tfrac{\mu}{\lambda + \mu E[h|x]}\right)^2 (h - E\left[h|x\right])^2\Big].$$

We can therefore approximate $E\left[(\lambda + \mu h)^c |x\right]$ as a linear function of the conditional mean and variance of h:

$$E\left[(\lambda + \mu h)^c |x\right] \simeq (\lambda + \mu E\left[h|x\right])^c \left[1 + \tfrac{c(c-1)}{2}\left(\tfrac{\mu}{\lambda + \mu E[h|x]}\right)^2 V\left[h|x\right]\right]. \tag{A10}$$

Note that $h = h(V, \hat{x})$ is normally distributed with conditional mean and variance (with respect to the distribution $f(V, x)$) given by

$$E\left[h|x\right] = \frac{\phi}{\sigma\hat{x}}\left(\frac{\phi \ln x - \phi \ln \hat{x}}{\sigma}\right) - \frac{\delta(1-\gamma)}{\hat{x}}, \tag{A11}$$

$$V\left[h|x\right] = \left(\frac{\phi}{\sigma\hat{x}}\right)^2. \tag{A12}$$

A.3.1 The Incentive Constraint

Using Taylor approximations, the expected utility of the contract to the agent appearing in equation (11) is approximated by

$$\begin{aligned}
&E\left[u(w(V,\hat{x}),x)\,|\,x\right] \\
&= \tfrac{1}{1-\gamma}E\left[(\lambda+\mu h(V))^{\frac{1-\gamma}{\gamma}}\;|\,x\right]\hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}}x^{-\delta(1-\gamma)} \\
&\simeq \tfrac{1}{1-\gamma}\left(\lambda+\mu E\left[h|x\right]\right)^{\frac{1-\gamma}{\gamma}}\left[1+\kappa_1\left(\tfrac{\mu}{\lambda+\mu E[h|x]}\right)^2\left(\tfrac{\phi}{\sigma\hat{x}}\right)^2\right]\hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}}x^{-\delta(1-\gamma)}
\end{aligned} \tag{A13}$$

where $\kappa_1=\frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)\left(\frac{1-2\gamma}{\gamma}\right)$. The incentive compatibility constraint is evaluated by setting the derivative of this expression with respect to x equal to zero, and evaluating the result at $x=\hat{x}$. Ignoring the multiplicatively constant term $\frac{1}{1-\gamma}\hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}}$, the derivative of this expression is

$$\begin{aligned}
\frac{\partial E\left[u(w(V,\hat{x}),x)\,|\,x\right]}{\partial x} \;=\;& \left(\tfrac{1-\gamma}{\gamma}\right)\left(\lambda+\mu\hat{h}\right)^{\frac{1-2\gamma}{\gamma}}\mu\hat{h}_x\left(1+\kappa_1\left(\tfrac{\mu}{\lambda+\mu\hat{h}}\right)^2\left(\tfrac{\phi}{\sigma\hat{x}}\right)^2\right)x^{-\delta(1-\gamma)} \\
&-\left(\lambda+\mu\hat{h}\right)^{\frac{1-\gamma}{\gamma}}\left(2\kappa_1\left(\tfrac{\mu}{\lambda+\mu\hat{h}}\right)\tfrac{\mu^2}{\left(\lambda+\mu\hat{h}\right)^2}\hat{h}_x\left(\tfrac{\phi}{\sigma\hat{x}}\right)^2\right)x^{-\delta(1-\gamma)} \\
&+\left(\lambda+\mu\hat{h}\right)^{\frac{1-\gamma}{\gamma}}\left(1+\kappa_1\left(\tfrac{\mu}{\lambda+\mu\hat{h}}\right)^2\left(\tfrac{\phi}{\sigma\hat{x}}\right)^2\right)\left(-\delta(1-\gamma)\right)x^{-\delta(1-\gamma)-1}.
\end{aligned}$$

Now simplify by dividing through by $x^{-\delta(1-\gamma)-1}$ and $\left(\lambda+\mu\hat{h}\right)^{\frac{1-\gamma}{\gamma}}$, and substitute B for $\left(\frac{\mu}{\lambda+\mu\hat{h}}\right)\left(\frac{\phi}{\sigma\hat{x}}\right)$ to get

$$\left(\tfrac{1-\gamma}{\gamma}\right)\left(\tfrac{\mu}{\lambda+\mu\hat{h}}\right)\hat{h}_x\left(1+\kappa_1B^2\right)x=\left(2\kappa_1\left(\tfrac{\mu}{\lambda+\mu\hat{h}}\right)\hat{h}_xB^2\right)x+\left(1+\kappa_1B^2\right)\delta(1-\gamma).$$

Finally, use the fact that $xh_x=\frac{\phi}{\sigma}\left(\frac{\phi}{\sigma\hat{x}}\right)$ to see that $\left(\frac{\mu}{\lambda+\mu\hat{h}}\right)x\hat{h}_x=B\frac{\phi}{\sigma}$, so the first-order condition simplifies to

$$\left(\tfrac{1-\gamma}{\gamma}\right)B\tfrac{\phi}{\sigma}\left(1+\kappa_1B^2\right)=2\kappa_1\tfrac{\phi}{\sigma}B^3+\left(1+\kappa_1B^2\right)\delta(1-\gamma).$$

Or more elegantly,

$$B\tfrac{\phi}{\sigma\gamma}\left(1-\kappa_1B^2\right)=\left(1+\kappa_1B^2\right)\delta.$$

Notice that if γ is near 0.5 or 1.0, or if B^2 is "small," then κ_1B^2 is approximately zero, and we get an elegant solution for the "slope" coefficient B, namely

$$B=\frac{\sigma\delta\gamma}{\phi}.$$

Recall that the coefficient on $\ln(1+r)$ for the optimal contract is $B/\gamma\sigma$. Therefore, by substituting from the above first-order condition, we have shown that the optimal contract is approximately

$$\ln w = \ln A + \frac{\delta}{\phi} \ln (1 + r) \, .$$

A.3.2 The Participation Constraint

By evaluating the expected utility at $x = \hat{x}$, letting $\hat{h} = E[h|x]$, and setting the result equal to $\bar{u}$, the approximation to the participation constraint in equation (11) is

$$\frac{1}{1-\gamma} \left(\lambda + \mu \hat{h}\right)^{\frac{1-\gamma}{\gamma}} \left[1 + \kappa_1 \left(\tfrac{\mu}{\lambda + \mu \hat{h}}\right)^2 \left(\tfrac{\phi}{\sigma \hat{x}}\right)^2\right] \hat{x}^{\frac{-\delta(1-\gamma)^2}{\gamma}} x^{-\delta(1-\gamma)} = \bar{u}.$$

¿From the discussion of the optimal contract, recall that the parameters A and B are defined as $A = \left(\lambda + \mu \hat{h}\right)^{\frac{1}{\gamma}} \hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}}$ and $B = \left(\frac{\mu}{\lambda + \mu \hat{h}}\right)\left(\frac{\phi}{\sigma \hat{x}}\right)$. Recalling that we have assumed $\gamma > 1$, and noting that $\bar{u}(1-\gamma) > 0$, we can substitute into the above equation and rearrange to solve for A:

$$A = \left(\frac{\bar{u}\,(1-\gamma)}{1 + \kappa_1 B^2}\right)^{\frac{1}{1-\gamma}} \hat{x}^{\delta}.$$

This expression gives the value of the parameter A for the contract the minimizes the cost of implementing the action $\hat{x}$. Since κ_1 and B are constants that, as we argued above, depend only on the exogenous parameters γ, δ, ϕ and σ, this equation says that A varies with $\bar{u}$ and $\hat{x}$. To say more than this, we need to characterize the firm's optimal choice of $\hat{x}$.

A.3.3 The Principal's Objective Function

The expected cost to the principal of inducing effort level $\hat{x}$ is

$$E\left[w(V, \hat{x}) \,|\, \hat{x}\right] = E\left[\left(\lambda + \mu h(V, \hat{x})\right)^{\frac{1}{\gamma}} \hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}} \,|\, \hat{x}\right].$$

Using the Taylor series approximation derived in equation (**??**), and evaluating at $x = \hat{x}$, we get

$$E\left[w(V, \hat{x}) \,|\, \hat{x}\right] \simeq \left(\lambda + \mu \hat{h}\right)^{\frac{1}{\gamma}} \left[1 + \left(\tfrac{1-\gamma}{2\gamma^2}\right)\left(\tfrac{\mu}{\lambda + \mu \hat{h}}\right)^2 \left(\tfrac{\phi}{\sigma \hat{x}}\right)^2\right] \hat{x}^{\frac{-\delta(1-\gamma)}{\gamma}}$$

Substituting from the IC and PC says that the cost of implementing action $\hat{x}$ is given by

$$E\left[w(V, \hat{x}) \,|\, \hat{x}\right] \simeq \left[1 + \left(\tfrac{1-\gamma}{2\gamma^2}\right) B^2\right] A.$$

Substituting for A – which still depends on the level of x to be implemented, $\hat{x}$ – yields

$$E\left[w(V,\hat{x})\,|\,\hat{x}\right] \simeq C\hat{x}^{\delta},$$

where

$$C = \left[1 + \left(\tfrac{1-\gamma}{2\gamma^2}\right)B^2\right]\left(\frac{\bar{u}\,(1-\gamma)}{1+\kappa_1 B^2}\right)^{\frac{1}{1-\gamma}}.$$

Therefore the firm's optimization problem is

$$\max_{\hat{x}}\left(\Theta x^{\phi} - C\hat{x}^{\delta}\right).$$

where

$$\Theta = \tau^{\phi}K\exp\left(\theta_{-1} + \beta' m + \frac{1}{2}\sigma^2\right).$$

The first-order condition is

$$\Theta\phi x^{\phi-1} - C\delta x^{\delta-1} = 0,$$

or

$$x^{\phi-1}\left(\Theta\phi - C\delta x^{\delta-\phi}\right) = 0.$$

which with some rearranging of terms implies principal's desired effort choice (or "demand for effort") is

$$\hat{x} = \left(\frac{\phi\Theta}{C\delta}\right)^{\frac{1}{\delta-\phi}}. \tag{A14}$$

A.4 The Optimal Contract, Revisited

With equation (A14), we can characterize the optimal contract as

$$\ln w = \ln A + \frac{\delta}{\phi}\ln\left(1+r\right),$$

where we have now shown that the parameter A is determined by the contracting environment as follows:

$$
\begin{aligned}
A &= \left(\frac{\bar{u}\,(1-\gamma)}{1+\kappa_1 B^2}\right)^{\frac{1}{1-\gamma}} \left(\frac{\phi\Theta}{C\delta}\right)^{\frac{\delta}{\delta-\phi}}, \\
\Theta &= \tau^{\phi} K \exp\left(\theta_{-1} + \beta' m + \frac{1}{2}\sigma^2\right), \\
C &= \left[1 + \left(\tfrac{1-\gamma}{2\gamma^2}\right) B^2\right] \left(\frac{\bar{u}\,(1-\gamma)}{1+\kappa_1 B^2}\right)^{\frac{1}{1-\gamma}}.
\end{aligned}
$$

B Data Construction

B.1 CEO Compensation Measures

To construct our measures of CEO wealth tied to the firm, we use the October 1999 release of Standard & Poor's ExecuComp data file. This file contains executive compensation data as reported by firms on their proxy statements for the years 1992 to 1998. Firms are required to report salary, bonus, the value of stocks and options granted, and other miscellaneous compensation items during the year. They are also required to report the existing number of shares held by executives in addition to those granted in the current year. Unfortunately, SEC reporting requirements require firms to report stock options held only if the options are currently in the money. Moreover, those options are valued at the money, and without the full time series of options granted and exercised, it is not possible to obtain a precise measure of the stock of options held by the CEO.[25] The stock market rose dramatically over our sample period, so it is likely that a large percentage of the options held by CEOs during this period were in the money.

With these caveats in mind, we construct the variable $cwealth_t$, which is the wealth of the CEO held in the firm, as follows (In this appendix, ExecuComp's acronyms are indicated in capital letters, and our data constructions are indicated by lower case):

$$cwealth_t = SALARY + BONUS + OTHANN + LTIP + ALLOTHTO + eqv_t + opv_t,$$

where $OTHANN$, $LTIP$, and $ALLOTHTO$ denote other annual compensation, the total value of stock options granted (using Black-Scholes), and long-term incentive payouts, respectively ($SALARY$ and $BONUS$ are self-explanatory). The last two terms, eqv_t and opv_t, denote the value of stocks and stock options, respectively. These value of equity is easily constructed from ExecuComp data as

$$eqv_t = SHROWN_t \times PRCCF_t \times AJEX_t,$$

where $SHROWN_t$ and $PRCCF_t \times AJEX_t$ are the number of shares owned by the CEO and price per share, respectively ($AJEX_t$ is the adjustment factor for stock splits, etc.)

To estimate opv_t, we can use the ExecuComp variable BLK_VALUE_t to measure the Black-Scholes value of *current* grants. It is more difficult to construct the Black-Scholes value of options outstanding, because the necessary detailed on the current holdings of previous grants is not reported on the proxy. Firms do, however, report the number of shares, and this number is broken down into shares that are exercisable vs. those that are not. We assign values to each group of shares. Executive stock options are usually granted with a four-year vesting period during which usually 25 percent of the options vest in each year. Hence knowing whether an option is exercisable or not is useful information because it provides a crude indicator of the year in which the options were granted.

To estimate the value of the stock of previously held options, we assume that none of the options granted in the current year were exercisable in the current year. We also assume

[25] Hall and Liebman (1998) confront this problem by collecting data on option grants for the ten-year period prior to their sample.

that 25 percent of the options that were unexercisable last period became exercisable this period. Hence we assume that $UEXNUMEX_t$ and $UEXNUMUN_t$ evolve according to

$$\begin{aligned} UEXNUMUN_t &= (1-0.25)\, UEXNUMUN_{t-1} + SOPTGRNT_t, \\ UEXNUMEX_t &= 0.25 \times UEXNUMUN_{t-1} + UEXNUMEX_{t-1} - SOPTEXSH_t, \end{aligned}$$

where $SOPTEXSH_t$ and $SOPTGRNT_t$, respectively, are the ExecuComp variables indicating the number of options exercised and granted during the current year. We use this information to solve for the beginning-of-period values:

$$\begin{aligned} UEXNUMUN_{t-1} &= (UEXNUMUN_t - SOPTGRNT_t)/(1-0.25), \\ UEXNUMEX_{t-1} &= UEXNUMEX_t - 0.25\, UEXNUMUN_t + SOPTEXSH_t. \end{aligned}$$

We then use estimate opv_t as

$$opv_t = BLK\_VALUE_t + (UEXNUMEX_{t-1} \times bsv\_ex_t) + (UEXNUMUN_t \times bsv\_un_t),$$

where bsv_ex_t and bsv_un_t denote our estimates of the option value of the two parts. We apply the Black Scholes formula making the following assumptions (which closely follow the assumptions used by ExecuComp). We do not include options exercised during the year as wealth in the firm because they are either measured by eqv_t (if the CEO kept the stock), or they are held as part of the CEO's private wealth.

1. We do not observe strike prices for the stock of previously granted options, so we estimate strike prices assuming that all options are granted at the money (which is approximately true) and then using estimates of the date on which the current stocks were granted. For exercisable options at the beginning of the year, $UEXNUMEX_{t-1}$, we assume an average strike price of $0.1p_{t-1} + 0.3p_{t-2} + 0.6p_{t-3}$, where p_t is the stock price at the end of the fiscal year. For $UEXNUMUN_{t-1}$,we assume an average strike price of $0.6p_{t-1} + 0.3p_{t-2} + 0.1p_{t-3}$. The price weighting reflects the fact that exercisable options tend to be older grants, and we truncate at three lags to reduce the impact of missing values on prices.

2. ExecuComp assumes the term of the grant is 7 years, because executives rarely wait until the expiration date to exercise their options. Following this assumption, we assume an average of term of 3 years for $UEXNUMEX_{t-1}$ and 5 years for $UEXNUMUN_{t-1}$.

3. The risk-free rates are the values used by ExecuComp, which are the approximate average yields that could have been earned by investing in a U.S. Treasury bond carrying a seven-year term for the years 1992 to 1998: 6.43%, 5.53%, 7.84%, 5.49%, 6.34%, 5.77% and 4.73%.

4. The expected volatility is the variance of the monthly total return for all available observations in the preceding 5-year period. This number is reported by ExecuComp as BS_VOLAT.

5. The expected dividend yield is also the same as that used by ExecuComp, *BS_YIELD*.

B.2 Firm and CEO Characteristics

The Execucomp variable BECAMECEO identifies the date on which a particular executive became CEO. We use this variable both to identify CEOs and to measure their tenure as CEO.[26] Similarly, we used JOINED_CO to determine how long the CEO had been with the company when appointed to CEO (this is our measure of "outsidership" discussed in the paper). We used the variable PEXECDIR to identify whether the CEO served on the board of directors.

We merge Execucomp data with Compustat to get additional firm-level variables; the variables used (and their Compustat identification codes) are listed in Table 1.

B.2.1 Diversification Measure

To construct a measure of the diversity of firms' across different industries, we use the Compustat Industry Segment files. The Segment files provide the total sales generated by each business segment for each firm-year (sdata1). The share of each segment's sales is taken as the ratio of the business segment sales to total firm sales as reported in the Compustat annual file (data12). We calculate a Herfindahl-like index of diversity by calculating one minus the squared shares all the business industries for each firm-year.

B.2.2 Construction of Industry-Weighted Herfindahl Indices

Using a dataset which provides measures of manufacturing industries' Herfindahl index (HI) in 19xx by SIC code, we calculate the sales weighted average of HI for each firm. Specifically, if a firm has only one business segment (or is not listed in the Compustat Industry Segment files) we assign the HI of its sole industry (using the variable SIC from the Annual Compustat files) to that firm-year. If a firm has more than one industry segment in a particular year, then we weight the various HI's according to the ratio of the business segment sales to total firm sales as reported in the Compustat annual file (data12).

B.3 Performance Measures

To construct our relative performance measures (RPM) and market measures (MKT), we use a variety of methods to insure robustness. Our baseline measure is calculate 60-month betas for each firm-year in our sample using data from the monthly CRSP files and the time series of Fama French factors created above. The estimation period for each firm-year ends in the month of its fiscal year end. For example, for a firm with June as its fiscal year end, the 1995 estimation period runs from July of 1990 through June of 1995. Two such time series regressions are run for each firm-year, a CAPM styled regression and a Fama-French three factor type regression. First, firm returns are regressed on a constant and the market

[26]There is also a dummy variable that to idenfiy the CEO in each year, but we found that this variable was often missing during the early years of the Execucomp data. Instead, we identify CEO using the ExecuComp variable BECAMECEO, which is a cross-sectional (executive-firm-specific) variable giving the date on which the executive became the CEO, and which is available for every executive in the data.

return (implemented as CRSP's value-weighted return). Second firm returns are regressed on a constant, the value-weighted market return, and the Fama-French factors SMB and HML calculated above. We also calculate a 12 month mean return for each firm year with similarly constructed estimation windows. The systematic stock return for a firm-year is defined as the 12 month predicted return from either the CAPM or Fama-French factor regressions. The idiosyncratic return is then the observed 12-month mean return less the systematic return.

B.3.1 Industry, Size, and Fama-French Factors

We construct size-weighted portfolios of industry returns based on 2-digit SIC codes. Each firm is assigned an industry portfolio corresponding to its 2-digit industry code. For each firm, we then construct a size-weighted portfolio of industry returns that excludes the that firm. Hence, there is, to some degree, cross-sectional variation in the definition of industry returns. The same procedure is used for each firm to construct size-weighted portfolios based on size deciles.

To construct Fama-French styled factors, we form portfolios based on market capitalizations and book-to-market ratios. Portfolios are created using the universe of non-financial NYSE listed firms with non-missing size and book-to-market quasi-. We use the annual Compustat files to calculate size (S) and book-to-market (BM) as DATA199×DATA25 and DATA216/(DATA12 + DATA10) respectively. BM values greater than 100 are set to 100. We then convert this annual dataset to a monthly frequency based on the fiscal year end month for each firm. For example, a firm with June as its fiscal year end is counted as having its 1995 annual S and BM values from June 1995 until May 1996. Firms are then grouped each month into S and BM classes. To ensure that portfolio formation is based on information available in the current month, we form classes based on S and BM values from the prior month.

Following Fama and French, we use three BM classes and two S classes. A firm is classified as large if its S is greater than the median firm in the sample and small otherwise. Similarly, a firm's book-to-market ratio is classified as low if its BM value is in the bottom 30 percent, medium if the value is between the 30th and 70th percentiles, and high if its BM is greater than that of the 70th-percentile firm. Six classes of firms are formed based on the two size and three book-to-market classifications. Portfolio returns for each of the six classes are calculated using size weighted CRSP monthly returns for each stock in the class. Finally, two 'factor mimicking' portfolio returns are calculated: high minus low book-to-market (HML) and small minus large size (SMB). The HML portfolio is calculated as the equally weighted return of the two high book-to-market portfolios less the equally weighted return of the two low book-to-market portfolios (the two medium book-to-market portfolios are not included). Similarly, we construct the SMB portfolio as the difference between the three small firm portfolios less the three large firm portfolios.

References

Acemoglu, Daron, and Andrew Newman. "The Labor Market and Corporate Structure." Mimeograph, MIT, 1997.

Aggarwal, Rajesh, and Andrew A. Samwick. "Executive Compensation, Strategic Competition, and Relative Evaluation: Theory and Evidence." Journal of Finance 44 (December 1999).

Aggarwal, Rajesh, and Andrew A. Samwick. "The Other Side of the Tradeoff: The Impact of Risk on Executive Compensation." Journal of Political Economy 107 (February 1999): 65-105.

Antle, Rick, and Abbie Smith. "An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives." Journal of Accounting Research 24 (1986): 1-39.

Baker, George P., and Brian J. Hall. "CEO Incentives and Firm Size." Working Paper No. 6868, National Bureau of Economic Research, December 1998.

Baker, George, Michael C. Jensen, and Kevin J. Murphy. "Compensation and Incentives: Practice versus Theory." Journal of Finance 43 (September 1988): 593-616.

Baker, Malcolm and Paul A. Gompers. "An Analysis of Executive Compensation, Ownership, and Control in Closely Held Firms." Working Paper, Harvard Business School, March 1999.

Barro, Jason, and Robert J. Barro. "Pay, Performance, and Turnover of Bank CEOs." Journal of Labor Economics 8 (October 1990): 448-481.

Berle, Adolph A., and Gardiner C. Means. The Modern Corporation and Private Property. New York, 1932.

Bertrand, Marianne, and Sendhil Mullainathan. "Are CEOs Rewarded for Luck? A Test of Performance Filtering." Mimeograph, Princeton University, March 1999.

Blanchflower, David G., Andrew J. Oswald, and Peter Sanfey. "Wages, Profits, and Rent-Sharing." Quarterly Journal of Economics 111 (February 1996): 227-257.

Bushman, Robert, Raffi Indejikian, and Abbie Smith. "CEO Compensation: The Role of Individual Performance Evaluation." Journal of Accounting and Economics 21 (1996): 161-193.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay. The Econometrics of Financial Markets. Princeton: Princeton University Press, 1997.

Ciscel, David H. "Determination of Executive Compensation." Southern Economic Journal 40 (April 1974): 613-617.

Ciscel, David H., and Thomas M. Carroll. “The Determinants of Executive Salaries: An Economic Survey.” Review of Economics and Statistics 62 (February 1980): 7-13.

Crystal, Graef S. In Search of Excess: The Overcompensation of American Executives. New York: W.W. Norton, 1991.

Demsetz, Harold and Kenneth Lehn. “The Structure of Corporate Ownership: Causes and Consequences.” Journal of Political Economy 93 (1985), 1155-77.

Dickens, William T., and Lawrence F. Katz. “Inter-Industry Wage Differences and Industry Characteristics.” In Kevin Lang and Jonathan Leonard, eds., Underemployment and the Structure of Labor Markets. Oxford: Basil Blackwell, 1987.

Fama, Eugene F. “Agency Problems and the Theory of the Firm.” Journal of Political Economy 88 (April 1980): 288-307.

Fama, Eugene F., and Kenneth French. “Common Risk Factors in the Returns on Stock and Bonds,” Journal of Financial Economics 33 (1993): 3-56.

Fershtman, Chaim, and Kenneth L. Judd. “Equilibrium Incentives in Oligopoly.” American Economic Review 77 (1987): 927-940.

Garen, John E. “Executive Compensation and Principal-Agent Theory.” Journal of Political Economy 102 (December 1994): 1175-1199.

Gibbons, Robert S. “Incentives and Careers in Organizations.” In David Kreps and Kenneth Wallis, eds., Advances in Economic Theory: Seventh World Congress, volume 2. Cambridge: Cambridge University Press, 1997.

Gibbons, Robert S., and Kevin J. Murphy. “Relative Performance Evaluation for Chief Executive Officers.” Industrial and Labor Relations Review 43 (February 1990): 30s-51s.

Grossman, Sanford, and Oliver Hart. “An Analysis of the Principal-Agent Problem.” Econometrica 51 (January 1983): 7-45.

Hall, Brian J., and Jeffrey B. Liebman. “Are CEOs Really Paid Like Bureaucrats?” Quarterly Journal of Economics 112 (August 1998): 653-697.

Haubrich, Joseph G. “Risk-Aversion, Performance Pay, and the Principal-Agent Problem.” Journal of Political Economy 102 (April 1994): 258-276.

Hermalin, Benjamin and Michail Weisback. “Endogenously Chosen Boards of Directors and their Monitoring of the CEO.” American Economic Review 88 (1998), 96-118.

Himmelberg, Charles P., R. Glenn Hubbard, and Darius Palia. “Understanding the Determinants of Managerial Ownership and the Link Between Ownership and Performance.” Journal of Financial Economics 53 (1999): 353-384.

Holmstrom, Bengt. “Moral Hazard and Observability.” Bell Journal of Economics 10 (Spring 1979): 74-91.

Holmstrom, Bengt. "Moral Hazard in Teams." Bell Journal of Economics 13 (Fall 1982): 324-340.

Holmstrom, Bengt, and Paul Milgrom. "Aggregation and Linearity in the Provision of Intertemporal Incentives." Econometrica 55 (March 1987): 303-328.

Hubbard, R. Glenn, and Darius Palia. "Executive Pay and Performance: Evidence from the U.S. Banking Industry." Journal of Financial Economics 39 (1995): 105-130.

Janakiraman, Surya, Richard Lambert, and David Larcker. "An Empirical Investigation of the Relative Performance Evaluation Hypothesis." Journal of Accounting Research, 1992.

Jensen, Michael C., and William M. Meckling. "Theory of the Firm: Managerial Behavior, Agency Cost, and Ownership Structure." Journal of Financial Economics 1 (October 1976): 305-360.

Jensen, Michael C., and Kevin J. Murphy. "Performance Pay and Top-Management Incentives." Journal of Political Economy 98 (April 1990a): 225-264.)

Jensen, Michael C., and Kevin J. Murphy. "It's Not How Much You Pay, But How." Harvard Business School." Harvard Business Review 68 (May-June 1990b): 138-153.

Joskow, Paul L., Nancy L. Rose, and Andrea Shepard. "Regulatory Constraints on CEO Compensation." Brookings Papers on Economic Activity, Microeconomics (1993): 1-58.

Joskow, Paul L., Nancy L. Rose, and Catherine Wolfram. "Political Constraints on Executive Compensation: Evidence from the Electric Utility Industry." RAND Journal of Economics 27 (Spring 1996): 165-182.

Kaplan, David S. "Principal-Agent Models of CEO Pay-for-Performance Relationships." Working Paper 318, Office of Employment Research and Program Development, December 1998.

Katz, Lawrence F., and Lawrence H. Summers. "Industry Rents: Evidence and Implications." Brookings Papers on Economics, Microeconomics (1989): 209-275.

Kole, Stacey, and Kenneth Lehn. "Deregulation and the Adaptation of Governance Structure: The Case of the U.S. Airline Industry." Mimeograph, University of Rochester, 1996.

McConnell, John, and Henri Servaes. "Additional Evidence on Equity Ownership and Corporate Value." Journal of Financial Economics 27 (October 1990): 595-612.

McLaughlin, Kenneth J. "Rent Sharing in an Equilibrium Model of Matching and Turnover." Journal of Labor Economics 12 (1994): 499-523.

Mirrlees, James. "Notes on Welfare Economics, Information, and Uncertainty." In M. Balch, D. McFadden, and S. Wu, eds., Essays on Economic Behavior Under Uncertainty. Amsterdam: North Holland, 1974.

Mirrlees, James. "The Optimal Structure of Incentives and Authority Within an Organization." Bell Journal of Economics 7 (1976): 105-131.

Mørck, Randall, Andrei Shleifer, and Robert W. Vishny. "Management Ownership and Market Valuation: An Empirical Analysis." Journal of Financial Economics 20 (1988): 293-316.

Murphy, Kevin J. "Corporate Performance and Managerial Remuneration: An Empirical Analysis." Journal of Accounting and Economics 7 (April 1985):11-42.

Murphy, Kevin J. "Executive Compensation." In Orley Ashenfelter and David Card, eds., Handbook of Labor Economics, volume 3. Amsterdam: North Holland, 1998.

Prendergast, Canice. "The Provision of Incentives in Firms." Journal of Economic Literature 37 (March 1999): 7-63.

Rappaport, Alfred. "New Thinking on How to Link Executive Pay with Performance," Harvard Business Review 77 (March-April 1999); 91-101.

Reder, Melvin W. "Wage Differentials: Theory and Measurement." In Aspects of Labor Economics. Princeton: Princeton University Press (for the National Bureau of Economic Research), 1962.

Rose, Nancy L., and Andrea Shepard. "Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?" RAND Journal of Economics 28 (Autumn 1997): 489-514.

Rosen, Sherwin. "Authority, Control, and the Distribution of Earnings." Bell Journal of Economics 13 (Fall 1982): 311-323.

Rosen, Sherwin. "Contracts and the Market for Executives." In Lars Werin and Hans Wijkander, eds., Contract Economics. Oxford: Blackwell Publishers, 1992.

Schaefer, Scott. "The Dependence of Pay-Performance Sensitivity on the Size of the Firm." Review of Economics and Statistics (August 1998): 436-449.

Slichter, Sumner. "Notes on the Structure of Wages." Review of Economics and Statistics 32 (1950): 80-91.

C Tables

TABLE 1

Variable Definitions

SALES	Firm sales (Compustat variable DATA12)
COMP	Salary, bonus, and other long-term compensation (ExecuComp variables SALARY + BONUS + OTHANN + LTIP + ALLOTHTO)
COMPPG	COMP (see above) plus the value of equity grants and the estimated Black-Scholes value of new options grants (ExecuComp variables RSTKGRNT and BLK_VALU, respectively)
VEQUITY	Value of equity held at fiscal year end (ExecuComp variables SHROWN×PRCCF)
VOPTIONS	Estimated value of options held at fiscal year end. (see Appendix B.1)
TWEALTH (A)	The total wealth of the CEO held in the firm. The variable we use for our analysis in the paper is the sum of COMPPG, VEQUITY, and VOPTIONS.
TWEALTH (B)	Like TWEALTH (A), this is a measure of total wealth of the CEO held in the firm, but with an annuity adjustment to COMPPG to estimate the present value of future cash compensation. This is done simply by multiplying COMPPG by three, and adding it to VEQUITY, and VOPTIONS.
LTW	The log of TWEALTH (A).
LOGSIZE	The log of firm sales (Compustat variable DATA12)
CEOTENURE	Number of years CEO has held the top position (Constructed as the difference between ExecuComp variable BECAME_CE and the year and month of current fiscal year.
PCMARGIN	The firm's price-cost margin, constructed as the ratio of gross operating income plus R&D to sales (Compustat variables DATA13, DATA45, and DATA12, respectively. Missing values of R&D were set equal to zero.
FOREIGN	A dummy variable equal to one if the firm reported non-negligible values of foreign income taxes (DATA64), deferred foreign taxes (DATA269), or foreign currency adjustment (DATA150). Specifically, we used FOREIGN=1 if $\lvert DATA64/DATA6\rvert \geq 0.005$ or $\lvert DATA270/DATA6\rvert \geq 0.005$ or $\lvert DATA150/DATA6\rvert \geq 0.005$, and zero otherwise.
R&D/S	R&D-to-sales ratio (Compustat: DATA46/DATA12)
R&D-DUM	Dummy variable equal to one if R&D/S>0, and zero otherwise.
ADV/S	Advertising-to-sales ratio (Compustat: DATA45/DATA12).
ADV-DUM	Dummy variable equal to one if ADV/S>0, and zero otherwise.
I/K	Investment-to-capital ratio (Compustat: DATA30/DATA8. If DATA30 is missing, we use DATA128. If DATA30 is missing, we use 0.3×DATA6, which is total assets.)
DIVERSE	A 4-digit-SIC-code-based measure of firm diversification (see Appendix B)
LOGSIGMA	The log of the standard error of the residual from the asset pricing equation. (see Appendix B).
RAWRET	The log of the firm's gross annual stock return (including dividends), $ln(1 + r)$, computed by accumulating monthly total returns as defined by CRSP (CRSP variable RET).
RPM	The "Relative Performance Measure," constructed as RAWRET minus the predictable component using aggregate pricing factors. (see Appendix B).
MKT	The "Market Index," constructed as the predicable component of RAWRET (see Appendix B).

TABLE 2

Sample Summary Statistics.

Summary statistics for the 6410 observations in our sample for 1992-1998 for which there were no missing observations for any of the following variables.

	Mean	Std	Min	Percentiles 25^{th}	50^{th}	75^{th}	Max
SALES ($mil.)	3,639	10,010	0.5	376	1,016	3,043	168,738
COMP ($thous.)	1,288	1941	-443	528	854	1,467	99,531
COMPPG ($thous.)	2,667	5,220	0	769	1,434	2,764	202,185
VEQUITY ($thous.)	59,661	579,949	0	1,256	4,879	18,782	34,222,242
VOPTIONS ($thous.)	8,348	24,255	0	644	2,391	6,956	518,895
TWEALTH ($thous.)	69,298	582,581	147	4,324	11,107	32,265	34,222,833
LTW	9.435	1.558	4.992	8.372	9.315	10.382	17.348
LOGSIZE	6.965	1.582	-0.691	5.931	6.924	8.021	12.036
TENURE	8.478	7.588	0.000	3.167	6.417	10.917	53.917
PCMARGIN	0.198	0.138	-0.200	0.110	0.170	0.267	1.000
FOREIGN	0.298	0.457	0.000	0.000	0.000	1.000	1.000
R&D/S	0.036	0.082	0.000	0.000	0.000	0.034	0.500
R&D-DUM	0.394	0.489	0.000	0.000	0.000	1.000	1.000
ADV/S	0.010	0.028	0.000	0.000	0.000	0.000	0.200
ADV/S-DUM	0.200	0.400	0.000	0.000	0.000	0.000	1.000
I/K	0.244	0.158	0.000	0.134	0.205	0.317	1.000
DIVERSE	0.418	0.493	0.000	0.000	0.000	1.000	1.000
LOGSIGMA	-2.480	0.436	-3.604	-2.812	-2.481	-2.160	-0.969
RAWRET	0.107	0.381	-3.497	-0.078	0.122	0.314	2.275
RPM	-0.023	0.305	-2.428	-0.164	-0.012	0.130	1.771
MKT	0.130	0.238	-1.458	0.012	0.131	0.260	1.345

Notes: i) The minimum value of -443.726 for COMP is not a typo. In November 1998, the compensation committee for Hewlett-Packard that under the terms of his long-term incentive plan (LTIP), Lewis E. Platt, the CEO, was required to forfeit part of his share holdings, the value of which was reflected as a negative LTIP pay-out in fiscal 1998. ii) The highest value of CEO wealth held in the firm is for William Gates III, CEO of Microsoft Corp., who in 1997 held $34,222,833,000 of wealth in Microsoft (of which all but $591,350 was equity). iii) The CEO with the longest tenure in our sample is the CEO of Kaman Corp, Charles H. Kaman, who has been CEO for 53 years since founding the company. iv) The minimum values of RAWRET, RPM and MKT fall below -1.0 because the data are measured on logs of the gross return.

TABLE 3

Measures of CEO Compensation and Wealth, 1992-98.

Medians of firm size and components of CEO compensation and wealth over time. "S&B" denotes salary and bonus plus other sources of cash compensation. "Grants" refers to new grants of equity and stock options. "Equity Value" and "Option Value" refer to the respective values of stocks and options held at the end of the fiscal year (including the value of grants during the current year). "Total Wealth (A)" is the sum of S&B, Equity Value, and Option Value. "Total Wealth (B)" measures the sum of 3× S&B plus equity and option value.

		Medians ($thousands)						
Year	#obs.	Firm Sales	S&B	S&B+ Grants	Equity Value	Option Value	Total Wealth (A)	Total Wealth (B)
1992	638	1,170,000	762	1,126	4,165	1,470	8,014	10,228
1993	917	1,033,000	756	1,126	4,641	1,556	9,129	11,046
1994	957	936,000	815	1,318	4,260	1,777	8,963	11,453
1995	1003	911,000	832	1,302	4,817	2,257	10,843	13,430
1996	1084	930,000	846	1,520	5,078	2,897	11,521	14,239
1997	1045	1,076,000	988	1,900	6,670	4,369	16,408	19,095
1998	766	1,219,000	1,010	2,062	5,486	3,903	13,597	16,644

TABLE 4

Cash Compensation on Relative Performance (RPM) and Market Index (MKT)

Regressions of log CEO cash compensation on talent variables and performance measures, where relative performance (RPM) is measured as the abnormal return according to an asset-pricing model. We use a factor model estimated with ex ante monthly data using the value-weighted market return, the 90-day Treasury bill rate, a value-weighted portfolio of same-size-decile stock returns, and a value-weighted portfolio of same-industry (2-digit SIC) stock returns. The systematic or "market" component of the firm's return is denoted by the variable MKT (see Table 2 for more details). The dependent variable is the log of cash compensation plus the value of stocks and stock options granted during the current fiscal year (COMPPG – see Table 2). Year and industry dummies are included, but not reported.

	Full Sample	Year						
		1992	1993	1994	1995	1996	1997	1998
CONSTANT	4.081	4.005	4.419	4.683	4.440	4.775	4.081	3.516
	(0.153)	(0.446)	(0.362)	(0.365)	(0.332)	(0.367)	(0.397)	(0.571)
LOGSIZE	0.406	0.371	0.359	0.380	0.412	0.433	0.467	0.398
	(0.008)	(0.022)	(0.019)	(0.019)	(0.018)	(0.019)	(0.021)	(0.030)
CEOTENURE	0.002	0.013	0.007	0.000	0.000	-0.003	-0.001	0.001
	(0.001)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)
PCMARGIN	0.626	1.075	0.763	0.545	0.596	0.316	0.501	0.726
	(0.083)	(0.265)	(0.219)	(0.207)	(0.190)	(0.202)	(0.215)	(0.272)
R&D/S	1.818	2.166	1.994	1.989	1.459	1.903	1.993	2.062
	(0.164)	(0.655)	(0.482)	(0.438)	(0.369)	(0.387)	(0.404)	(0.503)
R&D-DUM	-0.013	0.029	-0.082	0.081	0.017	0.024	-0.013	-0.181
	(0.029)	(0.083)	(0.073)	(0.073)	(0.066)	(0.071)	(0.077)	(0.101)
ADV-DUM	0.065	0.040	0.012	0.036	0.069	0.196	0.194	-0.019
	(0.034)	(0.077)	(0.068)	(0.097)	(0.089)	(0.096)	(0.097)	(0.135)
ADV/S	1.516	2.745	2.125	2.266	1.674	0.182	-0.242	1.859
	(0.495)	(1.234)	(1.047)	(1.360)	(1.182)	(1.295)	(1.316)	(1.944)
I/K	0.324	-0.036	0.184	-0.026	0.227	0.187	0.532	0.792
	(0.073)	(0.228)	(0.180)	(0.183)	(0.160)	(0.177)	(0.197)	(0.265)
DIVERSE	0.106	0.100	0.067	0.174	0.107	0.127	-0.004	0.222
	(0.021)	(0.061)	(0.052)	(0.053)	(0.048)	(0.053)	(0.059)	(0.079)
FOREIGN	0.072	0.071	0.056	-0.030	0.068	0.007	0.064	0.251
	(0.023)	(0.068)	(0.058)	(0.061)	(0.052)	(0.057)	(0.061)	(0.082)
LOGSIGMA	0.262	0.212	0.205	0.377	0.261	0.375	0.247	0.154
	(0.032)	(0.088)	(0.079)	(0.080)	(0.072)	(0.078)	(0.084)	(0.121)
RPM	0.129	0.230	0.151	0.169	0.178	0.151	0.145	0.066
	(0.031)	(0.080)	(0.077)	(0.087)	(0.080)	(0.082)	(0.090)	(0.101)
MKT	0.463	0.661	0.477	0.462	0.591	0.747	0.526	0.278
	(0.046)	(0.244)	(0.162)	(0.149)	(0.101)	(0.119)	(0.123)	(0.130)
NOBS	6371	616	895	934	978	1060	1024	744
Adj-R^2	0.419	0.442	0.371	0.383	0.479	0.430	0.434	0.327

TABLE 5

CEO Wealth on Relative Performance (RPM) and Market Index (MKT)

Regressions of log CEO cash compensation on talent variables and performance measures, where relative performance (RPM) is measured as the abnormal return according to an asset-pricing model. We use a factor model estimated with ex-ante monthly data using the value-weighted market return, the 90-day Treasury bill rate, a value-weighted portfolio of same-size-decile stock returns, and a value-weighted portfolio of same-industry (2-digit SIC) stock returns. The systematic or "market" component of the firm's return is denoted by the variable MKT (see Table 2 for more details). The dependent variable is the log of cash compensation plus the value of stocks and stock options granted during the current fiscal year (COMPPG – see Table 2). Year and industry dummies are included, but not reported.

	Full	Year						
	Sample	1992	1993	1994	1995	1996	1997	1998
CONSTANT	7.548	7.160	7.519	8.178	7.595	7.634	7.285	8.924
	(0.225)	(0.800)	(0.624)	(0.565)	(0.547)	(0.530)	(0.535)	(0.632)
LOGSIZE	0.400	0.378	0.366	0.376	0.389	0.417	0.456	0.408
	(0.011)	(0.040)	(0.033)	(0.030)	(0.029)	(0.028)	(0.028)	(0.033)
CEOTENURE	0.074	0.079	0.075	0.070	0.077	0.076	0.074	0.065
	(0.001)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
PCMARGIN	1.336	1.514	1.550	0.981	1.344	1.230	1.487	1.288
	(0.122)	(0.476)	(0.377)	(0.322)	(0.314)	(0.292)	(0.290)	(0.301)
R&D/S	0.544	1.060	0.291	0.321	0.339	0.890	0.781	0.676
	(0.241)	(1.175)	(0.832)	(0.679)	(0.608)	(0.560)	(0.545)	(0.557)
R&D-DUM	-0.149	-0.341	-0.159	-0.222	-0.127	-0.096	-0.099	-0.100
	(0.043)	(0.149)	(0.126)	(0.114)	(0.108)	(0.103)	(0.104)	(0.112)
ADV-DUM	0.141	0.143	0.299	-0.013	0.172	0.065	0.068	0.090
	(0.051)	(0.138)	(0.118)	(0.151)	(0.147)	(0.138)	(0.131)	(0.150)
ADV/S	3.782	4.897	2.747	6.339	3.557	3.881	3.585	4.069
	(0.728)	(2.213)	(1.807)	(2.107)	(1.948)	(1.871)	(1.775)	(2.150)
I/K	1.232	2.371	1.327	1.170	1.285	0.723	0.877	1.258
	(0.107)	(0.409)	(0.311)	(0.284)	(0.264)	(0.256)	(0.265)	(0.293)
DIVERSE	-0.054	0.042	-0.005	-0.000	-0.056	-0.142	-0.168	0.044
	(0.032)	(0.109)	(0.090)	(0.083)	(0.080)	(0.077)	(0.079)	(0.088)
FOREIGN	0.064	-0.040	-0.021	0.054	0.075	0.102	0.130	0.028
	(0.034)	(0.122)	(0.101)	(0.095)	(0.087)	(0.082)	(0.082)	(0.091)
LOGSIGMA	0.706	0.686	0.722	0.839	0.724	0.724	0.632	0.646
	(0.047)	(0.159)	(0.137)	(0.124)	(0.120)	(0.113)	(0.114)	(0.133)
RPM	0.385	0.571	0.434	0.268	0.264	0.280	0.327	0.682
	(0.046)	(0.144)	(0.133)	(0.134)	(0.131)	(0.118)	(0.122)	(0.112)
MKT	1.871	1.501	1.766	2.196	1.828	2.223	1.865	1.696
	(0.068)	(0.438)	(0.279)	(0.230)	(0.167)	(0.172)	(0.165)	(0.143)
NOBS	6371	616	895	934	978	1060	1024	744
Adj-R^2	0.477	0.437	0.396	0.454	0.495	0.506	0.506	0.523

TABLE 6

Alternative Relative Performance Measures (RPMs) and Timing Assumptions

Models 1, 2 and 3 report results using alternative asset-pricing models to measure relative performance and market returns. The three models use the following sets of portfolios, respectively: Model 1: {VWM, T90}; Model 2: {VWM, T90, IND, SIZ}; Model 3: {VWM, T90, SMB, HML}. These portfolios are the value-weighted market (VWM), 90-day Treasury (T90), 2-digit industry (IND), size decile (SIZ), and the Fama-French "HML" and "SMB" portfolios (see Appendix B for details). Only coefficients for RPM and MKT are reported; the control variables reported in Table 5, as well as year and industry dummies, were included but are not reported.

	Log Cash Compensation					
	Model 1		Model 2		Model 3	
LOGSIGMA	0.276	0.278	0.266	0.267	0.276	0.276
	(0.033)	(0.032)	(0.031)	(0.030)	(0.033)	(0.032)
RPM	0.151	0.245	0.153	0.285	0.139	0.250
	(0.031)	(0.035)	(0.033)	(0.036)	(0.032)	(0.036)
RPM(-1)		0.162		0.154		0.168
		(0.033)		(0.035)		(0.034)
RPM(-2)		0.153		0.167		0.155
		(0.032)		(0.033)		(0.032)
MKT	0.838	0.531	0.520	0.311	0.691	0.441
	(0.067)	(0.083)	(0.052)	(0.056)	(0.058)	(0.066)
MKT(-1)		0.386		0.440		0.379
		(0.088)		(0.065)		(0.078)
MKT(-2)		0.168		0.234		0.215
		(0.080)		(0.059)		(0.068)
NOBS	5716	5712	5716	5712	5716	5712
Adj-R^2	0.445	0.450	0.440	0.450	0.443	0.450
	Log CEO Wealth					
	Model 1		Model 2		Model 3	
LOGSIGMA	0.693	0.710	0.712	0.716	0.688	0.681
	(0.049)	(0.047)	(0.046)	(0.045)	(0.049)	(0.047)
RPM	0.457	0.847	0.382	0.743	0.401	0.786
	(0.046)	(0.051)	(0.050)	(0.053)	(0.048)	(0.052)
RPM(-1)		0.439		0.397		0.383
		(0.049)		(0.051)		(0.050)
RPM(-2)		0.183		0.177		0.193
		(0.046)		(0.048)		(0.047)
MKT	2.721	1.528	1.885	1.360	2.289	1.476
	(0.099)	(0.120)	(0.077)	(0.081)	(0.086)	(0.097)
MKT(-1)		1.413		1.035		1.334
		(0.127)		(0.095)		(0.113)
MKT(-2)		1.211		1.046		0.930
		(0.116)		(0.087)		(0.099)
NOBS	5716	5712	5716	5712	5716	5712
Adj-R^2	0.500	0.525	0.491	0.521	0.496	0.520

TABLE 7

RPM and MKT Sensitivities by Size Quartile and Market Index (MKT)

Using portfolio set {VWM, T90, IND, SIZ}, the sample is divided (in each year) into four parts according to size quartile (indicated in the table as Q1-Q4). Year and industry dummies are included, but not reported.

	Log Cash Compensation				**Log CEO Wealth**			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
CONSTANT	5.710	3.489	3.658	3.725	8.325	4.686	6.644	9.093
	(0.320)	(0.581)	(0.486)	(0.367)	(0.441)	(0.837)	(0.809)	(0.492)
LOGSIZE	0.281	0.482	0.425	0.388	0.142	0.517	0.536	0.444
	(0.032)	(0.069)	(0.055)	(0.026)	(0.044)	(0.100)	(0.092)	(0.035)
CEOTENURE	-0.003	0.000	0.005	0.009	0.081	0.062	0.071	0.091
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
PCMARGIN	0.478	1.225	1.569	1.516	1.017	2.583	2.777	1.357
	(0.123)	(0.216)	(0.213)	(0.281)	(0.169)	(0.312)	(0.355)	(0.376)
R&D	1.179	1.047	-2.883	-1.802	-0.489	-3.027	-2.730	-0.472
	(0.251)	(0.595)	(0.682)	(0.986)	(0.346)	(0.858)	(1.136)	(1.320)
R&D-DUM	-0.003	-0.094	0.182	0.014	-0.229	0.069	0.044	-0.324
	(0.062)	(0.063)	(0.058)	(0.064)	(0.085)	(0.090)	(0.097)	(0.086)
ADV/S	-0.123	2.618	-1.290	3.518	4.893	1.482	2.038	6.647
	(1.190)	(1.078)	(0.871)	(0.915)	(1.640)	(1.555)	(1.450)	(1.225)
ADV-DUM	0.190	0.002	0.185	-0.066	-0.008	0.315	0.168	-0.001
	(0.087)	(0.071)	(0.064)	(0.062)	(0.120)	(0.103)	(0.107)	(0.083)
I/K	0.286	0.255	0.529	0.678	0.939	1.242	1.803	1.692
	(0.119)	(0.143)	(0.152)	(0.206)	(0.164)	(0.207)	(0.253)	(0.275)
DIVERSE	0.014	0.052	0.119	0.148	-0.226	-0.085	0.040	-0.001
	(0.050)	(0.044)	(0.039)	(0.043)	(0.069)	(0.063)	(0.066)	(0.057)
FOREIGN	-4.678	0.126	0.067	0.099	0.018	0.000	-0.037	0.204
	(0.054)	(0.046)	(0.044)	(0.047)	(0.075)	(0.067)	(0.074)	(0.063)
LOGSIGMA	0.377	0.318	0.441	0.163	0.825	0.794	0.707	0.672
	(0.067)	(0.064)	(0.061)	(0.073)	(0.093)	(0.093)	(0.101)	(0.098)
RPM	0.027	0.108	0.307	0.319	0.267	0.438	0.329	0.740
	(0.049)	(0.062)	(0.069)	(0.089)	(0.067)	(0.090)	(0.116)	(0.119)
MKT	0.292	0.447	0.515	0.783	1.741	1.617	2.039	2.679
	(0.076)	(0.090)	(0.101)	(0.130)	(0.105)	(0.130)	(0.169)	(0.174)
NOBS	1408	1595	1647	1640	1408	1595	1647	1640
Adj-R^2	0.220	0.223	0.266	0.343	0.495	0.460	0.427	0.551

TABLE 8

RPM and MKT Sensitivities by Size Quartile and Market Index (MKT)

This table reports interactions of RPM and MKT with size quartiles as a parsimonious alternative to the sample splits by size quartile reported in Table 7. The robustness of the size interactions is investigated by adding intereactions with tenure quartiles, a 0-1 dummy variable indicating whether the CEO was an "outsider" when hired, and the Herfindahl index for the firm's 4-digit industry. The measures of RPM and MKT are based on the portfolio set {VWM, T90, IND, SIZ}. Year and industry dummies are included, but not reported.

	Log Cash Compensation				**Log CEO Wealth**			
CONSTANT	4.244	4.236	4.764	2.974	7.867	7.874	6.636	8.691
	(0.153)	(0.153)	(0.276)	(0.377)	(0.225)	(0.225)	(0.412)	(0.564)
LOGSIZE	0.385	0.385	0.404	0.396	0.371	0.372	0.364	0.306
	(0.009)	(0.009)	(0.011)	(0.012)	(0.013)	(0.013)	(0.016)	(0.018)
CEOTENURE	0.002	0.002	0.002	0.003	0.074	0.071	0.078	0.070
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
PCMARGIN	0.706	0.704	0.794	0.533	1.499	1.494	1.482	1.607
	(0.084)	(0.084)	(0.109)	(0.109)	(0.123)	(0.123)	(0.163)	(0.164)
R&D/S	1.684	1.700	1.607	2.329	0.346	0.316	0.066	0.841
	(0.164)	(0.164)	(0.212)	(0.178)	(0.241)	(0.241)	(0.317)	(0.266)
R&D-DUM	-0.024	-0.025	-0.044	0.015	-0.177	-0.176	-0.177	-0.113
	(0.029)	(0.029)	(0.036)	(0.030)	(0.042)	(0.042)	(0.054)	(0.045)
ADV/S	1.431	1.403	1.851	1.668	3.678	3.701	3.118	4.655
	(0.493)	(0.493)	(0.570)	(0.527)	(0.723)	(0.722)	(0.851)	(0.787)
ADV-DUM	0.057	0.058	0.017	0.012	0.125	0.124	0.168	0.076
	(0.034)	(0.034)	(0.042)	(0.041)	(0.050)	(0.050)	(0.063)	(0.062)
I/K	0.333	0.329	0.335	0.343	1.209	1.181	1.449	1.009
	(0.073)	(0.073)	(0.095)	(0.099)	(0.107)	(0.107)	(0.141)	(0.149)
DIVERSE	0.104	0.103	0.164	0.108	-0.057	-0.057	-0.117	0.022
	(0.021)	(0.021)	(0.027)	(0.028)	(0.031)	(0.031)	(0.041)	(0.041)
FOREIGN	0.070	0.070	0.045	0.035	0.067	0.070	0.100	0.043
	(0.023)	(0.023)	(0.029)	(0.026)	(0.034)	(0.034)	(0.044)	(0.039)
LOGSIGMA	0.258	0.256	0.258	0.097	0.734	0.730	0.631	0.273
	(0.030)	(0.030)	(0.039)	(0.041)	(0.044)	(0.044)	(0.058)	(0.062)
RPM	0.011	0.063	-0.058	0.094	0.263	0.247	0.237	0.373
	(0.049)	(0.073)	(0.090)	(0.070)	(0.071)	(0.106)	(0.135)	(0.104)
MKT	0.207	0.114	0.304	0.341	1.584	1.105	1.368	1.609
	(0.067)	(0.094)	(0.118)	(0.100)	(0.098)	(0.138)	(0.177)	(0.149)
NOBS	6365	6359	3784	3649	6365	6359	3784	3649
Adj-R^2	0.424	0.425	0.456	0.473	0.486	0.488	0.516	0.443

(Table 8 Continued on Next Page →)

TABLE 8, Continued.

	Log Cash Compensation				Log CEO Wealth			
Interactions with Size Quartile Dummies.								
RPM×S2	0.099	0.113	0.188	0.248	0.194	0.201	0.300	0.325
	(0.078)	(0.078)	(0.108)	(0.099)	(0.114)	(0.115)	(0.162)	(0.148)
RPM×S3	0.289	0.298	0.362	0.093	0.074	0.075	0.060	0.175
	(0.086)	(0.086)	(0.114)	(0.112)	(0.127)	(0.126)	(0.170)	(0.168)
RPM×S4	0.290	0.300	0.389	0.260	0.423	0.428	0.378	0.598
	(0.099)	(0.100)	(0.120)	(0.124)	(0.146)	(0.146)	(0.179)	(0.185)
MKT×S2	0.214	0.216	0.196	0.183	0.002	-0.059	0.312	-0.157
	(0.091)	(0.092)	(0.126)	(0.123)	(0.134)	(0.134)	(0.188)	(0.184)
MKT×S3	0.470	0.443	0.271	0.291	0.610	0.575	0.759	0.405
	(0.103)	(0.103)	(0.143)	(0.137)	(0.151)	(0.152)	(0.213)	(0.204)
MKT×S4	0.763	0.746	0.578	0.470	1.170	1.130	1.267	0.837
	(0.121)	(0.121)	(0.150)	(0.158)	(0.178)	(0.178)	(0.223)	(0.236)
Interactions with Tenure Quartile Dummies.								
RPM×T2		-0.123				-0.030		
		(0.095)				(0.139)		
RPM×T3		-0.031				0.066		
		(0.086)				(0.127)		
RPM×T4		-0.092				-0.000		
		(0.089)				(0.130)		
MKT×T2		0.294				0.498		
		(0.107)				(0.157)		
MKT×T3		0.098				0.639		
		(0.102)				(0.149)		
MKT×T4		0.020				0.811		
		(0.111)				(0.163)		
Interactions with "Outsider" Dummy.								
OUT			0.221				0.046	
			(0.031)				(0.047)	
RPM×OUT			0.107				-0.002	
			(0.086)				(0.129)	
MKT×OUT			-0.284				0.102	
			(0.112)				(0.167)	
Interactions with Herfindahl Index.								
HRF				-0.032				-0.067
				(0.016)				(0.024)
RPM×HRF				-0.163				-0.186
				(0.056)				(0.084)
MKT×HRF				0.039				0.057
				(0.078)				(0.116)