

Department of Finance Working Paper Series 1999

FIN-98-090

The Impact of the Likelihood of Turnover on Executive Compensation

Jay C. Hartzell

October 1998

**This working paper series has been generously supported by a grant from
CDC Investment Management Corporation**

THE IMPACT OF THE LIKELIHOOD OF TURNOVER ON EXECUTIVE COMPENSATION

Jay C. Hartzell*
Stern School of Business
New York University

First Draft: August 1997
This Draft: October 1998

Abstract: This study analyzes the role of three incentive devices in managerial compensation: pay for performance, termination, and career concerns. A model is derived which shows that the three incentives are substitutes; where the termination (or career concerns) incentive is low, the optimal contract contains stronger pay-for-performance incentives. The empirical implication, then, is that the pay-for-performance sensitivity of managers should be decreasing (increasing) in the probability of termination (retirement). To test the model's predictions, I first use a sample of CEOs to estimate the probabilities of forced and voluntary turnover. Then, these estimated probabilities are compared to the CEOs' estimated pay-for-performance sensitivity. The evidence is consistent with the hypothesis that boards consider the likelihood of termination when setting the compensation contract; the relationship between changes in CEO compensation and firm performance is decreasing in the estimated probability of forced turnover. While CEOs nearing retirement do not appear to have compensation that is increasingly sensitive to performance, their wealth does have increased sensitivity. Consistent with the model's intuition, the sensitivity of total CEO firm-related wealth to performance is positively related to the probability of voluntary turnover.

JEL Classification: G34

Keywords: Compensation, Executive turnover, Agency problems, Optimal contracts

*Please address correspondence to Jay Hartzell, Department of Finance, NYU – Stern School of Business, 44 West 4th Street, Suite 9-190, New York, NY 10012; e-mail: jhartzel@stern.nyu.edu, phone: (212) 998-0359, fax: (212) 995-4233. I would like to thank David Blackwell, Greg Brown, John Butler, David Chapman, Greg Hallman, John Hund, Mark Huson, Wendell Licon, Robert Parrino, Dan Quan, Dan Slesnick, and Sheridan Titman for their helpful insights and comments, as well as seminar participants at the 1998 FMA meetings, the University of Alberta, University of Florida, Georgia State University, University of Michigan – Ann Arbor, New York University, University of North Carolina – Chapel Hill, Penn State University, Rice University, Southern Methodist University, Stanford University, University of Texas, and Tulane University. I would especially like to thank Robert Parrino, for graciously supplying data and Laura Starks, the chair of my dissertation committee, for her guidance and assistance. The Eugene and Dora Bonham Memorial Fund provided financial assistance.

THE IMPACT OF THE LIKELIHOOD OF TURNOVER ON EXECUTIVE COMPENSATION

Abstract: This study analyzes the role of three incentive devices in managerial compensation: pay for performance, termination, and career concerns. A model is derived which shows that the three incentives are substitutes; where the termination (or career concerns) incentive is low, the optimal contract contains stronger pay-for-performance incentives. The empirical implication, then, is that the pay-for-performance sensitivity of managers should be decreasing (increasing) in the probability of termination (retirement). To test the model's predictions, I first use a sample of CEOs to estimate the probabilities of forced and voluntary turnover. Then, these estimated probabilities are compared to the CEOs' estimated pay-for-performance sensitivity. The evidence is consistent with the hypothesis that boards consider the likelihood of termination when setting the compensation contract; the relationship between changes in CEO compensation and firm performance is decreasing in the estimated probability of forced turnover. While CEOs nearing retirement do not appear to have compensation that is increasingly sensitive to performance, their wealth does have increased sensitivity. Consistent with the model's intuition, the sensitivity of total CEO firm-related wealth to performance is positively related to the probability of voluntary turnover.

1. Introduction

The potential motivation problem inherent in managerial relationships has long been realized (e.g., Jensen and Meckling (1976)). Managers acting in their own self-interest may make decisions that are suboptimal in the eyes of shareholders. By providing incentives so that managers' interests are aligned with those of the shareholders, the problem can be ameliorated. The most important of these incentives discussed in the agency-theory literature derive from the external and internal labor markets, the threat of termination, and compensation tied to observable performance.

In this paper, I study the interactions among a set of these incentive devices. Specifically, I develop a theoretical model that demonstrates that the incentives provided by the labor market, the threat of termination, and variable pay are substitutes. This is an intuitive result; the three incentive devices are different means of connecting a manager's current or future pay to performance. As such, the weakening of one component is associated with the strengthening of another, with the result that managers are still encouraged to take costly actions. For example, all else equal, managers who are less likely to be terminated have pay structures more sensitive to performance. Similarly, managers approaching retirement (who therefore place less weight on future compensation) also have more sensitive pay.

These predictions arise from a four-period model with two-period contracts. The use of a sequence of multi-period contracts is atypical in the literature, and the advantage of this approach is that it allows for the incorporation of both termination within contracts and career concerns across contracts. The contracts specify both the safe salary and the percentage of profits that will be paid to the manager in each period. The profit-sharing component of the contract provides the classic pay-for-performance incentive, leading managers to exert more (costly) effort on the margin, because they expect to receive higher wages from doing so.

Firm profits are the sum of the manager's unobserved ability, the effort he or she exerts, and a random error term. Both the firm and manager use observed profit to update their beliefs about managerial ability.¹ If in the middle of a contract, the revised estimate of the incumbent manager's ability is sufficiently low, the firm fires the manager.² Managers fired after the first period of a contract receive no wages in the second period of that contract. This possibility provides motivation for managers; on the margin, they work harder in order to increase the firm's estimate of their ability, thereby decreasing the probability of being fired.

¹ In order to isolate the moral hazard problem, I assume that the firm and manager have identical information about the manager's ability. If the manager had superior information, there would also be an adverse selection problem, which could obfuscate the issues addressed here.

During the first contract, managers have an additional incentive provided by the labor market. New contracts are negotiated after the first contract expires (at the end of the second period), and managers believed to have higher ability receive better second contracts. Thus, by exerting more effort in the first and second periods, the manager increases firm profits and the estimate of his or her ability. This in turn leads to higher expected utility from the second contract.

The manager may also leave the model by retiring, which occurs at the beginning of each period with some known exogenous probability. More likely retirement weakens both the termination and career concerns incentives. If a manager is likely to retire, then being fired poses less of a threat. Similarly, when choosing their effort level, managers nearing retirement place less weight on the utility provided by future contracts, rendering that source of motivation less effective.

When predicting the manager's response to the offered contract, the firm recognizes the effects of the three incentives, pay for performance, termination, and career concerns. Since the pay-for-performance incentive is the only one of these three devices under the control of the firm, it is set to work in conjunction with the firm's assessments of the other two incentives in order to elicit the desired level of managerial effort. If the firm has a high level of impediments to firing the manager, making termination unlikely, it offers a contract with a high sensitivity of pay to performance. Similarly, if the manager is likely to retire, the firm also offers a higher profit-sharing component. Finally, the analysis shows an interaction between the probabilities of these two types of turnover; the relationship between the probability of termination and pay-for-performance sensitivity is decreasing in the probability of retirement.

These implications are tested using a sample of CEOs, their patterns of turnover and compensation. The test is conducted in two stages. First, using a competing-risk hazard model, the respective probabilities of forced and voluntary turnover are estimated as functions of performance and characteristics of the CEO, his or her firm, and its industry. The likelihood of forced turnover is negatively related to performance, and this relationship is stronger for firms in homogeneous industries and firms whose officers and directors (other than the CEO) own less of the firm's stock. The likelihood of voluntary CEO departure is increasing in the age of the CEO. CEOs of poor-performing firms with either outsider-dominated boards or high levels of ownership by investment companies are also more likely to leave voluntarily.

² Here, "sufficiently low" means that the expected wages under the incumbent manager exceed expected profits by some known amount.

Next, the fitted values from the empirical turnover models are compared to the CEOs' compensation over the 1992-1995 period. Changes in CEO pay are regressed on returns, where the sensitivity of changes in CEO wealth to changes in shareholder wealth is allowed to be a function of the forecasted likelihood of turnover. The evidence is consistent with the prediction that the sensitivity of pay to performance is decreasing in the probability of *forced* departure. While the sensitivity of year-to-year compensation to performance does not appear to be related to probability of *voluntary* turnover, the sensitivity of changes in total CEO firm-related wealth to performance is increasing in this probability, consistent with the model's predictions.

Most of the previous research on the managerial agency problem focuses on one incentive device at a time. The relationship between compensation and performance is modeled by many, including Holmström (1979), Shavell (1979), Harris and Raviv (1979), Grossman and Hart (1983), and Murphy (1986). The use of termination as an incentive device is studied in Stiglitz and Weiss (1983) and Shapiro and Stiglitz (1984).³ The effect of the labor market on managerial motivation is addressed by Fama (1980), among others.

Comparatively little research analyzes the potential interactions between incentive devices. My work incorporates ideas from two exceptions, Gibbons and Murphy (1992) and Hallman and Hartzell (1997). Gibbons and Murphy (1992) investigate the role of career concerns in optimal incentive compensation. They argue that as wages from future contracts become less important to the manager, the firm should tie the manager's pay more closely to performance. Their empirical tests focus on CEOs in their last years before retirement and find that these CEOs have pay more sensitive to performance than do CEOs who are not near retirement. Hallman and Hartzell (1997) also study the relationship between incentive devices; they show that the sensitivity of a manager's pay to performance should be increasing in the cost of firing. As an empirical test, Hallman and Hartzell (1997) compare the pay-performance sensitivity of a sample of CEOs of real estate investment trusts (REITs) to that of general partners of real estate partnerships (who are more costly to fire). The results support their hypothesis; they find that the pay-performance sensitivity of the REIT managers is much less than that of the general partners.

Unlike Gibbons and Murphy (1992), the model here makes predictions about the effect of termination on the pay-for-performance relationship. This model also differs from that of Hallman and Hartzell (1997), who ignore the complicating effects of career concerns by analyzing a single contract. The model constructed here incorporates both of these important labor-market incentives, allowing for an

³ A study that combines firing, retiring, tenuring, and compensation is Acharya (1992). The focus of that paper, though, is on the optimal firing and tenuring rules, and the impact of these decisions on firm value, rather than the structure and incentive effects of the optimal contract.

analysis of the optimal contract in a more realistic setting, where managers face retirement as well as the threat of termination. This setting also provides a framework for predicting any cross-effects on pay for performance between the respective probabilities of termination and retirement.

The remainder of this paper is organized as follows: The next section discusses the model of compensation. Section three analyzes the model's comparative statics and empirical implications. The fourth section presents the estimated models of CEO turnover and section five discusses tests of the relationship between the estimated probabilities of turnover and the sensitivity of CEO pay to performance. Section six concludes.

2. A Model of Optimal Compensation

The model captures the intuition that the optimal compensation contract incorporates the probabilities of retirement and dismissal. To find the equilibrium, I first solve the manager's problem, the choice of effort, taking the parameters of the labor contract as given. Then, I solve the firm's problem, the choice of contract parameters, taking the manager's effort response as given. The intersection of these two solutions gives a Nash equilibrium.

The intuition behind the manager's problem is straightforward. Exerting more effort is costly, but has two consequences on the margin: (1) the firm has higher expected profits (of which the manager receives a percentage), and (2) since effort and ability are unobserved, these higher profits lead to a higher consensus expectation of the manager's ability. As a result, as the manager exerts more effort, he or she receives greater expected income from three sources: (1) the percentage of higher expected profits specified by the contract, (2) a lower probability of being fired and therefore, higher expected future wages, and (3) a more lucrative expected contract in the future.

Because the managers' wages are derived from these three sources, substitution effects arise. The results of the model show that as the probability of retirement increases or the probability of being fired decreases (due to increases in the impediments to firing), the optimal profit-share component of the contract increases (i.e., pay is more sensitive to performance).

There are four periods in the model and the two parties, risk-averse managers and risk-neutral firms, enter into two-period labor contracts. The contracts specify the manager's wage in each period as a linear function of the firm's profits. Managers receive both a salary, a_t , and a percentage of profits, b_t ; thus, manager's wages, w_t , are given by $w_t = a_t + b_t\pi_t$, where π_t is the firm's profit at time t .

The firm's profits are the sum of the manager's ability in that period, the effort he or she exerts, and a noise term: $\pi_t = \alpha_t + e_t + \varepsilon_t$, where α_t is the manager's ability, e_t is the manager's (positive) effort, and ε_t is a zero-mean random variable, distributed normally with variance σ_ε^2 .⁴ While the firm knows the realized profit level, it cannot observe the three separate components. The manager knows the effort exerted, but not his or her ability. Although manager ability is not known by either party, both know that first-period ability is distributed normally with mean μ_α and variance σ_α^2 , and that ability is autocorrelated, evolving as $\alpha_t = \rho\alpha_{t-1} + v_t$, where ρ is positive, and v_t is a zero-mean random normal variable with variance σ_v^2 . The error terms, ε_t and v_t , are independent over time and of each other.

Managers may stop receiving wages for two reasons: they can be fired or they can retire. While contracts last for two periods, the firm has the option of firing the manager after it observes the profit for the first period covered by the contract. Managers fired after the first period of the first contract forgo wages only in the next (second) period (i.e., they may still receive a second contract). Managers fired after the first period of the second contract (i.e., at the end of period three) receive no wages for the only remaining period (period four).

The firm will fire the manager under contract if its updated expectation of next period's profits under the incumbent manager is less than the expected wages (as specified by the contract) by more than some known amount h_i , where i indexes the number of the contract (i.e., 1 or 2). These parameters, h_1 and h_2 , are given in the model, but can be thought of more generally as functions of the firm-specific impediments to firing the incumbent. These impediments can include variables that affect the cost of firing a manager or the firm's ability to do so. Costs of firing can depend on many factors, such as the manager's severance package and the ease with which a firm can identify a potential replacement. Similarly, the firm's ability to fire a manager can depend on several factors, including the monitoring

⁴ Gibbons and Murphy (1992) also use this profit function.

mechanism in place, such as the manager's influence over the board of directors, or the presence of strong outside monitoring.

Let P_t be the *ex ante* probability that the manager will be fired at the end of period t (after π_{t-1} has been observed and w_{t-1} has been paid to the manager). This probability is a function of the manager's expected ability conditional on observed profit, and is therefore a function of the effort exerted in prior periods (since effort affects realized profit). If the manager is in the middle of the i th contract, the probability of being fired is also a function of h_i .

From the manager's perspective, retirement has a similar outcome to being fired; he or she forgoes future wages, but always for the duration of the model. The probability of retirement is exogenous and known to both parties. Let R_t be the probability of a manager retiring at the beginning of period t (before he or she exerts effort).

2.1. The Manager's Problem

The manager's objective is to choose an effort level that maximizes the sum of his or her expected utility over the remaining periods, where the expectation is conditional on the available information at the time of the decision.⁵ Manager's utility is a function of the difference between the wages received and the disutility or cost of exerting the chosen level of effort, given by $c(e_t)$. The manager is assumed to have a constant absolute risk aversion (CARA) utility function, so the problem can be reduced to maximizing the sum of his or her certainty equivalents, where the time t certainty equivalent is given by $a_t + b_t E(\pi_t) - c(e_t) - \frac{1}{2} r b_t^2 \sigma_{\pi_t}^2$, where r is the coefficient of absolute risk aversion, and $\sigma_{\pi_t}^2$ is the variance of the firm's time t profits.⁶ I assume that exerting more effort increases the cost to the manager at an increasing rate, i.e., $c'(e) > 0$ and $c''(e) > 0$. Thus, the manager's problem at the beginning of the model is

$$\begin{aligned}
& \underset{e_1, e_2, e_3, e_4}{\text{Max}} \quad a_1 + b_1 (E(\alpha_1) + e_1) - c(e_1) - \frac{1}{2} r b_1^2 \sigma_{\pi_1}^2 \\
& + (1 - R_2)(1 - P_2(e, h_1)) \left[a_2 + b_2 (E(\alpha_2) + e_2) - c(e_2) - \frac{1}{2} r b_2^2 \sigma_{\pi_2}^2 \right] \\
& + (1 - R_2)(1 - R_3) \left[a_3 + b_3 (E(\alpha_3) + e_3) - c(e_3) - \frac{1}{2} r b_3^2 \sigma_{\pi_3}^2 \right] \\
& + (1 - R_2)(1 - R_3)(1 - R_4)(1 - P_4(e, h_2)) \left[a_4 + b_4 (E(\alpha_4) + e_4) - c(e_4) - \frac{1}{2} r b_4^2 \sigma_{\pi_4}^2 \right]
\end{aligned} \tag{1}$$

⁵ For ease of computation, I ignore discounting effects. The model's conclusions are robust to a reasonable discount rate.

⁶ Note that both moments in this certainty equivalent are conditional. Section A.1. of the appendix shows how the conditional variance of profit evolves over time.

To solve (1), I start with the manager's problem at the beginning of the fourth period, where the only decision is to choose fourth-period effort to maximize remaining utility, where both the expectation and variance terms are conditional on the information available at the end of period three (e.g., π_3). The first-order condition for this problem is given by

$$e_4^* = c'^{-1}(b_4), \quad (2)$$

assuming that $c'^{-1}(\cdot)$ exists and is well-defined. (Section A.2. of the appendix shows the detail for this step and subsequent details of the model.) Working backwards in time, one can substitute this solution into the period-three problem, which yields the first-order condition,

$$b_3 - c'(e_3^*) - (1 - R_4) \frac{\partial P_4}{\partial e_3} u_4(e_4^*) = 0, \quad (3)$$

where for ease of notation, $u_4(e_4^*)$ represents the fourth-period certainty equivalent as a function of the optimal fourth-period effort. In order to solve (3) for e_3^* , one must know the functional form of the probability that the manager is fired, $P_4(e, h_2)$.

As mentioned above, at the end of odd-numbered periods (i.e., when the firm and manager are in the middle of a contract), the firm can fire the manager. It does so if the expected labor costs (given the prior period's profits) are greater than the expected profits plus the firm's level of impediments to firing the manager (h_i). Expected profits will be low when profit in the first period of the contract is low, causing the firm to revise downward its estimate of the manager's ability. Then, if this revised (conditional) estimate of next period's profit is sufficiently less than the expected wages as dictated by the agreed-upon contract, the firm opts to fire the manager rather than suffer the loss (in expectation).

The appendix derives of the probability that the manager is fired after the third period, $P_4(e, h_2)$. The comparative statics for $P_4(e, h_2)$ are what one would expect; the probability of being fired is decreasing in the effort put forth by the manager, and decreasing in the impediments to firing, i.e., $\frac{\partial P_4}{\partial e_3} < 0$, and $\frac{\partial P_4}{\partial h_2} < 0$. It can easily be shown that if the firm expects effort to be close to the optimal level, the threat of termination increases optimal effort beyond the level reached without the termination incentive (i.e., $e_3^* > c'^{-1}(b_3)$). The optimal effort level is decreasing in the probability of retirement, due to the fact that likely retirement weakens the threat of termination. In other words, a manager who

expects to retire is unlikely to be concerned about the prospect of being fired. Mathematically, this is shown by implicitly differentiating (3), which yields $\frac{\partial e_3^*}{\partial R_4} < 0$.⁷

2.2. The Firm's Problem: Periods 3 and 4

Assume that the firm wants the manager to exert effort $e_t^* = k$.⁸ The choice of optimal k is outside the model, and is held fixed as other parameters change. This can be thought of as arising from the assumption that firms do not adjust the optimal effort they want the manager to exert due to small changes in variables, like the cost of firing. In other words, the firm's problem of *effort* choice has a corner solution; the firm wants the manager to exert some (high) level of effort, which does not vary for small changes in the model's parameters.

In equilibrium, the manager's first-order conditions must be satisfied, and the contract must meet or exceed the manager's reservation wage for each period (i.e., it must satisfy the participation constraints).⁹ I choose to use a separate reservation utility level in each period instead of one *two-period* reservation wage.¹⁰ This implies that the optimal contract must satisfy four equations, two first-order conditions and two participation constraints. Solving these four equations for the four contract parameters of interest yields

$$b_4^* = c'(k), \quad (4)$$

$$a_4^* = w_4^* - b_4^* \rho \hat{\alpha}_3 - b_4^* k + c(k) + \frac{1}{2} r b_4^{*2} \sigma_{\pi_4}^2, \quad (5)$$

$$b_3^* = c'(k) - (1 - R_4) \left. \frac{\partial P_4}{\partial e_3} \right|_{e_3=k} w_4^*, \quad (6)$$

$$a_3^* = w_3^* - b_3^* (\hat{\alpha}_3 + k) + c(k) + \frac{1}{2} r b_3^{*2} \sigma_{\pi_3}^2, \quad (7)$$

⁷ Technically, this result requires that the expected effort is close to optimal effort, which should hold in equilibrium. For a detailed discussion of why this ensure the given sign, see the appendix.

⁸ The firm's desired effort level, k , could easily be allowed to vary across periods. Here, for simplicity, I assume k is constant across all four periods.

⁹ I use an equality in the participation constraint to ensure a minimum-cost contract. As explained later, the second contract's reservation wages are functions of estimated managerial ability, conditional on the outcomes of the first and second periods.

¹⁰ This avoids possible solutions that pay the worker almost all of his or her reservation wage in the first period of the contract at the expense of very little salary (utility) in the second period. It seems likely that the worker would break such a contract in the second period (e.g., by quitting).

where w_3^* and w_4^* are the manager's third- and fourth-period reservation utility levels, respectively.¹¹ Here, I have used the fact that in equilibrium, expected effort must equal the optimal choice of effort, implying that $E(e_3) = e_3^* = k$.

2.3. The Manager's problem - periods 1 & 2

Now, the manager's second-period problem is to maximize his or her certainty equivalent over periods two through four. For simplicity, let $w_{3,4}^*(\hat{\alpha}_3)$ be the manager's *two-period* expected utility level of the second contract (upon entering into the contract at the beginning of period three) as a function of his or her expected ability at that point in time. This represents the wages that the manager is able to command in negotiating the second contract, and managers believed to be of higher ability should be able to negotiate better contracts in the second period. Accordingly, while these functions are not specified in the model, I assume that $\frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} > 0$ (i.e., managers believed to be of higher ability can command a higher two-period reservation wage) and that $E(w_{3,4}^*(\hat{\alpha}_3)) - \frac{1}{2} r \text{Var}(w_{3,4}^*(\hat{\alpha}_3)) > 0$ (i.e., the certainty equivalent of the second contract is positive). For tractability, I also assume that $\frac{\partial^2 E(w_{3,4}^*)}{\partial \hat{\alpha}_3^2} = 0$ (i.e., $w_{3,4}^*(\hat{\alpha}_3)$ is a linear function).

Given this, (as the appendix shows) the manager's problem in period two has the first-order condition

$$c'(e_2^*) = b_2 + (1 - R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_2}. \quad (8)$$

A comparison of the third-period first-order condition in (3) and the second product term in (8) shows that the third-period termination incentive is replaced in the second period by motivation due to the possibility of receiving a better contract through increased effort.

Working farther back in time, the manager's first-period problem gives the first-order condition

¹¹ Note that w_4^* is the reservation wage, conditional on the worker being employed that period, and in equilibrium, $u_4(e_4^*) = w_4^*$.

¹² I assume that the variance of the certainty equivalent of the second contract, $\text{Var}(w_{3,4}^*(\hat{\alpha}_3))$, is not a function of effort. Given that the conditional variance of estimated ability is deterministic, this seems to be a reasonable assumption.

$$c'(e_1^*) = b_1 + (1 - R_2) \frac{\partial P_2}{\partial e_1} u_2(e_2^*) + (1 - R_2)(1 - R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_1}, \quad (9)$$

where the period-two firing rule is entirely analogous to $F_4(e, h)$.

The first term on the right-hand side of (9) is the direct incentive provided by the profit-sharing rule. The second one is the termination incentive; higher effort implies a lower probability of being fired, which implies a higher probability of receiving second-period wages. The third is the career concerns component, the expectation of receiving a higher expected contract in period three.

Of the various partial derivatives in (8) and (9), $\frac{\partial \hat{\alpha}_3}{\partial e_2}$ and $\frac{\partial \hat{\alpha}_3}{\partial e_1}$ can be signed directly from the rule for updating estimated managerial ability shown in the appendix: $\frac{\partial \hat{\alpha}_3}{\partial e_2} > 0$, and $\frac{\partial \hat{\alpha}_3}{\partial e_1} > 0$. Given this and assuming that $E(e_1) \approx e_1^*$, one can see from both (8) and (9) that effort in periods one and two is increasing in $\frac{\partial w_{3,4}^*}{\partial \hat{\alpha}_3}$. This makes intuitive sense; $\frac{\partial w_{3,4}^*}{\partial \hat{\alpha}_3}$ dictates the portion of increased profit from higher expected ability retained by the manager in the form of higher wages. In the extreme example where $\frac{\partial w_{3,4}^*}{\partial \hat{\alpha}_3}$ approaches zero, the firm receives almost all of the rents from increased manager ability and the manager has virtually no career concerns incentive. In this case, from (8) and (9), the only significant remaining incentives in the first contract would be those from pay-for-performance and the threat of termination.

2.4. The Firm's Problem: Periods 1 & 2

As in the firm's problem for periods three and four, I assume that the firm desires the manager to put forth effort level k , and the optimal contract must meet the participation constraint for the manager's utility in periods one and two. Solving the manager's first order conditions (equations (8) and (9)) and the two participation constraints for the four unknown contract parameters leads to the following expressions for the optimal first contract:

$$b_2^* = c'(k) - (1 - R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_2}, \quad (10)$$

$$a_2^* = w_2^* - (b_2^*)(\rho\mu_\alpha + k) + c(k) + \frac{1}{2}rb_2^{*2}\sigma_{\pi_2}^2, \quad (11)$$

$$b_1^* = c'(k) - (1 - R_2) \left. \frac{\partial P_2}{\partial e_1} \right|_{e_1=k} w_2^* - (1 - R_2)(1 - R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_1}, \quad (12)$$

$$a_1^* = w_1^* - b_1^*(\mu_\alpha + k) + c(k) + \frac{1}{2}rb_1^{*2}\sigma_{\pi_1}^2, \quad (13)$$

where I have simplified the expectations of ability at time t given the information at time zero by using $E(\hat{\alpha}_t | I_0) = \rho^{t-1}\mu_\alpha$ (from the law of iterated expectations).¹³

3. Comparative Statics

Of primary interest is the behavior of the optimal pay-for-performance incentives across contracts and time. From the above solutions, one can deduce the following:

- (1) In the first and third periods, the pay-for-performance component is increasing in the respective impediments to firing. Mathematically, $\frac{\partial b_1^*}{\partial h_1} > 0$ and $\frac{\partial b_3^*}{\partial h_3} > 0$. Proof: See the appendix.
- (2) Since the probability of being fired is decreasing in the cost of firing, in the middle of each contract (i.e., periods one and three), the optimal pay-for-performance component is decreasing in the probability of being fired.
- (3) Assuming that the probability of retiring in any period is strictly less than one, and that the probability of termination after periods one and three is positive, $b_1^* < b_4^*$, $b_2^* < b_4^*$, and $b_3^* < b_4^*$. In other words, the presence of the termination and/or career concerns incentives in the first three periods reduces pay-for-performance component below the fourth-period level (when there is neither the threat of termination nor a future contract to induce more effort in hopes of higher future wages).
- (4) Unlike Gibbons and Murphy (1992), the model does not *necessarily* predict that the pay-for-performance component is monotonically increasing over time. However, this seems likely to hold for many parameterizations of the model, and one *can* see that the pay-for-performance component for time t

¹³ I have also used the fact that in equilibrium, $E(e_t) = k$.

is increasing in the probability of retirement in future periods. This is equivalent to $\frac{\partial b_3^*}{\partial R_4} > 0$,

$$\frac{\partial b_2^*}{\partial R_3} > 0, \quad \frac{\partial b_1^*}{\partial R_2} > 0, \quad \text{and} \quad \frac{\partial b_1^*}{\partial R_3} > 0.$$

(5) As the probability of retirement increases, the relationship between the probability of termination and the pay-for-performance incentive declines. In other words, a manager who is likely to retire has little to fear from the threat of termination, necessitating an increase in the pay-for-performance component.

Mathematically, this can be expressed as $\frac{\partial^2 b_3^*}{\partial h_2 \partial R_4} < 0$ and $\frac{\partial^2 b_1^*}{\partial h_1 \partial R_2} < 0$. (For detailed expressions of these derivatives and those in point (4) above, see the appendix.)

These comparative statics are the primary testable implications of the model. They predict that the sensitivity of a manager's pay to firm performance should be decreasing (increasing) in the probability of dismissal (retirement) for some given level of performance. Second, they also predict that there is an interaction between these probabilities of turnover; a higher likelihood of retirement weakens the termination incentive (as well as the pay-for-performance incentive).

4. Empirical Tests: Estimation of the Likelihood of Turnover

The implications derived above are tested in two stages. First, I estimate the two probabilities of managerial turnover by focusing on firms' CEOs and using their patterns of turnover to estimate the likelihood of a CEO departing as a function of several independent variables. With these estimates of the likelihood of CEO departure and a sample of their compensation data, I then test the model's implications by comparing the estimated probabilities of turnover to the estimated pay-for-performance sensitivity.

The theoretical model predicts that as the firm's hesitancy to fire increases, the probability that the manager is fired decreases. Of course the firm's actual hesitancy to fire is unobservable, but one can estimate the probability of dismissal as a function of observed variables. This is the approach used here; the model is tested by comparing CEOs' relative likelihood of forced and non-forced departure with their respective sensitivities of pay to performance.

To estimate the probabilities of turnover, I use a competing risk hazard model, a natural candidate for the problem at hand.¹⁴ Of interest is the length of time a CEO is in place and possible influences on that duration. Rather than treating each CEO-year as an independent observation in a pooled logit approach, hazard models explicitly allow the probability of a CEO departing to be a function of his or her tenure. Intuitively, when estimating the probability of a CEO departing, one should allow it to be a function of whether or not that CEO departed last period. In the current context, the "competing risks" are the different reasons why a CEO may leave the firm.

In this type of model, the dependent variable is the length of time a CEO is in office.¹⁵ CEOs can leave office for two reasons: a forced departure, a proxy for being fired in the theoretical model, or a voluntary departure, which proxies for retirement. Let $\lambda_j(t; Z)$ be the cause j -specific hazard rate at time t for a person with covariates Z (where $j = 1$ for forced departures and $j = 2$ for voluntary departures). In other words, $\lambda_j(t; Z)$ is the instantaneous rate of CEO departure of type j at time t given Z and in the presence of the other failure type.¹⁶

¹⁴ For a comprehensive reference on hazard models, see Kalbfleisch and Prentice (1980).

¹⁵ For those familiar with hazard model terminology, here each CEO is one "spell."

¹⁶ Mathematically, this hazard rate can be expressed as $\lambda_j(t; Z) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t, J = j | T \geq t, Z)}{\Delta t}$ for $j = 1, 2$, where T is the CEO's random departure time due to cause J .

The hazard function is parameterized in two ways. The first uses a Cox proportional hazard specification.¹⁷ Each cause of turnover has its own underlying likelihood of occurring (i.e., its own hazard rate, λ_{0j}), and the covariates are allowed to affect each type of turnover differently (through separate coefficient vectors, β_j).¹⁸ The advantage of this technique is that no distributional assumption is required. Instead, the model is estimated via maximization of the partial likelihood function; the terms that include the underlying hazard rates, λ_{0j} , are separated and do not enter the estimation routine. CEOs who are not observed from the beginning of their tenure (termed left-censored observations), and those still in place at the end of the observation period, December 31, 1995, (termed right-censored) are incorporated into the estimation.

The second hazard model used assumes a parametric form for the hazard function, specifically the Weibull distribution. This allows for the conditional probability of each type of departure to be monotonically increasing or decreasing.¹⁹ The benefit of the Weibull model is that it explicitly incorporates the manager's tenure into the estimation of the probability of turnover, obviously at the cost of making a distributional assumption.

4.1. Sample Construction

To construct the sample of CEOs, I start with a random sample of 250 firms from all firms in the 1992 *Forbes* magazine survey of executive compensation which also have the following data available: returns from the Center for Research in Security Prices, accounting variables from Compustat, and compensation data from the Standard and Poor's ExecuComp database. The hazard model technique used to estimate the likelihood of turnover focuses on CEOs as observations (as opposed to CEO-years,

¹⁷ In this case, where a CEO's covariates at time t are given by the vector $\mathbf{z}(t)$, the hazard rate for cause j at time t can be expressed as $\lambda_j(t, Z) = \lambda_{0j}(t) \exp[\mathbf{z}(t)\beta_j]$ for $j = 1, 2$.

¹⁸ Let $t_{j1} < \dots < t_{jk_j}$ be the k_j departures of type j and let $\mathbf{z}_{ji}(t_{ji})$ be the covariates for departure t_{ji} . Then, one can construct the partial likelihood function, $L(\beta_1, \beta_2) = \prod_{j=1}^2 \prod_{i=1}^{k_j} \exp[\mathbf{z}_{ji}(t_{ji})\beta_j] / \prod_{i \in R(t_{ji})} \exp[\mathbf{z}_i(t_{ji})\beta_j]$, where $R(t_{ji})$ is the risk set for departure t_{ji} (i.e., those CEOs who have not yet departed upon entering time t_{ji}). Technically, this is only the likelihood function in the absence of ties, or CEOs who depart after the same tenure. Since I have some ties in my data, I use the Breslow (1974) approximation to the likelihood function above. Let d_{t_j} be the number of CEOs departing at time t_j for cause j . This approximation does two things: it sums the covariate vector in the numerator above across the d_{t_j} departing CEOs, and it raises the denominator to the d_{t_j} power.

¹⁹ The hazard rates are given by $\lambda_j(t, Z) = pt^{p-1} \exp[\mathbf{z}(t)\beta_j]$, where p is a shape parameter to be estimated from the data. For the special case of $p = 1$, the hazard rate is constant implying an exponential distribution. If $p > 1$, the hazard rate is increasing, implying that the probability of turnover (given survival to date t) is increasing in time. Conversely, for $p < 1$, the probability of turnover is decreasing in managerial tenure.

for example). Accordingly, for each of the 250 firms, my sample includes the 1992 *Forbes* survey incumbent CEO, plus any CEO hired since 1980.

The *Forbes* surveys and the ExecuComp database provide each CEO's age and tenure as CEO. For each CEO change, the exact date of turnover is obtained from the *Wall Street Journal*.²⁰ If the succession announcement states that the CEO was forced out, the departure is classified as involuntary (or forced). As in Borokhovich et al. (1996), CEOs who are less than 60 years old when they leave the position for reasons other than health or other employment are also classified as leaving involuntarily. Four CEOs whose firms were merged and are not employed by the successor firm are also classified as having involuntarily departed as of the first merger announcement date. All other departures are voluntary. Four CEOs whose firms merged but were employed by the successor firm were classified as right-censored (i.e., still in place, but no longer observed) as of the merger announcement date.

This procedure results in an initial sample of 175 turnovers. Of these turnovers, 31 are classified as involuntary departures using the above procedure. Seven of the 31 involuntary departures have announcements explicitly referring to the CEO being forced out of the firm. Of the 144 voluntary turnovers, 77 announcements give retirement as the primary cause of CEO departure, not including 47 CEOs who resign the office of CEO but become or remain Chairman of the board of directors.

For each CEO, I collect stock ownership and board composition variables from the firm's proxy statements for approximately every three years through 1995.²¹ From each proxy statement, I calculate the percentage of total votes held by the CEO and by all other officers and directors combined. I also collect the composition of the board of directors and follow previous studies such as Weisbach (1988) and Byrd and Hickman (1992) by classifying directors as insiders, outsiders, or "greys." Insiders are directors who are also officers of the firm. Grey directors are those nonemployee directors who may have a close business relationship with the firm, such as lawyers and consultants.²² All other directors are classified as outsiders. Quarterly institutional ownership data is from CDA Spectrum.

²⁰ *Wall Street Journal* announcements were unavailable for 15 turnovers. For 3 of these observations, the *New York Times* announcement provided the announcement date and turnover classification. For 12 of these observations, local newspapers provided these variables. Of these 15 turnovers, 12 were classified as voluntary and 3 as forced.

²¹ I use proxy statements for every third year (or as close to that interval as possible), plus the proxy for the first year preceding a change in CEO and the earliest available sample-year proxy statement for each firm. For years between collection dates, intermediate values are interpolated, and for the year preceding the first collection date, the variables are extrapolated. Sampling every three years reduces the amount of data collection required at the cost of precisely recording the date of each change. To mitigate these costs, sampling prior to turnovers captures changes occurring immediately before the turnover date, when changes are more likely.

²² Grey directors are those nonemployee directors who are related to an officer of the firm, or are a former officer, consultant, commercial or investment banker, lawyer, or an executive of an insurance companies.

4.2. Potential Determinants of the Likelihood of CEO Turnover

The most obvious explanatory variables for the forced departure covariate vector are measures of stock market and accounting performance. Several studies have shown a negative relationship between firm performance and the likelihood of CEO turnover, including Benston (1985), Coughlan and Schmidt (1985), Warner, Watts and Wruck (1988), Gilson (1989), Morck, Shleifer and Vishny (1989), Jensen and Murphy (1990) and Murphy and Zimmerman (1993). As a stock-market based performance measure, I calculate the abnormal return at month t , $AbnormalReturn_t$, as the difference between the holding period return over the previous 12 months and the median return over the same period for all CRSP firms in the same 2-digit SIC code. For CEOs in office less than 12 months, I use the return over the CEO's tenure, excluding the month of turnover. This variable is observed every three months, plus the month prior to the turnover announcement, but does not include the announcement month. A lagged value, $AbnormalReturn_{t-12}$, is also included to capture the possibility that boards consider more than just the previous year's performance, an effect documented by Kim (1996), among others. If the CEO has been in place less than a year, I assume that there was no lagged *abnormal* return attributable to that CEO, i.e., it is set to zero.

The accounting performance measure used, $\Delta EBIT_t$, is the change in earnings before interest and taxes (EBIT), scaled by the prior year's total assets, less the industry median of that variable. Changes in this annual variable occur at the first quarterly-return observation that is at least four months after the firm's fiscal year end, to allow for the board to have access to that information. $\Delta EBIT_t$ is set to zero from the beginning of each CEO's tenure until the first change. The lagged value, $\Delta EBIT_{t-12}$ is also included, and is set to zero until two changes have occurred.

Given some level of performance, there are several potential influences on the probability that a CEO will be forced to leave the firm. According to the theoretical model, these factors determine the firm's impediments to firing and consequently drive cross-sectional differences in the observed pattern of turnover. To proxy for these cross-sectional factors, I use measures of CEO voting power, officer and director voting power, the composition of the board of directors, the homogeneity of the firm's industry, the level of institutional ownership, the CEO's age, the year of each observation and interactions between these factors and firm performance.

High levels of CEO stock ownership may reduce the likelihood of turnover. As Denis, Denis, and Sarin (1997) argue, this effect may be due to the correlation between CEO stock ownership and the level of his or her power, or through a reduced effectiveness of the market for corporate control due to the impediment to a successful takeover. The empirical evidence regarding the effects of CEO ownership is

mixed. Mikkelson and Partch (1997) and Mehran and Yermack (1996) find a negative relationship between the likelihood of CEO turnover and the stock ownership of officers and directors, and the CEO, respectively, but neither study proxies for the cause of turnover.²³ Denis, Denis, and Sarin (1997) document that the sensitivity of the likelihood of CEO turnover to performance is decreasing in the ownership level of officers and directors. The findings in Mikkelson and Partch (1997) and Denis, Denis, and Sarin (1997) are for all officers and directors; neither study separates the effects of the CEO's ownership from that of the other officers and directors. In contrast to those three studies, Weisbach (1988) does not find a relationship between managerial stock ownership and the likelihood of turnover.

To allow for the possible influence of managerial stock ownership, I include two variables in the turnover model, *CEOVoting_t* and *OfficerDirectorVoting_t*, the percentages of total votes controlled at time *t* by the CEO and all other officers and directors, respectively. These percentages are collected from the firm's proxy statements and are of course highly correlated with the percentage of the firm's common shares held. The measure includes all shares of beneficial interest (e.g., options exercisable within 60 days), per proxy disclosure rules.

The composition of the board of directors may also influence their hesitancy to replace the top executive. Weisbach (1988) shows that outsider-dominated boards are more likely to replace CEOs of firms that are performing poorly. In contrast, Denis, Denis, and Sarin (1997) and Mikkelson and Partch (1997) do not find a relationship between board composition and the likelihood of turnover.²⁴ In estimating the turnover model, I include the percentage of directors from outside the firm, *PercentOutside_t*.

Two aspects of the firm's industry may influence the likelihood of turnover. First, industries that have poor performance are more likely to exhibit increased voluntary or involuntary turnover because CEOs are more willing to leave, or are viewed as unable to insulate their firms from industry-wide effects. To allow for this, I include the median industry return over the past 12 months at time *t*, *IndustryReturn_t*, where industries are defined by two-digit SIC codes.²⁵ The lag of this variable, *IndustryReturn_{t-12}*, is also included, but is set to zero for CEOs in place less than one year.

Parrino (1997) discusses another possible effect of the firm's industry, arguing that industry homogeneity increases turnover frequency for two reasons. First, monitors can more easily identify poor

²³ Mehran and Yermack (1996) do proxy for voluntary and involuntary turnover in testing their main hypotheses. However, they only report the effect of CEO ownership for the entire sample of turnovers.

²⁴ Denis, Denis, and Sarin (1997) do find increased likelihood of turnover at the 0.10 significance level for outsider-dominated boards if they use Weisbach's (1988) cutoff levels for the dummy variables and employ one-tailed tests.

²⁵ As in the construction of the abnormal return variable, the industry returns only include the return since the month after the CEO entered office.

performance in a homogeneous industry. If firms in a given industry are relatively similar, then a more accurate yardstick for measuring managerial performance is available. Second, potential replacements are more abundant, lowering the search or replacement costs of turnover. Parrino provides empirical evidence consistent with his hypothesis; he finds that both voluntary and involuntary turnovers are more likely in industries consisting of similar firms than in heterogeneous industries.

To proxy for the level of homogeneity in an industry, I use Parrino's (1997) mean partial correlation proxy. The proxy is constructed by regressing monthly returns for each firm in an industry on an equally weighted market return index and an equally weighted industry return index (where industry is defined by two-digit SIC code). The slope coefficient for the industry return variable is averaged across all firms in the industry. This mean slope, *IndustryHomogeneity_t*, proxies for the similarity among firms in an industry by measuring the similarity of observed returns and is increasing in the level of homogeneity.

The presence of a strong outside monitor could strengthen the sensitivity of forced turnover to performance. Denis, Denis, and Sarin (1997) find that the presence of an outside blockholder tightens the relationship between performance and turnover, but find no such effect for the level of institutional ownership. There are at least two reasons why institutional ownership could play a stronger role in my sample. First, there has been an increased role of institutional owners since the late 1980s and Denis et al.'s sample ends in 1988 (see, for example, Gillan and Starks (1997)). Second, Denis et al. are unable to separate the ownership of institutions by type and it may be that certain types of institutions are more effective monitors.

To allow for such an effect, I collect the percentage of each firm's common shares held by institutions each quarter from the CDA Spectrum database. The data include the total held by all institutions in the database, and the percentage held by each type of institution.²⁶ The five types identified separately by CDA Spectrum are: type 1, banks; type 2, insurance companies; type 3, investment companies and their managers; type 4, independent investment advisors; and type 5, all others, including state and private pension funds, and university endowment funds.

I include four additional control variables in the model. First, given the results of Mikkelsen and Partch (1997) and Huson, Parrino, and Starks (1997) that turnover likelihood changes over time, I include

²⁶ The institutional ownership data has a strong time trend, evidenced by significant coefficients for all institution types from regressions of ownership on the observation year. In order to avoid spurious results arising from this trend, the data is de-trended by subtracting the product of the observation year and the estimated coefficient for that institution type's regression, i.e., $InstitutionalOwnershipPercentage_{year, type=j} - \hat{\beta}_j Year$ where the slope coefficient is from the regression $InstitutionalOwnershipPercentage_{year, type=j} = \alpha_j + \beta_j Year + \varepsilon$.

the year of the observation since 1980, $Year_t$. Second, the CEO's age is included, as it should be a strong factor in the likelihood of voluntary turnover. If involuntary departure has an underlying hazard rate that varies monotonically over time, age may also have a significant effect on involuntary turnover because of its correlation with a CEO's tenure (for the Cox specification). For the voluntary turnover models, I also include a dummy variable that equals one if a CEO is age 64 or greater. This is designed to account for the large number of retirements around age 65.

Finally, to control for firm size, I include the natural logarithm of annual net sales, $\ln(Sales)_t$. As Parrino (1997) argues, large firms have greater managerial depth, making it comparatively easier for the board of a large firm to identify a potential successor.

In constructing the final sample, I require that these data are available at some point during the CEO's tenure. I exclude from the sample any CEO who is a member of the firm's founding family. As these are the most likely cases of managerial entrenchment, they may unduly bias the hazard model. These restrictions reduce the initial sample to a final sample of 346 CEOs from 204 firms. Of these CEOs, 25 left involuntarily, 121 left voluntarily, and 200 were right censored.²⁷ Figure 1 shows the distributions of turnovers by type and year, including the number of CEOs whose observed tenures are right-censored in each year. Most of the observed turnovers are in the later part of the sample, which is partly a function of the sample design. The percentage of total turnovers that are forced is 17.1%, higher than Parrino's (1997) 13%, but approximately equal to Mehran and Yermack's (1996) 17.2%.

Table 1 gives summary statistics for the final CEO sample, for both the entire group of CEOs and by turnover type. In constructing the table, means are first taken within each CEO and then across CEOs. This procedure controls for the lack of independence across quarterly observations within each CEO. As the table shows, the sample firms performed very well compared to their industry counterparts (the average of $AbnormalReturn$ is 6.8% for the entire sample). This may be due in part to the sample selection procedure; firms in the 1992 *Forbes* survey must be large firms and thus the survey will likely include a disproportionate number of strong performers from the 1980s. Selecting strong performers may bias downward the probability of forced departure or the underlying hazard rate. What is of primary interest here, though, is the cross-sectional variation in the likelihood of turnover. Thus, I am most interested in estimating the relative likelihood of CEO j being fired compared to CEO k , in order to subsequently compare their pay-for-performance sensitivity. These relative probabilities are functions of the covariates and should not be biased by sampling relatively successful firms.

²⁷ Most of the CEOs in the initial sample of 175 turnovers that are not in the final sample of 146 turnovers are either (1) founding family members, or (2) CEOs of firms from industries which do not have enough members to allow calculation of the industry homogeneity proxy.

Average CEO stock ownership (proxied by percentage of votes held) is smaller than Mehran and Yermack's sample (mean of 1.2% versus 2.93%), but the medians are identical (both are 0.2%). Total managerial ownership is less than the sample of Denis, Denis, and Sarin (1997), whose managers own an average of 10.8% of their firms' stock, but the difference is probably due to the much larger firms (on average) in my sample. Further differences between my sample and those of Mehran and Yermack and Denis et al. could be due to my use of the percentage of votes held by management, versus their use of the percentage of common stock.

The percentage of outsiders on the board appears similar to that found by Weisbach (1988) and Denis et al. Institutional ownership is much higher in this sample than in the Denis et al. sample, whose mean total ownership is 33.3%, compared to my *de-trended* mean of 35.2% (the trend was positive, biasing the number presented here downward).

Table 1 also shows that during the tenure of a CEO who eventually leaves involuntarily, sample firms perform comparatively poorly and are typically in industries characterized by lower returns. As one would expect, CEOs who leave involuntarily are usually younger and leave with less tenure in office than CEOs who leave voluntarily. CEOs who leave involuntarily also are from firms with higher levels of institutional ownership compared to those in place at the end of the sample, or those who leave voluntarily.

4.3. Estimation Results

Table 2 presents the estimated hazard models of CEO turnover. Columns one and two present models for involuntary turnover, while columns three and four are voluntary turnover results. For many of the variables, it is not clear whether they will affect the likelihood of turnover directly, or through a difference in the sensitivity of turnover to performance. To allow both possibilities, *IndustryHomogeneity*, *CEOVoting*, *OfficerDirectorVoting*, and *PercentOutside*, *Year*, and the five institutional ownership variables are included alone and as interaction terms with one of the performance variables (i.e., multiplied by either *AbnormalReturn* or $\Delta EBIT$).²⁸ To lessen the impact of performance outliers, each observation of *AbnormalReturn* and $\Delta EBIT$ is replaced by its respective decile before being multiplied by the other covariates, where 10 is the strongest performing decile and one is the weakest performing. For all models, standard errors are adjusted to incorporate the fact that multiple error terms

²⁸ To select which performance variables to use in the interaction terms, I estimate models with only the performance variables, the covariate in question (e.g., *IndustryHomogeneity*), and interactions with both stock and accounting performance, as well as the lags of these variables. In all cases, the lags were insignificant. The stronger of the two interaction terms (between the stock and accounting decile performance) were used in the models of Tables 2 and 3. Due to a lack of data, I did not estimate a model including all possible interactions for all variables.

can be attributed to each CEO. These adjusted standard errors are calculated using the robust estimator of Lin and Wei (1989).

Column one shows that the probability of involuntary turnover is decreasing in the level of lagged accounting performance. Consistent with Parrino (1997), industry homogeneity (proxied by the mean correlation proxy) affects the likelihood of involuntary turnover in two ways. First, the probability of turnover in homogeneous industries is higher. Second, the sensitivity of turnover to performance is increasing in the level of homogeneity.

Interestingly, voting power affects the likelihood of forced departure but board composition does not. This is consistent with the findings of Denis, Denis, and Sarin (1997) and Mikkelsen and Partch (1997), but is in contrast to those of Weisbach (1988). Furthermore, the voting power that is significant is that of the officers and directors excluding the CEO; the voting power of the CEO has no significant effect. These results are robust to alternative specifications of the board composition variable and its interaction with performance (results not reported). For example, board composition is not significant if dummy variables for mixed and outside boards are constructed using Weisbach's (1988) cutoff levels of 40% and 60%.

The sign of the coefficient on the interaction between *OfficerDirectorVoting* and the decile of abnormal return is positive. This implies that the sensitivity of the probability of forced turnover to performance is decreasing in the voting power of officers and directors (other than the CEO). Unfortunately, the share ownership (voting) data is not separated by outside directors, inside directors, and non-director officers. However, given the negative correlation between *OfficerDirectorVoting* and the percentage of outsiders on the board, a high level of *OfficerDirectorVoting* appears consistent with strong managerial power (again, excluding the CEO) as opposed to the existence of many outside directors with significant share ownership.

Firms with high levels of ownership by type 4 institutions -- independent investment advisors -- have higher probabilities of forced turnover. This could be a causal effect (i.e., institutions actively encourage boards to replace management), or institutions could be investing in firms with active boards that are more willing to replace CEOs.

The probability of forced turnover is increasing over the sample period, as is the sensitivity of the probability of turnover to performance. Given the results in Figure 1 and Table 1, these results are not surprising. The remaining institutional holdings, the age of the CEO, and the size of the firm have no significant effect.

Column three presents the results of the voluntary turnover model using the same covariates. The most significant variables in this specification are the age of the CEO and the *Age64Dummy* variable, both of which should obviously be correlated with the probability of leaving the firm voluntarily. CEOs of firms with high levels of ownership by investment companies and their managers (type 3) are less likely to leave voluntarily, especially when the firm is performing poorly. The significance of these terms is weak (the 0.10 level), and there is no obvious interpretation for the interaction term, other than random chance.

The primary purpose of empirically modeling turnover is prediction, for use in the test of the theoretical model's implications. Accordingly, more parsimonious versions are presented in columns two and four of Table 2. Likelihood ratio tests of the models of columns two and four versus those in one and three, respectively, fail to reject the null that all of the omitted coefficients are zero. The exact forms of these parsimonious versions are reached by starting with the full models, then omitting the least significant coefficients, one at a time, until the remaining coefficients are significant at the 0.10 level.²⁹ A comparison of columns one and two of Table 2 shows that the changes in significance are for the *IndustryHomogeneity* interaction term, which becomes significant at the 0.05 level, and the *Type4InstitutionalOwnership* variable, which is no longer significant using a two-tailed test.

To aid in the interpretation of the estimated models, Table 3 presents the same results for the parsimonious models, where coefficients are replaced by hazard ratios, which are analogous to a partial derivative of the hazard function. To calculate these hazard ratios, I construct the portion of the hazard rate attributable to the covariates of a "base case" CEO, whose independent variables equal their respective means (except for the year which is set to 10, i.e., 1990, and the *Age64Dummy* which is set equal to zero).³⁰ Then, for a given covariate, the mean is replaced by the mean less one standard deviation, and the new hazard rate is calculated.³¹ (For the *Year* term, the modified year is set to 5 (i.e., 1985 in place of 1990), and *Age64Dummy* is set equal to one.) Changes in performance variables may have a secondary effect due to the interaction terms.

The hazard ratio for a given covariate is the ratio of the modified hazard rate to that of the base case, and can be interpreted as the relative likelihood of turnover. In other words, a hazard ratio of two for *ln(Sales)* would imply that a CEO whose firm size is one standard deviation below the mean is twice as likely to depart as an otherwise identical CEO whose firm size is equal to the mean. Hazard ratios below one imply a lower chance of turnover than the base case, while ratios greater than one imply a

²⁹ This is a one-tailed test for variables about which the predicted effect is clear (e.g., performance).

³⁰ The mean abnormal return is in decile 6, and the mean change in EBIT is in decile 7.

larger probability of turnover. Accordingly, Table 3 also presents p-values for Wald tests of the null hypothesis that the hazard ratios respectively equal one.

Column one of Table 3 shows a strong impact of accounting performance on forced turnover, with a hazard ratio greater than 5. Abnormal returns also have a strong effect, with a hazard ratio greater than 3. The economic impact of industry homogeneity and the voting power of officers and directors are understated by the hazard ratios. A big portion of these variables' impact occurs through interaction with performance, much like a cross-partial derivative, while the hazard ratios in Table 3 are analogous to first derivatives. Column two shows the predictably strong impact of age on the probability of voluntary departure. A 51-year-old CEO is about 30% as likely to voluntarily depart as an otherwise identical CEO of the mean age, 56. Further, the jump to age 64 increases the probability of voluntary departure by more than two and one-half times.

4.4. Estimation Results: Parametric Models

A potential problem in using the semiparametric Cox results as a proxy for the probabilities of turnover is the impact of tenure on the underlying hazard rate. If the underlying conditional probabilities of departure -- which the Cox proportional hazard models do not estimate -- are a function of CEO tenure, then comparing estimated hazard rates across CEOs at different points in their careers could lead to erroneous conclusions. Put simply, two CEOs with identical covariates could have different probabilities of retiring or being fired due to differences in their tenure. In order to check for this possibility, I re-estimate the hazard models of Table 3 using the Weibull specification.

While these results are not presented here for the sake of brevity, use of the Weibull specification does not qualitatively change the results. For the forced turnover model, the only change in significance is that the *IndustryHomogeneity* interaction variable is significant in the full Weibull model at the 0.05 level. For the voluntary turnover models, the interaction between board composition and performance is no longer significant in the parsimonious version, but CEO share ownership and industry return are in the full model. This implies that CEOs are more likely to retire when their industry is doing well, and less likely to retire when they control a large portion of the firm's voting rights, consistent with Mehran and Yermack (1996).

Perhaps the most important result of the Weibull estimation is that given the specification used, tenure does not have a significant relationship with the probability of either type of turnover. This is implied by the insignificant difference between the shape parameter, p , and one (equivalently, the natural

³¹ For the voting variables, the standard deviation is greater than the mean, and the modified covariate is set to zero.

log of p is not significantly different from zero). This implies that the use of the Cox proportional hazard models to compare CEO's probabilities of turnover should not induce significant bias.³²

For the next stage of the analysis conducted in section five, the parsimonious models of the Cox and Weibull specifications are used to produce estimated probabilities of firing and retirement.³³

5. The Sensitivity of Pay to Performance

The theoretical model predicts that the sensitivity of changes in a manager's pay to performance is a function of the probabilities of each type of turnover. The empirical turnover models of the previous section provide estimates of these probabilities and Standard and Poor's ExecuComp database provides information on managerial compensation from which measures of changes in managerial pay can be constructed. The estimated probabilities are then compared to the changes in managerial pay in order to test the model's implications.

5.1. Sample Data

The ExecuComp database contains compensation data collected from proxy statements over the 1991 to 1995 period. Due to the change in proxy disclosure requirements, option grant data is only available for all firms after 1992. It is also available for some firms in 1992. Three measures of the change in CEO pay from year $t-1$ to year t are calculated: the change in CEO salary over the year, the change in salary plus the CEO's bonus, and the change in the CEO's total direct compensation. Total direct compensation is defined as the sum of salary, bonus, other annual pay (e.g., tax gross-ups), the value of restricted stock granted, the value of stock options granted (using a modified Black-Scholes formula), long-term incentive plan payouts, and all other compensation. These measures exclude any changes in the value of stock and options held by the CEO, and therefore understate the true sensitivity of the CEO's wealth to changes in shareholder wealth. The benefit of using these variables is that they are the compensation components over which the board has direct control in designing the CEO's contract (i.e., they cannot directly control the level of stock ownership).

One standard deviation below the mean is in decile 1 of both the abnormal return and change in EBIT variables.

³² Two more alternative empirical turnover models, an exponential hazard model and a pooled logit model, yield qualitatively similar results. Coefficients from the logit model, which is used widely in the turnover literature, are generally estimated to be more significant than their hazard model counterparts.

³³ All of the section five tests, with the exception of those in Table 7, were also conducted using the "full" turnover models in columns (1) and (3) of Table 2 and the corresponding Weibull models. The results (not presented) were qualitatively very similar. The only real exceptions were the effects of forced departure on the sensitivity of salary plus bonus (the first two columns of Table 5 below), which were statistically insignificant using the full model's fitted hazard rates. In order to avoid mechanical results due to the inclusion of CEO shareholdings and firm size, the regressions in Table 7 were only estimated for the parsimonious versions of the hazard models.

A fourth measure, the change in the CEO's total firm-related wealth, is also used in the analysis. For a given year t , this variable is defined as the total direct compensation for year t , plus the dollar change in the value of exercisable options held during the year, plus the dollar change in the value of the CEO's shareholdings during the year, plus the dollar value realized from the exercise of stock options during the year. Data to calculate true option deltas (the change in option price with respect to the change in the underlying stock price) are not available. Instead, I assume an option delta of 0.60, following Jensen and Murphy (1990) and Hubbard and Palia (1995). Given this assumption, the change in the value of stock options for year t is calculated as the exercisable options held at time t times the dollar change in stock price during the year times 0.6.

Blackwell and Farrell (1997) examine CEO pay around turnover events, and find systematic differences in the pay packages of departing and incoming CEOs around the turnover date. Here, in constructing the sample of compensation data, firm-years during which a CEO turnover occurred are excluded. For the change in pay variables, I also require that a CEO have observations in consecutive years in order to be included. Two hundred and sixty-four CEOs from 239 firms have at least one observation in the final compensation sample.

Panel A of Table 4 presents summary statistics for the compensation data. Consistent with the stock market over the period and with the turnover sample, the firms experienced large average returns over the sample period (mean = 20.8%, median = 14.1%). The average bonus of \$629 thousand was slightly less than the average salary of \$677.39 thousand, and about two-thirds of the average option grant value, \$984.94 thousand. These three components make up most of the average total direct compensation, which has a mean of \$2.9 million. CEOs own an average of \$108 million in company stock, but this is highly skewed, as the median is only \$6.1 million. A large part of this skew is due to Warren Buffet of Berkshire Hathaway, who during one sample year owned \$16.6 billion in Berkshire Hathaway stock. The average gain on CEO share holdings is similarly skewed, with a mean of \$29 million and a median of \$303 thousand. The mean gain is also affected by Buffet's \$6 billion gain in one fiscal year. Due to this influence, Buffet is excluded from the tests that use the change in total wealth as the dependent variable. Consistent with Jensen and Murphy (1990), much of the change in CEO wealth appears to be coming from stock holdings. The median change in total firm-related wealth is \$3.4 million, compared to the median total direct compensation of \$2.1 million.

To test for the predicted relationships between the likelihood of turnover and pay-for-performance sensitivity, the different measures of changes in CEO wealth are used in the following regression:

$$\Delta \ln(CEOPay_t) = \alpha_0 + \alpha_1 Year94 Dummy + \alpha_2 Year95 Dummy + Return_t \left[\begin{array}{l} \gamma_1 + \gamma_2 \Pr(\text{fired}_{t-1} | Performance_t = \bar{p}) + \\ \gamma_3 \Pr(\text{retired}_{t-1} | Performance_t = \bar{p}) + \\ \gamma_4 \Pr(\text{fired}_{t-1} | Performance_t = \bar{p}) \cdot \Pr(\text{retired}_{t-1} | Performance_t = \bar{p}) \end{array} \right] \quad (14)$$

where $\Pr(\text{fired}_{t-1} | Return_t = \bar{R})$ is the estimated hazard rate for forced departure at time t , given the value of the covariates at time $t-1$, and assuming a time t performance level of \bar{p} . This can be thought of as arising from a thought experiment by the CEO where she asks herself, given the current state, how likely it is that she will be terminated next year if some level of performance is realized (e.g., an abnormal return of -10%). If the CEO thinks it will be a small probability (i.e., it would be difficult to replace her), then she would exert less effort on the margin. The compensation committee of the board of directors could also make this prediction, and the model predicts that it will include more pay-for-performance in the CEO's compensation package. This prediction is equivalent to a negative coefficient (γ_2) on the interaction between returns and $\Pr(\text{fired}_{t-1} | Return_t = \bar{R})$, implying that sensitivity is decreasing in the probability of termination. The probability of retirement, $\Pr(\text{retired}_{t-1} | Performance_t = \bar{p})$ is defined analogously, and the model predicts that the coefficient on its interaction with return, γ_3 , is positive. Finally, the model predicts that the relationship between the probability of forced turnover and the sensitivity of pay to performance is weakened as retirement becomes more likely, i.e., $\gamma_4 > 0$.

This method of estimating the probability of turnover does not condition on the previous performance of the CEO, except through his or her tenure. Thus, the relationship tested for here differs from a more direct reputation effect, like that of Milbourn (1997), who argues that CEOs with better reputations should have greater pay-for-performance sensitivity. This reputation story's intuition is based on more efficient risk sharing for CEOs with established reputations, combined with the option value of the termination decision. Milbourn's results are consistent with the intuition here if one believes that CEOs with strong prior performance are less likely to be terminated and therefore require more incentive compensation.

Dummy variables for each fiscal year are also included in the regressions to allow for economy-wide influences on changes in compensation.³⁴ Changes in the natural logs are used to control for the effect of firm size on the sensitivity of pay to performance (see Gibbons and Murphy (1992)). Firm returns are continuously-compounded; they equal the natural log of one plus the annual holding-period return.

Panel B of Table 4 summarizes these fitted hazard rates for both the Cox and Weibull models for both types of turnover. In calculating the fitted hazard rates, $\Pr(\text{fired}_{t-1} | \text{Return}_t = \bar{R})$ and $\Pr(\text{retired}_{t-1} | \text{Performance}_t = \bar{p})$, *IndustryReturn* and $\Delta EBIT$ are assumed to be two standard deviations below their respective means, and *AbnormalReturn* and $\Delta EBIT$ are in the lowest-performing decile (one). It is obvious that the Cox models do not produce true probabilities due to the unestimated underlying hazard rate. The parametric (Weibull) fitted values are easier to interpret. Through these, one can see the wide spread in the estimated hazard rates from a minimum of almost zero, to a maximum of about 14% for the probability of voluntarily leaving in the parametric model.³⁵

5.2. Empirical Results: Changes in Salary and Salary Plus Bonus

When change in salary is the measure of changes in pay in equation (14), the regression results (not presented) show that there is no significant relationship between the likelihood of turnover and the sensitivity of salary to performance. In fact, there is not a statistical relationship between changes in salary and performance. One cannot reject the joint hypothesis that all of the slope coefficients are zero, and the R-squareds are very low. This is not too surprising, given the infrequency of downward revisions in salary.

The first two columns of Table 5 show the results of regressions of changes in salary plus bonus on performance and the interaction terms. As the top row shows, changes in salary plus bonus are positively related to stock market returns. The measures of the probability of forced departure have a significant relationship with pay-for-performance sensitivity at the 0.01 level. (Significance for the returns and interaction terms, about which there are clear predictions, is calculated using one-tailed tests.) The probability of voluntary turnover has no significant impact on pay-for-performance sensitivity, which is inconsistent with the model's second prediction. Like the probability of forced turnover, the product of the two probabilities is significant at the 0.05 level.

One potential problem with these regressions is the use of fitted values as regressors. These random variables could have an impact on the true standard errors of the regression. In order to control for this possibility, Table 5 also shows p-values from a 1,000-repetition bootstrap procedure (in braces). As the table shows, the forced departure results are not robust to these standard errors, but the product of the probabilities remains significant. Thus, while revisions in salary plus bonus appear sensitive to

³⁴ For the change in CEO-pay variables available prior to 1994 (i.e., change in salary and change in salary plus bonus), a 1993 dummy is also included.

³⁵ As an alternative measure of the likelihood of turnover, for the Weibull models I conducted the tests using the probability of turnover during the following twelve months (rather than the instantaneous hazard rate). The results are qualitatively very similar and are thus not presented.

performance, the level of sensitivity appears only slightly related to the likelihood of turnover. Only the cross-product effect is robust to empirically estimated standard errors.

5.3. Empirical Results: Change in Total Direct Compensation

The pay-for-performance regressions using changes in total direct compensation are presented in the third and fourth columns of Table 5. Here again, revisions in compensation are positively related to performance at the 0.01 level. The evidence is more strongly consistent with the first prediction of the model; the pay-for-performance sensitivity is decreasing in the probability of being fired. These columns also show p-values from the bootstrap procedure (in braces). While this diminishes the significance somewhat, the results are still consistent with the firing hypothesis; the coefficient on the interaction with the probability of forced departure is significant, with p-values of 0.058 and 0.015.

Furthermore, this effect is economically significant. Table 6 shows estimated pay-for-performance sensitivity for three levels of the likelihood of turnover: the median probability, the 25th percentile and the 75th percentile (where higher percentiles imply increased likelihood of turnover). The table also shows the changes in pay-for-performance sensitivity for a movement from the 75th percentile to the 25th percentile of either fitted hazard rate. Holding the retirement hazard rate at its median, moving from the 75th percentile of the Cox forced-departure hazard rate to the 25th percentile increases the estimated pay-for-performance sensitivity of total direct compensation by more than 57% (0.49 to 0.77). Changes in the parametric forced-departure hazard rate have an even greater impact, changing the sensitivity from 0.47 to 0.83, an increase of almost 77%.

As with the change in salary plus bonus regressions, columns three and four of Table 5 show that the probability of voluntary departure and the cross-product term do not have significant effects on the pay-for-performance sensitivity of total direct compensation. Thus, the coefficients on the remaining interaction terms do not support the model's other predictions about the likelihood of retirement, which is in contrast to the findings of Gibbons and Murphy (1992). The lack of a statistical impact is supported by the small economic significance in Table 6. Moving from the 75th percentile to the 25th percentile of fitted voluntary hazard rates decreases pay-for-performance sensitivity by at most 1.3%.

Even if the probability of retirement is well-estimated, there are at least two possible explanations for the lack of an observed relationship. First, because the CEO pay measures used do not include shares or options owned, the underestimation of pay-for-performance sensitivity is increasing in the level of a CEO's stock or option holdings. These holdings typically increase over a CEO's tenure, as does the

probability of retirement.³⁶ So, the board may not need to explicitly adjust the sensitivity of the CEO's current pay as he or she nears retirement; his or her wealth is already more sensitive to performance due to increased stock and option holdings.

Second, the *costs* of increasing the sensitivity of a manager's pay could be especially high as he or she nears retirement, a feature not incorporated into the theoretical model. For example, due to the horizon problem, CEOs near retirement may be more likely to alter their investment decisions or manipulate earnings, and they may have increased incentives to do so depending on how their pay is tied to current performance (see, for example, Dechow and Sloan (1991) and Murphy and Zimmerman (1993)).

5.4. Empirical Results: Change in Total CEO Firm-Related Wealth

In an attempt to differentiate between these hypotheses, (14) is re-estimated, where the change in total CEO firm-related wealth is the dependent variable, and returns are replaced by the dollar change in shareholder wealth. Now, instead of an elasticity, the coefficients have the Jensen and Murphy (1990) interpretation of the effect of a \$1 change in shareholder wealth on CEO wealth. Given the results that pay-for-performance sensitivity is a function of firm size, these regressions also include an interaction between changes in shareholder wealth and the natural log of sales, and the level of the natural log of sales.

The results of these regressions are presented in Table 7. As they show, the current year's change in shareholder wealth has a significant impact on CEO wealth. Also, as found by Jensen and Murphy (1990), pay-for-performance sensitivity is decreasing in firm size. The results in columns one and three support the second implication of the model; the interaction coefficients for the probability of voluntary departure are significant and have the predicted sign. This significance is increased using bootstrapped p-values. However, when the cross-product term (i.e., the probability of forced departure times the probability of voluntary departure times the dollar change in shareholder wealth) is included in columns two and four, the effect of the likelihood of voluntary turnover is no longer significant. This is quite possibly due to multicollinearity between the regressors, as the cross-product effect is the product of the two probabilities of turnover, both of which are included separately.

Table 8 presents estimated total wealth pay-for-performance sensitivity as the likelihoods of turnover vary from the median to the 25th and 75th percentiles. These sensitivities are in levels, i.e., the dollar change in CEO wealth per \$1,000 change in shareholder wealth. The economic significance

³⁶ The average annual change in shares owned in my sample is significantly greater than zero at the 0.05 level.

presented here supports the hypothesis that the sensitivity of CEO wealth to changes in shareholder wealth is increasing in the probability of retirement. Moving from the 75th percentile of the likelihood of voluntary turnover to the 25th percentile decreases pay-for-performance sensitivity by about 20% (e.g., from \$4.68 per \$1,000 to \$3.76 per \$1,000).

This effect is virtually identical, whether the Cox or Weibull models are used, and with or without the cross-product terms. This lends support to the possibility of multicollinearity weakening the statistical significance of the models in columns two and four of Table 7, as the overall economic effect is fairly constant across specifications. Given the results in Tables 5 and 6, it appears that increased stock and/or option holdings, rather than revisions in year-to-year compensation, is driving the sensitivity of total wealth to the probability of voluntary departure.

The results in Tables 7 and 8 do not support the hypotheses regarding the probability of forced departure. Decreases in the likelihood of forced departure have no statistically significant effect on the sensitivity of total CEO wealth to performance (Table 7), and the economic impact has opposite the predicted sign (Table 8). Furthermore, the cross-product terms are insignificant (Table 7).

6. Conclusions

Although there is a substantial body of literature on both managerial compensation and turnover, there has been little research on the interaction between the two variables. This study is a step in providing this missing link. I develop a model of optimal compensation whose primary implications are that the optimal pay-for-performance component is (1) increasing in the probability that the manager will retire, and (2) decreasing in the probability that the manager will be fired at a given level of poor performance. A secondary implication is that the relationship between the probability of turnover and the sensitivity of pay to performance is decreasing in the probability of retirement.

The implications are tested by first estimating these two probabilities of managerial turnover, using hazard models with a sample of firms' CEOs and their observed pattern of turnover. The probability of forced departure is a function of performance, and the sensitivity of the probability to performance is primarily a function of the level of homogeneity in an industry and the voting power of a firm's officers and directors (excluding the CEO). The likelihood of voluntary departure is a function of the CEO's age, and the interactions between board composition and institutional ownership and stock market performance.

Given these empirical models of turnover, estimated probabilities of forced and voluntary departure at time t are calculated using the independent variables at time $t-1$ and an assumed level of

(poor) firm and industry performance at time t . I then test the model's predictions by calculating the sensitivity of changes in CEO pay to shareholder returns, allowing the sensitivity to be a function of the estimated probabilities of turnover.

The results are consistent with the model's implications for the probability of forced departure. CEOs who are less likely to be fired given some level of performance have total compensation that is more sensitive to performance. This is consistent with the idea that boards of directors consider the likelihood of termination (and the incentives it affects) when setting a manager's compensation scheme. The annual compensation results show no relationship between pay-for-performance sensitivity and the probability of voluntary departure, in contrast to the model's predictions and the empirical results of Gibbons and Murphy (1992). One possible explanation for the lack of a relationship is that the definition of changes in CEO wealth used understates the true sensitivity of CEO wealth by excluding stock and option holdings, which are likely to accumulate as the CEO approaches retirement. This conjecture is supported by the relationship between the probability of retirement and the sensitivity of total CEO firm-related wealth to performance.

Thus, there does appear to be an interaction between these three incentive devices. CEOs facing a small threat of termination have changes in year-to-year compensation that are more strongly related to performance. CEOs who are likely to retire have increased sensitivity of their total wealth to performance, compensating for the diminished effect of their career concerns.

Appendix A: Derivation of the Model

A.1. Variances

The variances are not functions of the data. The conditional expectation of manager's

ability, $S_{t+1|t} = E(\hat{\alpha}_t - \alpha_t)^2 = \rho^2 S_{t|t-1} \left(1 - \frac{S_{t|t-1}}{S_{t|t-1} + \sigma_\varepsilon^2} \right) + \sigma_v^2$, evolves as follows:

$$\begin{aligned}
 S_{1|0} &= \sigma_\alpha^2 \\
 S_{2|1} &= \rho^2 \frac{\sigma_\alpha^2 \sigma_\varepsilon^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2} + \sigma_v^2 \\
 S_{3|2} &= \rho^2 S_{2|1} \left(1 - \frac{S_{2|1}}{S_{2|1} + \sigma_\varepsilon^2} \right) + \sigma_v^2 \\
 S_{4|3} &= \rho^2 S_{3|2} \left(1 - \frac{S_{3|2}}{S_{3|2} + \sigma_\varepsilon^2} \right) + \sigma_v^2.
 \end{aligned} \tag{A1}$$

The conditional variance of profit, $\sigma_{\pi_t}^2 = E(\pi_t - E(\pi_t | \pi_{t-1}))^2 = S_{t|t-1} + \sigma_\varepsilon^2$, follows

$$\begin{aligned}
 \sigma_{\pi_1}^2 &= \sigma_\alpha^2 + \sigma_\varepsilon^2 \\
 \sigma_{\pi_2}^2 &= S_{2|1} + \sigma_\varepsilon^2 \\
 \sigma_{\pi_3}^2 &= S_{3|2} + \sigma_\varepsilon^2 \\
 \sigma_{\pi_4}^2 &= S_{4|3} + \sigma_\varepsilon^2.
 \end{aligned} \tag{A2}$$

A.2. Model details

A.2.2. The Manager's problem - periods three and four

To solve (1), I start with the manager's problem at the beginning of the fourth period, where the only decision is to choose fourth-period effort to maximize remaining utility:

$$\text{Max}_{e_4} a_4 + b_4 (E(\alpha_4) + e_4) - c(e_4) - \frac{1}{2} r b_4^2 \sigma_{\pi_4}^2, \tag{A3}$$

where both the expectation and variance terms are conditional on the information available at the end of period three (e.g., π_3). Differentiating (A3) with respect to effort gives the first-order condition,

$$e_4^* = c'^{-1}(b_4), \quad (\text{A4})$$

assuming that $c'^{-1}(\cdot)$ exists and is well-defined. Working backwards in time, the period-three problem becomes

$$\text{Max}_{e_3} a_3 + b_3(E(\alpha_3) + e_3) - c(e_3) - \frac{1}{2}rb_3^2\sigma_{\pi_3}^2 + (1-R_4)(1-P_4(e, h_2))u_4(e_4^*). \quad (\text{A5})$$

Here for ease of notation, $u_4(e_4^*)$ represents the fourth-period certainty equivalent as a function of the optimal fourth-period effort, or $a_4 + b_4(E(\alpha_4) + e_4^*) - c(e_4^*) - \frac{1}{2}rb_4^2\sigma_{\pi_4}^2$. Differentiating (A5) with respect to e_3 yields the first-order condition for the third period,

$$b_3 - c'(e_3^*) - (1-R_4)\frac{\partial P_4}{\partial e_3}u_4(e_4^*) = 0. \quad (\text{A6})$$

In order to solve (A6) for e_3^* , one must know the functional form of the probability that the manager is fired, $P_4(e, h_2)$.

The profit function implies that the conditional expectation of profit is

$$E(\pi_4 | \pi_3) = E(\alpha_4 | \pi_3) + e_4^*, \quad (\text{A7})$$

where the asterisk on the effort term denotes the manager's optimal fourth-period effort level. Using the Kalman filter, it can be shown that:

$$E(\pi_4 | \pi_3) = \rho E(\alpha_3 | \pi_2) + \rho \left(\frac{S_{3|2}}{S_{3|2} + \sigma_\epsilon^2} \right) (\pi_3 - E(e_3) - E(\alpha_3 | \pi_2)) + e_4^*, \quad (\text{A8})$$

where $S_{t|t-1} \equiv \text{var}(\alpha_t - \hat{\alpha}_t) | I_{t-1}$.³⁷ Given this notation, define $\Sigma_4 \equiv \frac{S_{3|2}}{S_{3|2} + \sigma_\epsilon^2}$, $\hat{\alpha}_t \equiv E(\alpha_t | \pi_{t-1})$ and

$E(e_3) \equiv \hat{e}_3$. This simplifies the notation in (A8) to

$$E(\pi_4 | \pi_3) = \rho \hat{\alpha}_3 + \rho \Sigma_4 (\pi_3 - \hat{e}_3 - \hat{\alpha}_3) + e_4^*. \quad (\text{A9})$$

So, the third-period firing rule becomes fire if

³⁷ Hamilton (1994) is one of many references for this result.

$$[\rho\hat{\alpha}_3 + \rho\Sigma_4(\pi_3 - \hat{e}_3 - \hat{\alpha}_3) + e_4^*](1 - b_4) - a_4 + h_2 \leq 0. \quad (\text{A10})$$

Using the definition of π_3 , and rearranging implies that the manager is fired if

$$\alpha_3 - \hat{\alpha}_3 + \varepsilon_3 \leq \frac{a_4 - h_2}{\rho\Sigma_4(1 - b_4)} - \frac{e_4^*}{\rho\Sigma_4} - \frac{\hat{\alpha}_3}{\Sigma_4} + \hat{e}_3 - e_3 \quad (\text{A11})$$

The left-hand side of (A11) is a mean-zero normally-distributed random variable, with variance $S_{3|2} + \sigma_\varepsilon^2$. Normalizing by the standard deviation gives the probability of being fired as

$$P_4(e, h) = \Phi\left(\left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{-\frac{1}{2}}\left[\frac{a_4 - h_2}{\rho(1 - b_4)} - \frac{e_4^*}{\rho} - \hat{\alpha}_3 + (\hat{e}_3 - e_3)\Sigma_4\right]\right) \quad (\text{A12})$$

where $\Phi(\cdot)$ is the standard normal cumulative density function.

The comparative statics for (A12) as follows:

$$\frac{\partial P_4}{\partial e_3} = -\phi\left(\left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{-\frac{1}{2}}\left[\frac{a_4 - h_2}{\rho(1 - b_4)} - \frac{e_4^*}{\rho} - \hat{\alpha}_3 + (\hat{e}_3 - e_3)\Sigma_4\right]\right)\left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{\frac{1}{2}} < 0, \quad (\text{A13})$$

where $\phi(\cdot)$ is the standard normal probability density function, and

$$\frac{\partial P_4}{\partial h_2} = -\phi\left(\left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{-\frac{1}{2}}\left[\frac{a_4 - h_2}{\rho(1 - b_4)} - \frac{e_4^*}{\rho} - \hat{\alpha}_3 + (\hat{e}_3 - e_3)\Sigma_4\right]\right)\frac{\left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{-\frac{1}{2}}}{\rho(1 - b_4)} < 0. \quad (\text{A14})$$

Substituting (A12) into (A6) shows that the optimal third-period effort must satisfy

$$c'(e_3^*) = b_3 + (1 - R_4)\phi(z)\left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{\frac{1}{2}}u_4(e_4^*), \quad (\text{A15})$$

$$\text{where } z = \left(S_{3|2}\right)^{-\frac{1}{2}}\left(\Sigma_4\right)^{-\frac{1}{2}}\left[\frac{a_4 - h_2}{\rho(1 - b_4)} - \frac{e_4^*}{\rho} - \hat{\alpha}_3 + (\hat{e}_3 - e_3^*)\Sigma_4\right].$$

Implicitly differentiating (A15) shows that optimal effort is decreasing in the probability of retirement:

$$\frac{\partial e_3^*}{\partial R_4} = \frac{-\phi(z)(S_{3|2})^{-\frac{1}{2}}(\Sigma_4)^{\frac{1}{2}}u_4(e_4^*)}{c'(e_3^*) - (1-R_4)\phi(z)(S_{3|2})^{-1}\Sigma_4u_4(e_4^*)z} < 0. \quad (A16)$$

A.2.2. The Manager's problem - periods one and two

The manager's second-period problem is to maximize his or her certainty equivalent over periods two through four:

$$\begin{aligned} \text{Max}_{e_2} \quad & a_2 + b_2[E(\alpha_2) + e_2] - c(e_2) - \frac{1}{2}rb_2^2\sigma_{\pi_2}^2 + \\ & (1-R_3) \left\{ E[w_3^*(\hat{\alpha}_3) + (1-R_4)(1-P_4(e, h_2))w_4^*(\hat{\alpha}_3)] - \right. \\ & \left. \frac{1}{2}r\text{Var}(w_3^*(\hat{\alpha}_3) + (1-R_4)(1-P_4(e, h_2))w_4^*(\hat{\alpha}_3)) \right\}, \end{aligned} \quad (A17)$$

where $E(\cdot)$ denotes expectation at the beginning of period two, and the future reservation utility levels, w_3^* and w_4^* , are written as functions of the manager's conditional expected ability and contain a random component (because the estimated manager's ability is based in part on realized profit. Given this, the manager's problem in (A17) has the first-order condition

$$c'(e_2^*) = b_2 + (1-R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_2}. \quad (A18)$$

Working farther back in time, the manager's first-period problem becomes

$$\begin{aligned} \text{Max}_{e_1} \quad & a_1 + b_1(E(\alpha_1) + e_1) - c(e_1) - \frac{1}{2}rb_1^2\sigma_{\pi_1}^2 + (1-R_2)(1-P_2(e, h_1))u_2(e_2^*) \\ & + (1-R_2)(1-R_3) \left[E(w_{3,4}^*(\hat{\alpha}_3)) - \frac{1}{2}r\text{Var}(w_{3,4}^*(\hat{\alpha}_3)) \right], \end{aligned} \quad (A19)$$

which gives the first-order condition

$$c'(e_1^*) = b_1 + (1-R_2)\phi(y)(S_{1|0})^{-\frac{1}{2}}(\Sigma_2)^{\frac{1}{2}}u_2(e_2^*) + (1-R_2)(1-R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_1}, \quad (A20)$$

³⁸ Technically, this result requires that the expected effort is close to optimal effort, which should hold in equilibrium. This condition ensures that $z < 0$, and the other terms in (A16) are positive. For a detailed discussion of why $z < 0$, see section A.2.3.

where $y = (S_{10})^{-\frac{1}{2}} (\Sigma_2)^{-\frac{1}{2}} \left[\frac{a_2 - h_1}{\rho(1 - b_2)} - \frac{e_2^*}{\rho} - \mu_\alpha + (\hat{e}_1 - e_1^*) \Sigma_2 \right]$, and the period-two firing rule is

entirely analogous to $F_4(e, h)$ in (A12) (after the obvious changes in notation).

Using (A8), $\frac{\partial \hat{\alpha}_3}{\partial e_2}$ and $\frac{\partial \hat{\alpha}_3}{\partial e_1}$ can be signed directly:

$$\frac{\partial \hat{\alpha}_3}{\partial e_2} = \rho \left(\frac{S_{21}}{S_{21} + \sigma_\varepsilon^2} \right) > 0, \quad (\text{A21})$$

$$\frac{\partial \hat{\alpha}_3}{\partial e_1} = \rho^2 \left(1 - \frac{S_{21}}{S_{21} + \sigma_\varepsilon^2} \right) \left(\frac{S_{10}}{S_{10} + \sigma_\varepsilon^2} \right) > 0. \quad (\text{A22})$$

A.2.3. Comparative Statics: Proof that $\frac{\partial b_1^}{\partial h_1} > 0$ and $\frac{\partial b_2^*}{\partial h_2} > 0$*

First, I will prove $\frac{\partial b_3^*}{\partial h_2} > 0$. Using (6), (A12), and the definition of the standard normal probability density function,

$$\frac{\partial b_3^*}{\partial h_2} = (1 - R_4) \phi(z^*) z^* \frac{\partial z^*}{\partial h_2} (S_{3|2})^{-\frac{1}{2}} (\Sigma_4)^{\frac{1}{2}} w_4^*. \quad (\text{A23})$$

We also know that

$$\frac{\partial z^*}{\partial h_2} = -(S_{3|2})^{-\frac{1}{2}} (\Sigma_4)^{-\frac{1}{2}} \left[\frac{1}{\rho(1 - b_4^*)} \right] < 0. \quad (\text{A24})$$

So, the sign of the expression in (A23) is the opposite of the sign of z^* . Re-arranging the expression for z shows that the sign of z^* is the sign of $a_4^* - (1 - b_4^*)(\rho \hat{\alpha}_3 + k) - h_2$. The product in this expression is the expected profit in the fourth period that is retained by the firm. As long as the expected profit retained by the firm is greater than the fourth-period salary, then assuming there are impediments to firing (i.e., $h_2 \geq 0$), $z^* < 0$.

To prove $\frac{\partial b_1^*}{\partial h_1} > 0$, first note that

$$\frac{\partial b_1^*}{\partial h_1} = (1 - R_2) \phi(y^*) y^* \frac{\partial y^*}{\partial h_1} (S_{10})^{-\frac{1}{2}} (\Sigma_2)^{\frac{1}{2}} w_2^*. \quad (\text{A25})$$

A similar argument to the one above shows that as long as the expected profit to the firm net of the wage contract is positive, $y^* < 0$. Also, by the similarity of y^* and z^* , $\frac{\partial y^*}{\partial h_1} < 0$ (see (A24)). The signs of all of the other terms are unambiguously positive. Q.E.D.

A.2.4. Comparative Statics: Effect of Probability of Retirement

The model above implies that the pay-for-performance component for time t is increasing in the probability of retirement in future periods. Mathematically, this is equivalent to

$$\frac{\partial b_3^*}{\partial R_4} = \phi(z^*) (S_{32})^{-\frac{1}{2}} (\Sigma_4)^{\frac{1}{2}} w_4^* > 0, \quad (\text{A26})$$

$$\frac{\partial b_2^*}{\partial R_3} = (1 - P_3(e)) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_2} > 0, \quad (\text{A27})$$

$$\frac{\partial b_1^*}{\partial R_2} = \phi(y^*) (S_{10})^{-\frac{1}{2}} (\Sigma_2)^{\frac{1}{2}} w_2^* + (1 - R_3) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_1} > 0, \quad (\text{A28})$$

$$\frac{\partial b_1^*}{\partial R_3} = (1 - R_2) \frac{\partial E(w_{3,4}^*)}{\partial \hat{\alpha}_3} \frac{\partial \hat{\alpha}_3}{\partial e_1} > 0. \quad (\text{A29})$$

Finally, the model predicts a cross effect, i.e., as the probability of retirement increases, the relationship between the probability of termination and the pay-for-performance incentive declines:

$$\frac{\partial^2 b_3^*}{\partial h_2 \partial R_4} = -\phi(z^*) z^* \frac{\partial z^*}{\partial h_2} (S_{32})^{-\frac{1}{2}} (\Sigma_4)^{\frac{1}{2}} w_4^* < 0, \quad (\text{A30})$$

$$\frac{\partial^2 b_1^*}{\partial h_1 \partial R_2} = -\phi(y^*) y^* \frac{\partial y^*}{\partial h_1} (S_{10})^{-\frac{1}{2}} (\Sigma_2)^{\frac{1}{2}} w_2^* < 0. \quad (\text{A31})$$

References

- Acharya, S., 1992. Maximizing the market value of a firm to choose dynamic policies for managerial hiring, compensation, firing and tenuring. *International Economic Review* 33, 373-397.
- Benston, G. J., 1985. The self-serving management hypothesis: Some evidence. *Journal of Accounting and Economics* 7, 67-84.
- Blackwell, D. W., and K. A. Farrell, 1997. Changes in CEO compensation and firm performance following CEO turnover. Working paper, The University of Nebraska - Lincoln.
- Borokhovich, K. A., R. Parrino, and T. Trapani, 1996. Outside directors and CEO selection. *Journal of Financial and Quantitative Analysis* 31, 337-355.
- Breslow, N. E., 1974. Covariance analysis of censored survival data. *Biometrics* 30, 89-99.
- Byrd, J. W., and K. A. Hickman, 1992. Do outside directors monitor managers? Evidence from tender offer bids. *Journal of Financial Economics* 32, 195-221.
- Coughlan, A. T., and R. M. Schmidt, 1985. Executive compensation, management turnover and firm performance. *Journal of Accounting and Economics* 7, 43-66.
- Dechow, P. M., and R. G. Sloan, 1991. Executive incentives and the horizon problem: An empirical investigation. *Journal of Accounting and Economics* 14, 51-89.
- Denis, D. J., D. K. Denis and A. Sarin, 1997. Ownership structure and top executive turnover. *Journal of Financial Economics* 45, 193-221.
- Fama, E. F., 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88, 288-307.
- Gibbons, R., and K. J. Murphy, 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy* 100, 468-505.
- Gillan, S. and L. T. Starks, 1997. Relationship Investing and shareholder activism by institutional investors: The wealth effects of corporate governance related proposals. Working paper, The University of Texas.
- Gilson, S. C., 1989. Management turnover and financial distress. *Journal of Financial Economics* 25, 241-262.

- Grossman, S. J., and O. D. Hart, 1983. An analysis of the principal agent problem. *Econometrica* 51, 7-45.
- Hallman, G., and J. C. Hartzell, 1997. Optimal compensation contracts with pay-for-performance and termination incentives. Working paper, New York University.
- Hamilton, J. D., 1994. *Time Series Analysis* (Princeton University Press, Princeton, N.J.).
- Harris, M., and A. Raviv, 1979. Optimal incentive contracts with imperfect information. *Journal of Economic Theory* 20, 231-259.
- Holmström, B., 1979. Moral hazard and observability. *Bell Journal of Economics* 10, 74-91.
- Hubbard, R. G., and D. Palia, 1995. Executive pay and performance: Evidence from the U.S. banking industry. *Journal of Financial Economics* 39, 105-130.
- Huson, M., R. Parrino and L. T. Starks, 1997. The effectiveness of internal monitoring mechanisms: Evidence from CEO turnover between 1971 and 1994. Working paper, the University of Alberta.
- Jensen, M. C., and W. Meckling, 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3, 305-360.
- Jensen, M. C., and K. J. Murphy, 1990. Performance pay and top-management incentives. *Journal of Political Economy* 98, 225-264.
- Kalbfleisch, J. D., and R. L. Prentice, 1980. *The Statistical Analysis of Failure Time Data* (John Wiley and Sons, New York, N. Y.).
- Kim, Y., 1996. Long-term firm performance and chief executive turnover: An empirical study of the dynamics. *Journal of Law, Economics, and Organization* 12, 480-496.
- Lin, D. Y., and L. J. Wei, 1989. The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* 72, 27-36.
- Mehran, H., and D. Yermack, 1996. Stock-based compensation and top management turnover. Working paper, Northwestern University.
- Mikkelsen, W. H., and M. M. Partch, 1997. The decline of takeovers and disciplinary management turnover. *Journal of Financial Economics* 44, 205-228.

- Milbourn, T. T., 1997. The Executive Compensation Puzzle: Theory and Evidence. Working paper, London Business School.
- Morck, R., A. Shleifer, and R. W. Vishny, 1989. Alternative mechanisms for corporate control. *American Economic Review* 79, 842-852.
- Murphy, K. J., 1986. Incentives, learning, and compensation: a theoretical and empirical investigation of managerial labor contracts. *Rand Journal of Economics* 17, 59-76.
- Murphy, K. J., and J. L. Zimmerman, 1993. Financial performance surrounding CEO turnover. *Journal of Accounting and Economics* 16, 273-315.
- Parrino, R., 1997. CEO turnover and outside succession: A cross-sectional analysis. *Journal of Financial Economics* (forthcoming).
- Shapiro, C., and J. E. Stiglitz, 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review* 74, 433-444.
- Shavell, S., 1979. Risk sharing and incentives in the principal and agent relationship. *Bell Journal of Economics* 10, 55-73.
- Stiglitz, J. E., and A. Weiss, 1983. Incentive effects of terminations: applications to the credit and labor markets. *American Economic Review* 73, 912-927.
- Warner, J., R. L. Watts, and K. H. Wruck, 1988. Stock prices and top management changes. *Journal of Financial Economics* 20, 461-492.
- Weisbach, M. S., 1988. Outside directors and CEO turnover. *Journal of Financial Economics* 20, 431-460.

Figure 1: Yearly Distribution of Turnovers By Type

This figure details the turnovers for each year from a sample of 346 Chief Executive Officers of 205 firms. Involuntary departures are those for which the Wall Street Journal announces that the CEO was forced out, or those for which the departing CEO is less than 60 years old and did not leave for health reasons or another job. Right-censored CEOs are still in place at the end of the observation period.

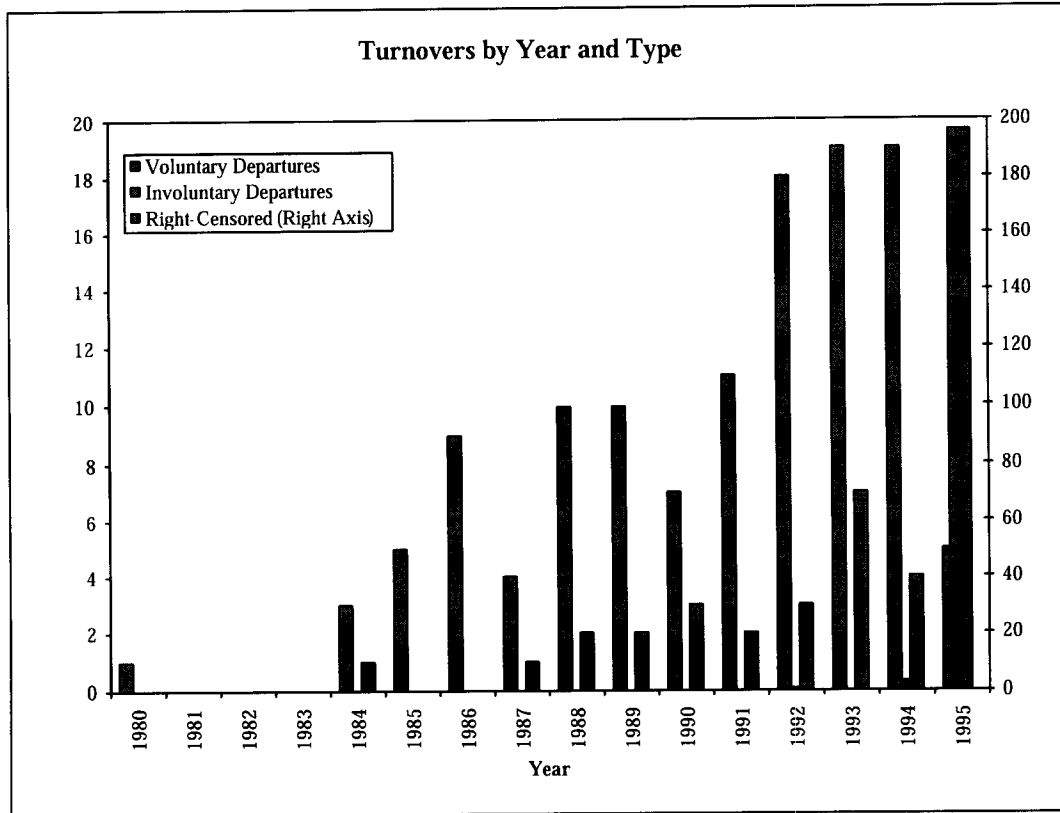


Table 1: Summary Statistics By Turnover Type Using Within-CEO Means

Summary statistics are for the means of each variable within the 346 CEOs in the sample. *AbnormalReturn* is the difference between the firm's twelve-month return and the median return for the two-digit SIC code industry, *IndustryReturn*. $\Delta Ebit$ is the change in the firm's earnings before interest and taxes, scaled by total assets, less the industry median of that variable. *CEOVoting* and *OfficerDirectorVoting* are the percentage of votes held by the CEO and all other officers and directors, respectively. *PercentOutside* is the percentage of a firm's directors that are from outside the firm (and are not classified as grey). *Ln(Sales)* is the natural logarithm of the firm's net sales. *IndustryHomogeneity* is the mean partial correlation proxy for the firm's two-digit SIC code industry. *InstitutionalOwnership* is the percentage of the firm's shares owned by institutions, net of a linear time trend over the sample period. This table also presents means and medians for the sample variables by turnover type, as well as tests of equality in means of each variable across turnover types. For the means tests, one, two, and three asterisks denote significance at the 0.10, 0.05, and 0.01 levels, respectively.

Variable	<u>Means</u>				<u>Medians</u>			
	All Observations (n=346)	Right-Censored Observations (n=200)	Involuntary Departures (n=25)	Voluntary Departures (n=121)	All Observations (n=346)	Right-Censored Observations (n=201)	Involuntary Departures (n=21)	Voluntary Departures (n=102)
<i>AbnormalReturn</i>	6.8%	7.2%	1.6%	7.3%	5.0%	5.0%	4.2%	6.2%
<i>IndustryReturn</i>	8.9%	8.5%	7.3%	9.9%	8.7%	8.3%	6.0%	9.8%
$\Delta Ebit$	0.2%	0.1%	-0.7%	0.4%	0.0%	0.0%	-0.1%	0.0%
<i>CEOVoting</i>	1.2%	1.8%	0.8%	0.4%	0.2%	0.2%	0.2%	0.1%
<i>OfficerDirectorVoting</i>	4.1%	4.3%	5.7%	3.5%	1.4%	1.5%	1.3%	1.2%
<i>PercentOutside</i>	58.2%	59.4%	55.1%	56.8%	59.7%	61.5%	57.1%	58.6%
<i>Ln(Sales)</i>	8.02	8.04	7.99	8.00	7.96	7.94	7.80	7.99
<i>IndustryHomogeneity</i>	0.31	0.30	0.33	0.31	0.29	0.29	0.32	0.29
<i>InstitutionalOwnership</i> (De-trended)	35.2%	34.7%	40.5%	35.1%	0.0%	38.0%	42.8%	36.5%
<i>CEOAge</i>	56.4	54.8	54.9	59.4	56.8	55.2	55.6	60.0
CEO Tenure at End of Spell (Months)	83.4	81.6	62.7	90.5	63.2	59.0	52.0	74.1

Variable	<u>t-Statistics for Tests of Equality in Means</u>		
	Involuntary vs. Voluntary	Involuntary vs. Right-Censored	Voluntary vs. Right-Censored
<i>AbnormalReturn</i>	-2.23 **	-2.14 **	-0.03
<i>IndustryReturn</i>	-1.18	-0.58	-0.85
$\Delta Ebit$	-1.84 *	-1.56	-0.60
<i>CEOVoting</i>	0.82	-1.27	0.96
<i>OfficerDirectorVoting</i>	1.01	0.60	0.42
<i>PercentOutside</i>	-0.54	-1.34	0.79
<i>Ln(Sales)</i>	-0.02	-0.17	0.16
<i>IndustryHomogeneity</i>	0.89	1.24	-0.34
<i>InstitutionalOwnership</i> (De-trended)	2.05 **	2.18 **	-0.10
<i>CEOAge</i>	-4.82 ***	0.06	-4.14 ***
CEO Tenure at End of Spell (Months)	-2.68 ***	-1.72 *	-0.57

Table 2: Estimated Cox Hazard Models of CEO Turnover

This table presents competing-risks Cox proportional hazard models. *AbnormalReturn* is the difference between the firm's twelve-month return and the median return for the two-digit SIC code industry, *IndustryReturn*. $\Delta Ebit$ is the change in the firm's earnings before interest and taxes, scaled by total assets, less the industry median of that variable. *AbRetDecile* and $\Delta Ebit$ Decile are the respective deciles in which the abnormal return and the change in EBIT fall, where 1 is the worst-performing decile. *IndustryHomogeneity* is the mean partial correlation proxy for the firm's two-digit SIC code industry. *CEOVoting* and *OfficerDirectorVoting* are the percentage of votes held by the CEO and all other officers and directors, respectively. *PercentOutside* is the percentage of a firm's directors that are from outside the firm (and are not classified as grey). *InstitutionalOwnershipType1* is the percentage of the firm's common shares held by banks, net of an estimated linear time trend, times 100. *InstitutionalOwnershipTypes 2 through 5* are similarly de-trended, and represent the percentage ownership of insurance companies (type 2), investment companies and their managers (type 3), independent investment advisors (type 4), and all other institutions (type 5). $Ln(Sales)$ is the natural logarithm of the firm's net sales. *Year* equals the years elapsed since 1980, and *CEOAge64Dummy* takes on the value of one if the CEO is at least 64 years old, and zero otherwise. t-statistics for the null hypothesis that the coefficient is equal to zero are in parentheses, using robust standard error estimates. One, two, and three asterisks denote statistical significance at the 0.10, 0.05 and 0.01 levels, respectively.

Variable	Forced Departures		Voluntary Departures	
	(1)	(2)	(3)	(4)
<i>AbnormalReturn_t</i>	1.061 (0.64)		-0.373 (-0.71)	
<i>AbnormalReturn_{t,12}</i>	1.085 (0.76)		-0.004 (-0.01)	
<i>IndustryReturn_t</i>	-1.861 (-1.39)		0.527 (1.06)	
<i>IndustryReturn_{t,12}</i>	-0.221 (-0.09)		0.571 (1.32)	
$\Delta Ebit_t$	-7.783 (-0.76)		3.659 (1.37)	
$\Delta Ebit_{t,12}$	-16.953 (-2.06) **	-15.710 (-2.39) **	1.172 (0.43)	
<i>IndustryHomogeneity</i>	7.364 (2.09) **	5.337 (2.15) **	-0.928 (-0.52)	
<i>IndustryHomogeneity X $\Delta Ebit$Decile</i>	-1.150 (-1.58)	-0.785 (-2.45) **	-0.110 (-0.38)	
<i>CEOVoting</i>	-0.857 (-0.15)		-3.428 (-0.52)	
<i>OfficerDirectorVoting</i>	-2.165 (0.45)		0.739 (0.30)	
<i>CEOVoting X AbRetDecile</i>	-0.847 (-0.50)		-0.328 (-0.35)	
<i>OfficerDirectorVoting X AbRetDecile</i>	0.939 (2.35) **	0.650 (3.30) ***	-0.269 (-0.66)	
<i>PercentOutside</i>	-1.373 (-0.81)		0.573 (0.56)	
<i>PercentOutside X $\Delta Ebit$Decile</i>	0.321 (0.96)		-0.120 (-0.80)	-0.085 (-1.92) *
<i>InstitutionalOwnershipType1</i>	0.011 (0.23)		-0.003 (-0.13)	
<i>InstitutionalOwnershipType2</i>	-0.156 (-0.90)		0.025 (0.35)	
<i>InstitutionalOwnershipType3</i>	-0.063 (0.69)		0.006 (0.12)	

Variable	Forced Departures		Voluntary Departures	
	(1)	(2)	(3)	(4)
<i>InstitutionalOwnershipType4</i>	0.064 (1.69) *	0.028 (1.62)	-0.044 (-1.81) *	
<i>InstitutionalOwnershipType5</i>	-0.020 (-0.17)		0.026 (0.44)	
<i>InstitutionalOwnershipType1</i>	0.0002 (0.02)		0.0011 (0.31)	
<i>X AbRet Decile</i>				
<i>InstitutionalOwnershipType2</i>	0.024 (0.98)		-0.015 (-1.23)	
<i>X AbRet Decile</i>				
<i>InstitutionalOwnershipType3</i>	0.006 (0.27)		-0.012 (-1.45)	-0.010 (-2.58) ***
<i>X AbRet Decile</i>				
<i>InstitutionalOwnershipType4</i>	-0.012 (-1.23)		0.006 (1.72) *	
<i>X AbRet Decile</i>				
<i>InstitutionalOwnershipType5</i>	0.019 (1.03)		-0.004 (-0.40)	
<i>X AbRet Decile</i>				
<i>Ln(Sales)</i>	-0.109 (-0.43)		-0.031 (-0.43)	
<i>Year</i>	0.294 (2.89) ***	0.205 (3.72) ***	-0.017 (-0.35)	
<i>Year X AbRetDecile</i>	-0.042 (-2.44) **	-0.028 (-3.39) ***	0.007 (1.02)	
<i>CEOAge</i>	0.037 (0.85)		0.233 (5.65) ***	0.232 (5.96) ***
<i>CEOAge64Dummy</i>			0.981 (3.42) ***	0.922 (3.29) ***
Log Likelihood	-99.5	-105.5	-455.7	-463.2
Number of Subjects	346	346	346	346
Number of Completed (Uncensored) Spells	25	25	121	121

Table 3: Estimated Cox Hazard Ratios for CEO Turnover

This table presents hazard ratios for the parsimonious competing-risks Cox proportional hazard models of Table 2. Columns (1) and (2) of this table correspond to the models in columns (2) and (4) of Table 2, respectively. The hazard ratio is the ratio of the hazard rate of an "adjusted" CEO to a "base case" CEO. The base case CEO has covariates equal to the mean for all CEOs except for $Year=10$ (i.e., 1990) and $CEOAge64Dummy=0$. The adjusted CEO has identical covariates, except for the variable in question which is replaced by the mean less one standard deviation ($Year=5$ and $CEOAge64Dummy=1$ for those variables). If the mean minus one standard deviation is less than the minimum observation, the minimum observation is used. Variables are as defined in Table 2. Interaction terms are not presented; their effects are included in the hazard ratios of the respective underlying variable. P-values for Wald tests of the null hypothesis that the ratio is one are in parentheses. One, two, and three asterisks denote statistical significance at the 0.10, 0.05 and 0.01 levels, respectively.

Variable	Forced Departures (1)	Voluntary Departures (2)
<i>AbnormalReturn_t</i>	3.613 (0.002) ***	
<i>ΔEbit_t</i>	5.042 (0.014) **	
<i>ΔEbit_{t,12}</i>	1.536 (0.017) **	
<i>IndustryHomogeneity</i>	1.013 (0.953)	
<i>OfficerDirectorVoting</i>	0.851 (0.001) ***	
<i>PercentOutside</i>		1.077 (0.055) *
<i>InstitutionalOwnershipType3</i>		1.218 (0.010) ***
<i>InstitutionalOwnershipType4</i>	0.778 (0.104)	
<i>Year</i>	0.840 (0.450)	
<i>CEOAge</i>		0.296 (0.000) ***
<i>CEOAge64Dummy</i>		2.515 (0.001) **

Table 4: Summary Statistics for Compensation Tests

Panel A details summary statistics for the compensation of the sample of 200 CEOs over the 1991-1995 period. The value of options granted is calculated using a modified Black-Scholes approach. Total direct compensation is the sum of salary, bonus, other annual compensation, value of restricted stock granted, value of stock options granted, long-term incentive plan payouts, and all other compensation. The change in total CEO wealth is defined as total direct compensation, plus the change in value of the CEO's shares held, plus the change in value of the CEO's exercisable options, plus the net value realized from exercise of options during the year. Due to changes in proxy disclosure rules, fewer observations are available for the value of options, CEO share holdings, total direct compensation, and total CEO firm-related wealth. The gain on CEO share holdings net of the market is the dollar gain on the CEO's share holdings, less the hypothetical dollar gain on an equivalent investment in the S&P 500 Index (i.e., the total return on the index). Dollar amounts are in thousands. Panel B presents summary statistics for the two methods of estimating the likelihood of CEO turnover, given a level of performance. These statistics only cover the period over which compensation data is available. Note that due to the Cox hazard model methodology, the numbers presented are not true hazard rates, because the underlying hazard rate is not estimated.

Panel A: Compensation Sample						
Variable	Mean	Median	Std. Dev.	Min.	Max.	Number of Obs.
<i>Annual Return</i>	20.8%	14.1%	0.38	-86.1%	303.9%	1,079
<i>Salary</i>	677.39	650.00	235.73	100.00	2,000.00	1,060
<i>Bonus</i>	628.58	408.64	868.43	0.00	8,606.22	1,060
<i>Salary + Bonus</i>	1,305.97	1,066.90	968.55	100.00	9,406.22	1,060
<i>Value of Options Granted</i>	984.94	439.70	1,663.88	0.00	14,433.44	782
<i>Long-Term Incentive Plan Payouts</i>	214.17	0.00	672.10	0.00	11,306.25	1,060
<i>Total Direct Compensation</i>	2,917.79	2,109.05	2,617.93	280.39	20,305.96	782
<i>Market Value</i>	6,537,061	2,877,950	10,895,270	35,891	103,073,300	1,081
<i>CEO Share Value</i>	108,378	6,148	805,563	0	16,569,640	852
<i>Gain on CEO Share Holdings</i>	29,139	303	304,255	-740,425	6,039,663	562
<i>Gain on CEO Share Holdings (Net of market)</i>	10,157	-38.89	161,633	-1,016,269	2,079,148	562
$\Delta(\ln(\text{Salary} + \text{Bonus}))$	0.090	0.080	0.282	-1.883	1.056	583
$\Delta(\ln(\text{Total Direct Compensation}))$	0.080	0.097	0.537	-2.297	1.964	495
$\Delta(\text{Value of Option Holdings})$	1,605	239	6,268	-31,859	70,372	736
$\Delta(\text{Total CEO Wealth})$	33,665	3,452	305,596	-736,317	6,039,987	560

Panel B: Fitted Hazard Rates						
Model	Mean	Median	Std. Dev.	Min.	Max.	Number of Obs.
<u>Cox Proportional Hazard Models</u>						
Forced Turnover	130	115	69	31	558	738
Voluntary Turnover	2,168,905	543,804	12,100,000	3,934	228,000,000	739
<u>Weibull Hazard Models</u>						
Forced Turnover	0.0167	0.0145	0.0097	0.0034	0.0809	738
Voluntary Turnover	0.0058	0.0033	0.0111	0.0003	0.1436	779

Table 5: Pay-for-Performance Sensitivity as Functions of Probability of Turnover,
Dependent Variables: $\Delta \ln(\text{Salary} + \text{Bonus})$ and $\Delta \ln(\text{Total Direct Compensation})$

This table presents regressions of changes in CEO salary plus bonus and total direct CEO compensation (as defined in Table 4) on returns and estimated probabilities of CEO departure (from the parsimonious model of Table 2 and the corresponding Weibull model). In calculating the estimated hazard rates, performance variables are set to two standard deviations below their mean and the remaining independent variables are set to their true values at time $t-1$. Standard OLS p-values are in parentheses, and p-values from bootstrapped regressions are in braces (using 1,000 repetitions). One, two, and three asterisks denote significance at the 0.10, 0.05, and 0.01 levels, respectively. For the return and interaction terms, significance is calculated using one-sided tests. All of the regressions are significant at the 0.01 level.

Variable	Dependent Variable: $\Delta \ln(\text{Salary} + \text{Bonus})$		Dependent Variable: $\Delta \ln(\text{Total Direct Compensation})$	
	Cox Proportional Hazard	Weibull Hazard	Cox Proportional Hazard	Weibull Hazard
<i>AnnualReturn_t</i>	0.5679 (0.000) *** {0.000} ***	0.6020 (0.000) *** {0.002} ***	0.9934 (0.000) *** {0.002} ***	1.1007 (0.000) *** {0.003} ***
<i>AnnualReturn_{t-1}</i>	0.0554 (0.110) {0.123}	0.0547 (0.113) {0.122}	0.1106 (0.198) {0.234}	0.1168 (0.185) {0.228}
<i>AnnualReturn_t X Pr(ForcedTurnover)</i>	-0.0015 (0.004) *** {0.174}	-13.3268 (0.003) *** {0.153}	-0.0039 (0.005) *** {0.058} *	-37.4331 (0.002) *** {0.015} **
<i>AnnualReturn_t X Pr(VoluntaryTurnover)</i>	-3.53E-08 (0.953) {0.976}	-29.1118 (0.993) {0.992}	1.85E-08 (0.421) {0.469}	-13.8052 (0.627) {0.691}
<i>AnnualReturn_t X Pr(ForcedTurn.) X Pr(VoluntaryTurnover)</i>	2.35E-10 (0.045) ** {0.028} **	1526.89 (0.014) ** {0.007} ***	-6.20E-11 (0.545) {0.496}	1170.53 (0.290) {0.208}
Dummy=1 if 1993 Fiscal Year	0.0358 (0.238) {0.252}	0.0327 (0.281) {0.272}		
Dummy=1 if 1994 Fiscal Year	0.0964 (0.003) *** {0.004} ***	0.0914 (0.005) *** {0.000} ***	0.2461 (0.000) *** {0.000} ***	0.2462 (0.000) *** {0.000} ***
Dummy=1 if 1995 Fiscal Year	-0.0314 (0.353) {0.244}	-0.0380 (0.263) {0.184}	0.0520 (0.479) {0.502}	0.0568 (0.439) {0.472}
Intercept	0.0249 (0.345)	0.0295 (0.259)	-0.0746 (0.209)	-0.0764 (0.196)
R ²	0.093	0.0943	0.063	0.071
Adjusted R ²	0.081	0.0825	0.047	0.055
Number of Observations	621	621	411	411

Table 6: Summary of Changes in Pay-for-Performance Sensitivity as Functions of Changes in Probability of Turnover, Dependent Variable: $\Delta \ln(\text{Total Compensation})$

This table presents estimated pay-for performance sensitivities for three levels of the various probabilities of turnover: the median, the 25th percentile, and the 75th percentile (where higher percentiles imply increased likelihood of turnover). The Difference and % Change columns give an indication of the economic significance of a change in the probability of forced or voluntary turnover from the 75th percentile to the 25th percentile. The columns indicated refer to the regressions in Table 5. Note that because the regressions are run in change-in-log form, the sensitivities can be interpreted as elasticities, rather than dollar sensitivities.

	<u>Vary $Pr(\text{Forced Turnover})$</u>				
	Predicted sign: (+)				
	Median Sensitivity	75th Percentile	25th Percentile	Difference	% Change
Cox Hazard Models - Column (3)	0.66	0.49	0.77	0.28	57.5%
Weibull Hazard Models - Column (4)	0.69	0.47	0.83	0.36	76.6%
	<u>Vary $Pr(\text{Voluntary Turnover})$</u>				
	Predicted sign: (-)				
	Median Sensitivity	75th Percentile	25th Percentile	Difference	% Change
Cox Hazard Models - Column (3)	0.59	0.59	0.59	0.01	0.9%
Weibull Hazard Models - Column (4)	0.69	0.69	0.68	-0.01	-1.3%

Table 7: Pay-for-Performance Sensitivity as Functions of Probability of Turnover,
Dependent Variable: $\Delta \ln(\text{Total CEO Firm-Related Wealth})$

This table presents regressions of changes in total CEO firm-related wealth (as defined in Table 4) on changes in shareholder wealth and estimated probabilities of CEO departure (from the parsimonious models above). $\Delta \text{Shareholder Wealth}$ is defined as the total return to the firm's common stock times the firm's market value at the beginning of the year. In calculating the estimated hazard rates, performance variables are set to two standard deviations below their mean and the remaining independent variables are set to their true values at time $t-1$. Standard OLS p-values are in parentheses, and p-values from bootstrapped regressions are in braces (using 1,000 repetitions). One, two, and three asterisks denote significance at the 0.10, 0.05, and 0.01 levels, respectively. For the return and interaction terms, significance is calculated using one-sided tests. All of the regressions are significant at the 0.01 level.

Variable	Cox Proportional Hazard Models		Weibull Hazard Models	
$\Delta \text{Shareholder Wealth}_t$	0.0150 (0.001) *** {0.000} ***	0.0149 (0.001) *** {0.000} ***	0.0144 (0.001) *** {0.000} ***	0.0145 (0.001) *** {0.001} ***
$\Delta \text{Shareholder Wealth}_{t-1}$	0.0001 (0.453) {0.403}	0.0001 (0.447) {0.398}	0.0001 (0.907) {0.403}	0.0001 (0.888) {0.395}
$\Delta \text{Shareholder Wealth}_t X$ <i>Pr(Forced Turnover)</i>	0.00001 (0.804) {0.921}	4.46E-06 (0.702) {0.815}	0.0458 (0.844) {0.921}	0.0217 (0.821) {0.699}
$\Delta \text{Shareholder Wealth}_t X$ <i>Pr(Voluntary Turnover)</i>	8.55E-10 (0.057) * {0.027} **	6.62E-10 (0.316) {0.279}	0.3462 (0.048) ** {0.027} **	0.2232 (0.322) {0.217}
$\Delta \text{Shareholder Wealth}_t X$ <i>Pr(Forced Turnover) X</i> <i>Pr(Voluntary Turnover)</i>		1.56E-12 (0.440) {0.450}		8.0 (0.389) {0.396}
$\Delta \text{Shareholder Wealth}_t X$ <i>Ln(Sales)</i>	-0.0015 (0.000) *** {0.000} ***	-0.0015 (0.001) *** {0.000} ***	-0.0015 (0.000) *** {0.000} ***	-0.0015 (0.001) *** {0.000} ***
Dummy=1 if 1994 Fiscal Year	-1586.1 (0.704) {0.454}	-1585.1 (0.704) {0.524}	-1682.9 (0.686) {0.454}	-1674.4 (0.688) {0.512}
Dummy=1 if 1995 Fiscal Year	4490.5 (0.300) {0.350}	4451.9 (0.306) {0.382}	4443.5 (0.305) {0.350}	4383.4 (0.313) {0.324}
<i>Ln(Sales)</i>	-1660.0 (0.359) {0.286}	-1693.4 (0.354) {0.250}	-1612.5 (0.374) {0.286}	-1667.3 (0.361) {0.260}
Intercept	18190.6 (0.217)	18451.8 (0.214)	17844.7 (0.226)	18266.6 (0.218)
R ²	0.087	0.088	0.088	0.089
Adjusted R ²	0.069	0.067	0.070	0.068
Number of Observations	407	407	407	407

Table 8: Summary of Changes in Pay-for-Performance Sensitivity as Functions of Changes in Probability of Turnover, Dependent Variable: $\Delta(\text{Total CEO Firm-Related Wealth})$

This table presents estimated pay-for performance sensitivities for three levels of the various probabilities of turnover: the median, the 25th percentile, and the 75th percentile (where higher percentiles imply increased likelihood of turnover). The Difference and % Change columns give an indication of the economic significance of a change in the probability of forced or voluntary turnover from the 75th percentile to the 25th percentile. The columns indicated refer to the regressions in Table 7. Note that because the regressions are run in change-in-wealth form, the sensitivities can be the dollar change in CEO wealth per \$1,000 change in shareholder wealth.

	<u>Vary Pr(Forced Turnover)</u>				
	Median Sensitivity	75th Percentile	25th Percentile	Difference	% Change
Predicted sign: (+)					
<u>Cox Hazard Models</u>					
No Cross-Product Term - Column (1)	4.05	4.27	3.89	(0.38)	-8.9%
Cross-Product Term - Column (2)	4.05	4.27	3.89	(0.38)	-8.8%
<u>Weibull Hazard Models</u>					
No Cross-Product Term - Column (3)	4.11	4.40	3.91	(0.49)	-11.2%
Cross-Product Term - Column (4)	4.11	4.41	3.90	(0.51)	-11.6%
	<u>Vary Pr(Voluntary Turnover)</u>				
	Median Sensitivity	75th Percentile	25th Percentile	Difference	% Change
Predicted sign: (-)					
<u>Cox Hazard Models</u>					
No Cross-Product Term - Column (1)	4.05	4.68	3.76	(0.92)	-19.6%
Cross-Product Term - Column (2)	4.05	4.68	3.77	(0.90)	-19.3%
<u>Weibull Hazard Models</u>					
No Cross-Product Term - Column (3)	4.11	4.63	3.69	(0.94)	-20.4%
Cross-Product Term - Column (4)	4.11	4.62	3.69	(0.92)	-20.0%