CEO Compensation and Private Information:
An Optimal Contracting Perspective[*]

Roman Inderst[†]   Holger M. Mueller[‡]

December 2006

Abstract

We consider the joint optimal design of CEOs' severance pay and on-the-job pay in a model in which the CEO has interim private information about the likely success of his strategy. The board faces a tradeoff between reducing the likelihood that the firm forgoes an efficient strategy change and limiting the CEO's informational rents. The optimal truthtelling mechanism takes a simple form: it consists of fixed severance pay and high-powered, non-linear on-the-job pay, such as a bonus scheme or option grant. Our model makes testable predictions linking CEOs' severance pay and on-the-job pay to each other as well as to the firm's external business environment, firm size, and corporate governance.

1   Introduction

A tenet of the management literature is that top executives are assumed to act on the firm's behalf as "champions of change" (Geletkanycz and Black, 2001)—ready and willing to adjust strategic policy as demanded by the firm's changing environment. Yet evidence suggests that top executives often act as impediments to change, showing instead a strong "commitment to the strategic status quo" (Geletkanycz, 1997; Finkelstein and Hambrick, 1990).

Why do top executives cling to the strategic status quo? One possible reason is that each executive has his own style and strategy (e.g., Bertrand and Schoar, 2003), so that a strategy change is best accomplished by hiring a replacement with a different strategy.[1] As empirical research suggests, major strategy changes tend to occur only after a new CEO has been appointed (Miller and Friesen, 1980; Tushman, Virany, and Romanelli, 1985), while CEO replacement is one of the primary means by which firms adapt to major changes in their environments (Tushman, Newman, and Romanelli, 1986; Wiersema and Bantel, 1992).[2]

A potential difficulty with implementing a strategy change is that CEOs are likely to have better information than the board as to whether a strategy change is desirable.[3] What is more, the CEO may withhold information indicating that a strategy change is desirable to safeguard his own position. In this paper, we examine the optimal design of CEO pay in the light of this dilemma. We find that the uniquely optimal CEO compensation package is a combination of severance pay and high-powered on-the-job pay, such as bonus schemes and option grants.

In our model, the CEO privately observes a signal indicating the interim success likelihood of his strategy. Optimal truthtelling mechanisms take a simple form. If the signal is above a

---

[1]Firms tend to choose successors with different career experiences, educational backgrounds, and personal characteristics than the previous CEO to achieve a better "fit" with the firm's new strategic orientation (White, Smith, and Barnett, 1997). What this suggests is that the incumbent CEO cannot easily change his style and strategy. For example, it is difficult to imagine that former Sunbeam CEO Al Dunlop—better known as "Chainsaw Al"—could easily shake off his ruthless cost-cutting image. Relatedly, Rotemberg and Saloner (2000) argue in a theoretical contribution that hiring a CEO with a particular vision is a credible signal to the firm's employees that the firm will pursue a particular strategy. If the CEO could change his strategy at will, the signal would not be credible.

[2]Hofer (1980) argues that a long-standing CEO must be replaced for a successful turnaround to occur.

[3]"The CEO most always determines the agenda and the information given to the board. This limitation severely hinders the ability of even highly talented board members to contribute effectively to the monitoring of the CEO and the company's strategy" (Jensen, 1993).

certain threshold, the strategic status quo is maintained, and the CEO receives his on-the-job pay. If the signal is below the threshold, a replacement with a different strategy is hired, and the CEO receives his severance pay. It is uniquely optimal *not* to make the CEO's on-the-job pay or his severance pay contingent on his reported signal. That is, the uniquely optimal menu consists of fixed severance pay and a single on-the-job pay scheme.

The role of severance pay is to induce truthful revelation in low-signal states.[4] The higher is the CEO's severance pay, the higher is the threshold signal above which the strategic status quo is maintained. The board faces a simple tradeoff. On the one hand, it wants to push up the threshold signal towards the first-best threshold, minimizing the range of signals at which the firm forgoes an efficient strategy change. On the other hand, it wants to minimize the CEO's informational rents. In our model, the CEO's informational rents are proportional to his severance pay.[5] Thus, pushing up the threshold signal towards the first best improves efficiency, but it also implies that the CEO extracts more informational rents.

In the face of this tradeoff, the board's problem of designing the CEO's pay can be stated in two ways. Holding the CEO's severance pay—and thus his informational rents—fixed, the optimal on-the-job pay scheme minimizes the range of signals at which the firm forgoes an efficient strategy change. Conversely, holding the threshold signal fixed, the optimal on-the-job pay scheme minimizes the amount of severance pay—and thus informational rents—needed to implement a given threshold signal. One problem is the dual of the other.

The uniquely optimal on-the-job pay scheme is a high-powered, non-linear pay scheme that shifts all of the CEO's pay into the highest cash-flow states, which implies that—depending on the constraints imposed on the model—it is either a bonus scheme or an option-like contract.[6] Intuitively, a scheme that shifts all of the CEO's on-the-job pay into the highest cash-flow states minimizes the CEO's *expected* on-the-job pay at low signals, thus minimizing the informational

---

[4]Lambert and Larcker (1985) and Harris (1990) are two early models in which severance pay (or golden parachutes) provides executives with insurance against the loss of their jobs.

[5]The CEO's on-the-job pay is (endogenously) tied to his severance pay in our model, implying that the two instruments move in the same direction. Hence, *both* the CEO's severance pay *and* his on-the-job pay are measures of his informational rents.

[6]We first consider a relatively unconstrained setting in which the optimal on-the-job pay scheme is a discontinuous bonus scheme. If we introduce a simple manipulation problem in which the CEO can falsify the firm's cash flows, the optimal on-the-job pay scheme becomes an option-like contract. Both schemes have the property that—subject to binding constraints—all of the CEO's on-the-job pay is shifted into the highest cash-flow states.

rents needed to accomplish a given threshold signal.

To the best of our knowledge, this is the first paper to consider the joint design of CEOs' on-the-job pay and severance pay in an optimal contracting framework—i.e., a setting with more than two cash flows (here: a continuum) in which contracts are not restricted to a narrow class.[7] Considering a continuum of cash flows allows us to distinguish between linear and non-linear, as well as different types of non-linear, pay schemes. Considering the joint optimal design of severance pay and on-the-job pay in a model with interim private information allows us to derive a novel argument for high-powered, non-linear CEO pay such as bonus schemes and option grants: such pay minimizes the CEO's expected on-the-job pay at low signals, thereby minimizing the CEO's informational rents. The argument differs from existing arguments based on moral hazard (e.g., Holmström, 1979; Innes, 1990) and risk taking (e.g., Smith and Stulz, 1985). In a recent paper, Dittmann and Maug (2006) estimate the standard principal-agent model for a sample of 598 CEOs and conclude that neither moral hazard nor risk taking can satisfactorily explain the use of option grants for CEOs.[8]

Consistent with empirical evidence, our model predicts that CEOs' on-the-job pay and severance pay should move in the same direction. In a study of contractual—i.e., not ex-post negotiated—severance pay agreements, Rusticus (2006) finds a significant positive relation between the amount of cash severance pay and CEOs' on-the-job pay. Likewise, Schwab and Thomas (2004) find that CEOs' on-the-job pay is positively related to their contractual severance pay, while Lefanowicz, Robinson, and Smith (2000) find that managers whose employment contracts stipulate generous golden parachutes are more highly compensated on their jobs.[9] Another prediction is that CEOs of firms in volatile business environments should earn more

---

[7]Almazan and Suarez (2003) consider the interplay between severance pay and on-the-job pay in a moral hazard model in which the firm's cash flow is either zero or $R > 0$.

[8]Dittmann and Maug (2006) conclude "we need a different contracting model to understand salient features of executive compensation contracts." Likewise, Jenter (2002) argues that moral hazard alone cannot explain the use of option grants in executive compensation, while Carpenter (2000) and Ross (2004) cast doubt on the argument that options are granted to increase risk-taking incentives.

[9]Schwab and Thomas (2004) also document the pervasiveness of contractual severance pay: more than 93 percent of CEO employment contracts in their sample formally stipulate severance pay. Our prediction that CEOs' on-the-job pay and severance pay should move in the same direction is orthogonal to the "outrage-constraint" view by Bebchuk and Fried (2004), who argue that severance pay is merely a form of "stealth compensation" acting as a substitute for more visible on-the-job pay.

informational rents—and receive both higher severance pay and higher on-the-job pay—than CEOs of firms in stable business environments. Consistent with this prediction, Rusticus (2006) finds that "firms with more stability and less frequent strategic change have less need for everance agreements." Moreover, our model predicts that the rise in strategic uncertainty since the late 1970s goes hand in hand with a rise in CEO compensation—both in terms of severance pay and on-the-job pay—which is again consistent with the evidence.[10]

There is a large literature, both empirical and theoretical, on CEO compensation. While we cannot do justice to all papers, three papers seem particularly relevant for our paper. Almazan and Suarez (2003) also have a model in which CEOs receive both severance pay and on-the-job pay. There is no private information, so the board's decision to hire a replacement is always first-best efficient. Rather, the inefficiency is that the CEO has insufficient incentives to create value, leading to a novel argument for severance pay: by creating a link between the CEO's payoff from renegotiating with the board and his ex-ante investment, severance pay is a cheaper way to provide effort incentives than is on-the-job incentive pay. Dow and Raposo (2005) also consider the relation between CEO pay and strategy change. In their paper, the problem is not that CEOs impede strategy changes, but rather that they may choose overly "dramatic" strategy changes.[11] None of the available strategy choices involves the CEO's replacement, implying that there is no need for any severance pay. Finally, Eisfeldt and Rampini (2006) also consider a model with interim private information. In their model, managers have private information about the productivity of the assets under their control. To induce a low-productivity manager to relinquish control of his assets, shareholders must pay him a bonus, which is similar to the role of severance pay in our model. A further but important difference between all three papers and ours is that we focus on the optimal functional form—in particular, the non-linearity—of on-the-job compensation schemes.

The rest of this paper is organized as follows. Section 2 presents the model and preliminary results. The main analysis is in Section 3. Section 4 derives comparative static results. Section 5 concludes. All proofs are in the Appendix.

---

[10]See Section 4 for references. The rise in strategic uncertainty is also consistent with the increase in firm-level volatility in the United States over the past 30 years (e.g., Comin and Philippon, 2005).

[11]See also Dow and Raposo (2005b) for a related model in which successful organizational changes require that the CEO obtains the support of other senior managers.

2   CEO Compensation, Private Information, and Entrenchment

2.1   The Basic Model

Information-Based Entrenchment

In $t = 0$ a CEO with a particular style and strategy is hired to manage the firm. While the CEO is the best available candidate for the job, there is uncertainty as to whether his strategy is a good match for the firm. At some interim date, $t = 1$, the CEO privately observes a noisy signal of the quality of this match.[12] On the other hand, the firm's board—when deciding whether to initiate a strategy change by hiring a replacement with a different strategy—must rely on the information it is given by the CEO. Hence, the CEO can always secure his stay in office by simply withholding unfavorable information from the board. The assumption that the CEO privately observes the signal is meant to capture his informational advantage vis-à-vis the board (see Introduction). The firm's cash flow is realized in $t = 2$.

The CEO's ability to secure his stay in office by withholding unfavorable information is a parsimonious way to model entrenchment. One could think of a richer strategy space, e.g., one in which the CEO—after privately observing his signal—can entrench himself by undertaking an irreversible investment that makes it prohibitively costly for the board to replace him, as in Shleifer and Vishny (1989) and other models of managerial entrenchment.[13] While a richer strategy space has interesting practical implications, it does not affect the way how we solve the model: by the revelation principle, we can without loss of generality restrict attention to direct mechanisms in which the CEO simply reports his private signal.

Technology and Beliefs

We model the uncertainty as to whether the CEO's strategy is a good match for the firm by assuming that the firm's cash flow under the CEO's strategy is a random variable $s \in S := [\underline{s}, \overline{s}]$, where $\underline{s} \geq 0$, and where $\overline{s}$ may be finite or infinite. The CEO's private interim signal is denoted by $\theta \in \Theta := [\underline{\theta}, \overline{\theta}]$. Each signal is associated with a conditional distribution function $G_\theta(s)$, whose density $g_\theta(s)$ is continuous in both θ and s. The signal is informative about s in the sense of the Monotone Likelihood Ratio Property (MLRP), implying that $g_{\theta'}(s)/g_\theta(s)$ is strictly

---

[12]Hermalin and Weisbach (1998, p.100) were the first to argue that what matters for the board's decision to replace the CEO is the quality of the *match*, not the CEO's "ability" in an absolute sense.

[13]See, e.g., Bagwell and Zechner (1993), Edlin and Stiglitz (1995), and Fulghieri and Hodrick (2005).

increasing in s for all $\theta' > \theta$ in Θ.[14] Among other things, this implies that the firm's expected cash flow under the CEO's strategy, $E[s \mid \theta]$, is increasing in θ. Initially when the CEO is hired, everyone has the same beliefs $F(\theta)$, where the density $f(\theta)$ is assumed to be positive for all θ. Hence, the common belief that the firm's cash flow under the CEO's strategy equals $s = s'$ is $\Pr[s = s'] = \int_{\Theta} g_{\theta}(s')f(\theta)d\theta$.

Realistically, whether the CEO's strategy is likely to be successful also depends on how dedicated the CEO is to implementing his strategy. We assume the following parsimonious effort problem: if the CEO puts in high effort, everything is as described above. That is, the distribution function $F(\theta)$ holds on the equilibrium path where the CEO puts in high effort. By contrast, if the CEO puts in low effort, we assume that this results in a sufficiently low signal under which the CEO's strategy has little chance of being successful, so that a strategy change becomes optimal.[15] Finally, putting in high effort is costly: it means that the CEO must forgo private benefits of $B > 0$.

To initiate a strategy change, the board must hire a new CEO with a different strategy (see Introduction). Let $V > 0$ denote the firm's expected cash flow under a potential replacement. There is no need to make distributional assumptions about V—what is important is only the firm's *expected* cash flow under a potential replacement. Moreover, note that V is independent of θ : the interim signal is match-specific in the sense that it only indicates whether the current CEO is a good match for the firm, but it contains no information as to how good a match a potential replacment might be. Finally, to make the problem nontrivial, we assume that a strategy change is optimal for some, but not all, values of $\theta \in \Theta$.[16]

CEO Compensation

The CEO's compensation package consists of his on-the-job pay $w(s, \widehat{\theta})$—which may depend on the firm's cash flow—and his severance pay $W(\widehat{\theta})$ if a replacement with a different strategy is hired. As we use an optimal contracting approach to solve the problem of the CEO having interim private information, both w and W may in principle depend on the CEO's reported signal

---

[14]MLRP is satisfied by many standard probability distributions (Milgrom, 1981).

[15]We may assume that low effort results in $\theta = \underline{\theta}$, although as we will see shortly, the precise specification is not important. What is important is only that θ is sufficiently low so that a strategy change becomes optimal.

[16]For instance, this is the case if $E[s \mid \theta]$ is sufficiently greater than V at high values of θ and sufficiently smaller than V at low values of θ, while $F(\theta)$ puts sufficient probability mass on both high and low values of θ.

$\widehat{\theta}$. We impose two constraints on the CEO's on-the-job pay scheme. The first constraint is that $w \leq s$, i.e., the CEO's on-the-job pay cannot exceed the firm's cash flow. The second constraint is that w must be nondecreasing in s. This latter constraint is only for expositional convenience. While it simplifies the derivation of our results, the constraint that w be nondecreasing in s does not bind at the optimal solution.

2.2   Preliminary Analysis

First-Best Benchmark

Let us briefly derive the first-best benchmark. The question is for what signals should the board maintain the strategic status quo, and for what signals should it initiate a strategy change. The first-best optimal decision is to maintain the status quo if $E[s \mid \theta] := \int_S s g_\theta(s) ds \geq V$ and to initiate a strategy change if $E[s \mid \theta] < V$. Given that $E[s \mid \theta]$ is strictly increasing in θ, and given that a strategy change is optimal for some, but not all, values of $\theta \in \Theta$, there exists a unique interior cutoff signal $\theta_{FB} \in (\underline{\theta}, \overline{\theta})$ defined by $E[s \mid \theta_{FB}] = V$ such that it is first-best optimal to maintain the status quo if $\theta \geq \theta_{FB}$ and to initiate a strategy change if $\theta < \theta_{FB}$.[17]

The Implementation Problem

We use an optimal contracting approach to solve the problem of the CEO having interim private information about his signal. Let $\Theta_+ \subset \Theta$ denote the set of all signals for which the status quo is maintained, and let $\Theta_- := \Theta \backslash \Theta_+$ denote the set of all signals for which a strategy change is initiated. From our earlier assumption that a strategy change is optimal for some, but not all, values of $\theta \in \Theta$, we have that both Θ_+ and Θ_- must be non-empty.

As we will show in Section 3.4, it is not optimal to make the CEO's on-the-job pay $w(s, \widehat{\theta})$— or, likewise, his severance pay $W(\widehat{\theta})$—contingent on his reported signal $\widehat{\theta}$. Rather, it is uniquely optimal to restrict the CEO's compensation package to a single on-the-job pay scheme $w(s)$ and a fixed severance payment W. To keep the exposition simple, we proceed as follows. We first consider a *restricted problem* in which we assume that $w(s, \theta) = w(s)$ for all $\theta \in \Theta_+$ and $W(\theta) = W$ for all $\theta \in \Theta_-$. Subsequently, we show that the solution to this restricted problem is also the unique solution to the general problem in which both w and W may depend on the reported signal $\widehat{\theta}$. That is, we will show that offering a richer menu of compensation schemes is

---

[17] As $G_\theta(s)$ satisfies MLRP, $E[s \mid \theta]$ is strictly increasing in θ. In conjunction with continuity of $g_\theta(s)$, this implies that there exists a unique interior cutoff signal $\theta_{FB} \in (\underline{\theta}, \overline{\theta})$ satisfying $E[s \mid \theta_{FB}] = V$.

not optimal—in fact, it is *strictly* suboptimal in our model.

In the restricted problem, the CEO's truthtelling constraints take the simple form

$$W \geq E[w(s) \mid \theta] \tag{1}$$

for all $\theta \in \Theta_-$ and

$$E[w(s) \mid \theta] \geq W \tag{2}$$

for all $\theta \in \Theta_+$.

Decisions (see Fudenberg and Tirole, 1992, Ch. 7) also take a simple form: either the status quo is maintained, in which case the CEO receives his on-the-job pay $w(s)$, or a strategy change is initiated, in which case the CEO receives his severance pay W. Hence there are only two possible outcomes. A *decision rule* is a mapping from the signal space Θ into the set of possible outcomes. Using standard terminology, a decision rule is *implementable* if it satisfies the truthtelling constraints (1)-(2). In what follows, we will show that the only decision rules that are implementable are simple cutoff rules—characterized by a unique interior cutoff signal $\theta^* \in (\underline{\theta}, \overline{\theta})$—where the status quo is maintained for all $\theta \geq \theta^*$ while a strategy change is initiated for all $\theta < \theta^*$.

The argument is as follows. First, we show that the CEO's expected on-the-job pay $E[w(s) \mid \theta]$ must be strictly increasing in θ. Clearly, it cannot be optimal to give the CEO a fixed wage $w(s) = w \neq W$ for all $s \in S$; he would otherwise always—i.e., for *all* signals $\theta \in \Theta$—either prefer w or W, depending on which of the two is higher, violating the requirement that both Θ_+ and Θ_- are non-empty. Arguably, if $w(s) = w = W$ the CEO is indifferent between w and W. In this case, however, the CEO would have no incentive to put in high effort. Given that $w(s)$ is nondecreasing, ruling out that $w(s) = w$ implies that $w(s)$ must be strictly increasing on a set of positive measure. In conjunction with the fact that $G_\theta(s)$ satisfies MLRP, this in turn implies that $E[w(s) \mid \theta]$ must be strictly increasing in θ.

That $E[w(s) \mid \theta]$ is strictly increasing in θ implies that if $E[w(s) \mid \theta'] \geq W$ for some value $\theta = \theta'$, then it must also hold that $E[w(s) \mid \theta] \geq W$ for all higher values $\theta > \theta'$. From continuity of $g_\theta(s)$, it then follows that there exists a unique interior cutoff signal $\theta^* = \theta^*(w(s), W) \in (\underline{\theta}, \overline{\theta})$ characterized by

$$E[w(s) \mid \theta^*] = W \tag{3}$$

such that $E[w(s) \mid \theta] \geq W$ if $\theta \geq \theta^*$ and $E[w(s) \mid \theta] < W$ if $\theta < \theta^*$, implying that the only decision rules that are implementable are simple cutoff rules with $\Theta_+ = [\theta^*, \overline{\theta}]$ and $\Theta_- = [\underline{\theta}, \theta^*)$.

9

Endogenous Benefits of Entrenchment

There are no exogenous (e.g., private) benefits of entrenchment in our model. The CEO's benefit from staying in office is that he receives his on-the-job pay $w(s)$, while his (opportunity) cost of staying in office is that he forgoes his severance pay W. Assuming that all costs and benefits of entrenchment are monetary implies that they can be jointly—and optimally—designed by the firm's board.

What makes the optimal design of $w(s)$ and W a nontrivial problem is that the CEO is biased in favor of staying in office. Precisely, the CEO's bias to stay in office derives from the simple effort problem introduced above. Consider the incentive constraint of inducing the CEO to put in high effort. If the CEO puts in low effort, we have assumed that this results in a low signal under which the CEO's strategy has little chance of being successful, so that a strategy change becomes optimal. But even "on the equilibrium path" where the CEO puts in high effort, there is a nontrivial probability $F(\theta^*) > 0$ that a strategy change will be initiated, simply because the CEO's strategy might prove to be a poor match for the firm. The CEO's effort constraint inducing him to put in high effort is therefore

$$\int_{\theta^*}^{\overline{\theta}} E[w(s) \mid \theta] f(\theta) d\theta + F(\theta^*) W \geq W + B, \tag{4}$$

which can be rewritten as

$$\int_{\theta^*}^{\overline{\theta}} \left(E[w(s) \mid \theta] - W \right) f(\theta) d\theta \geq B. \tag{5}$$

For the CEO to put in high effort, there must be consequently a wedge between his expected on-the-job pay $E[w(s) \mid \theta]$ and his severance pay W. Intuitively, since low effort results in the CEO receiving severance pay, severance pay constitutes a "reward" for the CEO putting in low effort. To induce the CEO to put in high effort, his expected on-the-job pay must consequently exceed his severance pay by a sufficient margin, similar to efficiency-wage models in which workers are motivated by the fact that their on-the-job pay exceeds their income if they are fired (e.g., Shapiro and Stiglitz, 1984; Yellen, 1994).

It is important to note that the wedge between $E[w(s) \mid \theta]$ and W required by (5) imposes no constraint on the functional form of $w(s)$. All this wedge requires is that "on average"—i.e., across all values of $\theta \in [\theta^*, \overline{\theta}]$—the CEO's *expected* on-the-job pay must exceed his severance pay by a sufficient margin, which in turn biases the CEO in favor of staying in office. For example, if $W = 0$, the CEO prefers to stay in office for *all* values of $\theta \in \Theta$, because $E[w(s) \mid \theta] > W = 0$.

10

Accordingly, there is no direct relation between the effort problem and the optimal functional form of $w(s)$—in contrast to "plain" moral hazard models, e.g., Innes (1990)—implying that our results regarding the optimal functional form of $w(s)$ are solely driven the problem of the eliciting the CEO's private information, not by the problem of motivating the CEO to put in high effort. All the effort problem does in our model is to create a bias on the part of the CEO to stay in office, which in turn makes the truthtelling problem nontrivial.[18]

3   The Optimal CEO Compensation Package

3.1   Basic Results

To motivate the tradeoff underlying the optimal design of the CEO's compensation scheme, consider again the effort constraint (4). By standard arguments, (4) must bind at the optimal solution, implying that the CEO's expected pay in case he puts in high effort—the left-hand side in (4)—must equal $W + B$. As the CEO's opportunity cost of putting in high effort is B, this implies that the CEO earns an informational rent equal to W.

Observation. *The CEO earns an informational rent equal to his severance pay W.*

To see why the CEO earns an informational rent, we can rewrite (4) more conveniently as

$$\int_{\theta^*}^{\overline{\theta}} E[w(s) \mid \theta] \frac{f(\theta)}{1 - F(\theta^*)} d\theta = W + \frac{B}{1 - F(\theta^*)}. \tag{6}$$

Accordingly, if the board wants to increase the CEO's severance pay by one dollar, it must also increase his expected on-the-job pay by one dollar.[19] Regardless of whether the status quo is maintained or a strategy change is initiated, the CEO is thus better off by one dollar. Given that there is a one-to-one relation between the CEO's severance pay and his on-the-job pay, both instruments are consequently measures of the CEO's informational rents. As for testable

---

[18]Another way of seeing this is as follows. One can easily show that if θ is observable and contractible—in which case there is no truthtelling problem—there exists an infinite number of compensation schemes $\{w(s), W\}$ that satisfy the effort constraint (5) and implement the first-best decision rule. For example, one can simply set $W = 0$ and pay the CEO a fixed wage of $w = B/[1 - F(\theta_{FB})]$.

[19]Precisely, the left-hand side in (6) represents the CEO's expected on-the-job pay *conditional* on maintaining the status quo.

predictions, (6) implies that changes in CEOs' severance pay and on-the-job pay must move in the same direction.

Granting the CEO severance pay is costly as it leaves him valuable informational rents. Without any severance pay, however, there would never be any strategy change. For example, if $W = 0$ the CEO would always want to stay in office, since $E[w(s) \mid \theta] > W = 0$ for all $\theta \in \Theta$. To analyze the tradeoff between initiating an efficient strategy change and leaving the CEO informational rents, we now turn to the board's maximization problem. The board chooses $w(s)$ and W to maximize

$$F(\theta^*)(V - W) + \int_{\theta^*}^{\overline{\theta}} E[s - w(s) \mid \theta]f(\theta)d\theta, \qquad (7)$$

subject to the CEO's binding effort constraint (4) and the constraint (3) characterizing the unique interior cutoff signal θ^*, which replaces the truthtelling constraints (1)-(2). Inserting the binding effort constraint into (7), we can rewrite the board's objective function as

$$\int_{\theta^*}^{\overline{\theta}} E[s \mid \theta]f(\theta)d\theta + F(\theta^*)V - B - W. \qquad (8)$$

The board has two conflicting objectives. The first three terms in (8) represent the total surplus created, implying that the board seeks to maximize overall efficiency. Precisely, since $E[s \mid \theta] - V$ is negative and decreasing in θ for $\theta < \theta_{FB}$ and positive and increasing in θ for $\theta > \theta_{FB}$, the board wants to push the cutoff signal θ^* as close as possible towards the first-best benchmark θ_{FB}. The first objective is thus to minimize the range of signals at which the firm forgoes an efficient strategy change. (As it is costly to push up θ^*, it will never be optimal to set $\theta^* > \theta_{FB}$.) The second objective follows from the last term in (8): minimize the CEO's severance pay and thus his informational rents.

Based on this tradeoff, we can state the board's problem of designing the optimal CEO compensation package in two equivalent ways. Holding the CEO's severance pay fixed, the optimal on-the-job pay scheme pushes up the cutoff signal θ^* as close as possible towards θ_{FB}, thus minimizing the range of signals at which the firm forgoes an efficient strategy change. Conversely, holding the cutoff signal θ^* fixed, the optimal on-the-job pay scheme minimizes the severance pay needed to implement θ^*, thus minimizing the CEO's informational rents. One problem is the dual of the other. We obtain the following result.

Proposition 1. *The uniquely optimal CEO compensation package consists of severance pay $W > 0$ and a bonus scheme paying $w(s) = 0$ if $s < \widehat{s}$ and $w(s) = s$ if $s \geq \widehat{s}$ for some $\widehat{s} \in (\underline{s}, \overline{s})$.*

12

Proof. See Appendix.

The uniquely optimal on-the-job pay scheme is a high-powered, discontinuous bonus scheme that shifts all of the CEO's on-the-job pay into the highest cash-flow states. The intuition is as follows. As low cash flows are relatively more likely after low signals—due to the fact that $G_\theta(s)$ satisfies MLRP—a bonus scheme of the sort described in Proposition 1 minimizes the CEO's *expected* on-the-job pay at low signals. As the prospect of receiving his on-the-job pay is why the CEO wants to stay in office, a bonus scheme of the sort described in Proposition 1 thus minimizes the CEO's incentives to entrench himself at low signals. This in turn implies that a relatively smaller severance payment is needed to induce the CEO *not* to entrench himself. In short, a bonus scheme of the sort described in Proposition 1 minimizes the informational rents that must be left to the CEO to implement a *given* cutoff signal θ^*.

The flip side of shifting all of the CEO's pay into the highest cash-flow states is that the CEO has strong incentives to stay in office when the signal is high. This is inconsequential, however, because at high signals both the CEO and the board prefer the status quo.

3.2 Continuous Versus Discontinuous On-The-Job Pay Schemes

As the previous analysis has shown, the cheapest (i.e., information-rent minimizing) way to implement a given cutoff signal θ^* is to shift all of the CEO's on-the-job pay into the highest cash-flow states. In a relatively unconstrained setting like the previous one, this implies that the optimal on-the-job pay scheme is a discontinuous bonus scheme paying $w(s) = 0$ if $s < \widehat{s}$ and $w(s) = s$ if $s \geq \widehat{s}$ for some $\widehat{s} \in (\underline{s}, \overline{s})$.[20] Such a discontinuous on-the-job pay scheme may entail problems of its own, however. For instance, if the firm's cash flow increases only slightly from $\widehat{s} - \varepsilon$ to $\widehat{s}$, the CEO's on-the-job pay jumps from zero to $w(s) = \widehat{s}$. To the extent that the CEO is able to manipulate the firm's cash flow, he will have strong incentives to do so.

In what follows, we examine how the optimal CEO on-the-job pay scheme ought to be modified so that it becomes immune to manipulation problems of the sort just described. Suppose that the CEO is able to manipulate the firm's cash flow at private cost $h(\Delta)$, where $\Delta = |s' - s|$, and where s' and s denote the manipulated and original cash flow, respectively. The cost function $h(\Delta)$ is assumed to be nondecreasing and convex with $h(0) = 0$ and $h'(0) = \gamma$, where γ is

---

[20]The only binding constraint in Proposition 1 is $w(s) \leq s$. As it is optimal to shift as much as possible of the CEO's on-the-job pay into the highest cash-flow states, the constraint that $w(s)$ be nondecreasing does not bind.

positive but small.[21]

Given that $h(\Delta)$ is nondecreasing and convex, any on-the-job pay scheme that is immune to "small" manipulations around $\Delta = 0$ is automatically also immune to "larger" manipulations $\Delta > 0$, implying that the manipulation problem gives rise to an additional constraint requiring that the CEO's marginal manipulation cost around $\Delta = 0$ must equal or exceed his marginal benefit from manipulation for all $s \in S$.[22] Given that optimality requires to shift as much as possible of the CEO's on-the-job pay into the highest cash-flow states, this additional constraint binds at the optimal solution. We obtain the following result.

Proposition 2. *The uniquely optimal compensation package if the CEO can manipulate the firm's cash flow consists of severance pay $W > 0$ and an "option-like" on-the-job pay scheme paying $w(s) = 0$ if $s < \widehat{s}$ and $w(s) = \gamma(s - \widehat{s})$ if $s \geq \widehat{s}$ for some $\widehat{s} \in (\underline{s}, \overline{s})$.*

Proof. See Appendix.

Because the constraint imposed by the manipulation problem binds, the optimal on-the-job pay scheme from Proposition 1 is no longer feasible. Instead of a discontinuous bonus scheme, the new optimal on-the-job pay scheme is now a continuous, "option-like" pay scheme under which the CEO receives $w(s) = 0$ up to some threshold $\widehat{s}$, while for all $s \geq \widehat{s}$ he receives a (small) fraction γ of the incremental cash flow above $\widehat{s}$.[23] Not only is the on-the-job pay scheme from Proposition 2 immune to manipulation, but it also has the appealing property that it is continuous. For this reason, our subsequent comparative static results are based on this scheme. The first result is immediate given the preceding discussion.

---

[21] The CEO's private cost of manipulating the firm's cash flows by Δ is likely to be significantly smaller than Δ—at least for small values of Δ—implying that $\gamma << 1$. The special case $\gamma = 1$ coincides with an assumption frequently found in the security design literature that $s - w(s)$ be nondecreasing (e.g., Innes, 1990; Nachman and Noe, 1994).

[22] Any on-the-job pay scheme that is immune to "small" manipulations from s to $s' = s + \varepsilon$ is also immune to "larger" manipulations from $s'' < s$ to s', because the CEO's private manipulation cost for the last increment from s to s' is weakly higher under the "larger" manipulation due to the convexity of the manipulation cost function $h(\Delta)$. An analogous argument holds for the case where $s' = s - \varepsilon$.

[23] The "strike price" $\widehat{s}$ is uniquely pinned down by the requirement that the effort constraint (4) binds. If γ is close to zero, (4) cannot be satisfied for any $\widehat{s} > \underline{s}$. Precisely, as γ goes to zero, the "strike price" $\widehat{s}$ approaches $\underline{s}$, implying that the optimal on-the-job pay scheme becomes $w(s) = F + \gamma s$, which can be implemented by giving the CEO a fixed wage of F plus stock.

Corollary 1. *As the CEO's marginal manipulation cost γ increases, the optimal on-the-job pay scheme becomes "steeper"—i.e., the pay-for-performance sensitivity increases—and less informational rents are needed to implement a given cutoff signal θ^*.*

Proof. See Appendix.

3.3 Severance Pay and Strategy Change

Our focus thus far has been on the design of the CEO's on-the-job pay scheme. For a given amount of severance pay W, the optimal on-the-job pay scheme pushes up the cutoff signal θ^* as close as possible toward θ_{FB}, minimizing the range of signals at which the firm forgoes an efficient strategy change. Conversely, for a given cutoff signal θ^*, the optimal on-the-job pay scheme minimizes the amount of severance pay—and thus informational rents—needed to implement θ^*. We now examine how an increase in severance pay affects the cutoff signal *given* that the CEO's on-the-job pay scheme is chosen optimally.

An increase in W reduces the CEO's incentives to stay in office for any given signal θ. At the same time, an increase in severance pay must be accompanied by a simultaneous increase in the CEO's expected on-the-job pay to preserve the wedge required by (4), which makes it *more* attractive for the CEO to stay in office. One can show that under the *optimal* on-the-job pay scheme the first effect outweighs the second, implying that an increase in severance pay pushes the cutoff signal θ^* upward, thus narrowing the range of signals at which the firm forgoes an efficient strategy change.

The intuition is as follows. Under the optimal on-the-job pay scheme, the increase in $w(s)$ needed to match an increase in W occurs at relatively high cash flows, implying that the CEO's expected on-the-job pay $E[w(s) \mid \theta]$ increases primarily at high signals. (This follows from MLRP and holds for both the optimal on-the-job pay scheme from Propositions 1 and 2.) On the other hand, $E[w(s) \mid \theta]$ increases only relatively little at low signals. Hence, while "on average" the CEO's expected on-the-job pay $E[w(s) \mid \theta]$ increases one-for-one along with his severance pay to satisfy (4), it increases by more than W at high signals and by less than W at low signals. At relatively low signals the difference $E[w(s) \mid \theta] - W$ thus *decreases*, implying that an increase in W pushes the cutoff signal θ^* upward and closer towards θ_{FB}.[24]

---

[24] While the optimal value of W trades off the benefits of reducing the range of signals at which the firm forgoes an efficient strategy change against the costs of leaving the CEO informational rents, the *specific* solution depends

Proposition 3. *Under the optimal CEO on-the-job pay scheme, an increase in severance pay narrows the range of signals at which the firm forgoes an efficient strategy change. Implementing the first best is too costly, however: at the optimal solution it holds that $\theta^* < \theta_{FB}$.*

Proof. See Appendix.

According to Proposition 3, it is too costly to implement the first best. Intuitively, a marginal decrease in θ^* at $\theta^* = \theta_{FB}$ has a negligible effect on shareholders' wealth. A marginal reduction in the CEO's informational rents, however, constitutes a first-order cost saving.

The magnitude of executives' severance packages has been the subject of much recent discussion. A commonly found argument is that severance pay constitutes a "reward" for failure, which is not only unnecessary but also counterproductive in the sense that it reduces CEOs' incentives to put in effort.[25] This is precisely what would also happen in our model *if* the CEO's severance pay were set too high relative to his on-the-job pay. On the other hand, our model shows that introducing a binding cap on CEOs' severance pay may come at a high cost: while a binding cap reduces the CEO's informational rents, it pushes the cutoff signal θ^* downward, thereby widening the range of signals at which the firm forgoes an efficient strategy change.

3.4   Menu of Compensation Schemes

Hitherto we have considered a restricted problem in which $w(s, \theta) = w(s)$ for all $\theta \in \Theta_+$ and $W(\theta) = W$ for all $\theta \in \Theta_-$. It remains to show that the solution to this restricted problem is also the unique solution to the general problem in which both w and W may directly depend on the CEO's reported signal $\widehat{\theta}$.

It is easy to show that it cannot be optimal to make the CEO's severance pay contingent on his reported signal: conditional on reporting some $\widehat{\theta} \in \Theta_-$, the CEO would always

---

on distributional assumptions. Once the optimal value of W is pinned down, the only remaining choice variable $\widehat{s}$ (and hence θ^*) is uniquely pinned down by (4).

[25] A prominent example is the lawsuit by Walt Disney shareholders against the company for awarding Michael Ovitz severance pay worth $130 million after being only 14 months with Disney. Public pressure against high severance packages is not limited to the United States. The United Kingdom, for instance, had a public inquiry about "rewards for failure" (DTI, 2003), and it has witnessed substantial shareholder activity against high severance packages. As a result, listing rules were amended in 2002 to require firms to publish their directors' remuneration reports, which must be approved by shareholders.

report the signal that yields him the highest severance pay, i.e., he would always report $\widehat{\theta} \in \arg\max_{\theta' \in \Theta_-} W(\theta')$. Therefore, the CEO's severance pay must be a constant $W(\widehat{\theta}) = W$.

There is a similarly straightforward argument why there is no point in making the CEO's on-the-job pay contingent on his reported signal. As the choice between maintaining the status quo and initiating a strategy change is a binary choice, all that matters for this choice is whether θ is an element of Θ_- or Θ_+. Conditional on knowing that $\theta \in \Theta_+$, there is no value from having more precise information about the CEO's signal.

We can prove an even stronger claim, namely, that making the CEO's on-the-job pay contingent on his reported signal is *strictly* suboptimal in our setting. The intuition is as follows. By construction, the "single" optimal on-the-job pay scheme from Propositions 1 and 2 *minimizes* the CEO's expected on-the-job pay at low signals. Any richer menu of on-the-job pay schemes $w(s, \widehat{\theta})$—regardless of whether this menu includes the "single" optimal on-the-job pay scheme—must therefore shift some of the CEO's expected on-the-job pay "back" into low-signal states. That is, if the CEO can choose from a richer menu $w(s, \widehat{\theta})$, he will not choose the "single" optimal on-the-job pay scheme from Propositions 1 and 2 at low signals, but rather some other on-the-job pay scheme that provides him with a higher expected pay at low signals. Consequently, a richer menu $w(s, \widehat{\theta})$ no longer minimizes the CEO's expected on-the-job pay $E[w \mid \theta]$ at low signals. But this property of minimizing $E[w \mid \theta]$ at low signals is what drives optimality in our model, for it minimizes the informational rents needed to implement a given cutoff signal θ^*. By implication, if the board wanted to offer the CEO a richer menu $w(s, \widehat{\theta})$, it would either have to leave the CEO more informational rents to implement the same cutoff signal, or—holding informational rents constant—it would have to tolerate a lower cutoff signal and thus a wider range of signals at which the firm forgoes an efficient strategy change.

Proposition 4. *It is strictly suboptimal to make either the CEO's on-the-job pay or his severance pay contingent on his reported signal.*

Proof. See Appendix.

4 Comparative Statics

The preceding analysis has shown that if the board wants to implement a higher cutoff signal θ^*, it must leave the CEO more informational rents. In our model, the magnitude of the CEO's

informational rents is equal to his severance pay. Moreover, there is a positive, one-to-one mapping between the CEO's severance pay and his expected on-the-job pay, implying that both instruments are measures of the CEO's informational rents. We now examine factors that determine the cutoff signal θ^* and thus, by implication, the CEO's informational rents. The factors we consider are the uncertainty of the firm's business environment, firm size, and the firm's corporate governance.

External Business Environment

For firms in stable business environments—where it is unlikely that a strategy change becomes optimal—the expected cost of forgoing an efficient strategy change is relatively small. Conversely, for firms in unstable business environments the expected cost of forgoing an efficient strategy change is relatively large. In our model, the likelihood that a strategy change becomes optimal is $F(\theta_{FB})$. Consider an increase in θ_{FB}.[26] To increase the likelihood that a strategy change will be indeed initiated, the board must raise the second-best cutoff signal θ^*. Raising θ^* is costly, however. It implies that the board must leave the CEO more informational rents and, by implication, more severance pay and more on-the-job pay.

To ensure a uniquely optimal value of W, we assume that the board's objective function is strictly quasiconcave in W. Once the optimal value of W is pinned down, the only remaining choice variable $\hat{s}$ (and hence θ^*) is uniquely determined. We have the following result.

Proposition 5. *CEOs of firms in unstable business environments should earn more informational rents—and receive both higher severance pay and higher on-the-job pay—than CEOs of firms in stable business environments.*

Proof. See Appendix.

Proposition 5 is consistent with evidence by Rusticus (2006) showing that CEOs' (contractual) severance pay is positively related to measures proxying for the uncertainty of firms' operating environments and the likelihood of strategy changes, as well as evidence by Rusticus (2006), Schwab and Thomas (2004), and Lefanowicz, Robinson, and Smith (2000) showing that CEOs' severance pay and golden parachutes are positively related to CEOs' on-the-job pay.

---

[26] An increase in uncertainty alone—e.g., in the form of a mean-preserving spread—does not necessarily imply that the likelihood that a strategy change becomes optimal increases. What matters is that θ_{FB} increases.

The result may also help to shed light on some developments that have taken place over the past decades. As Holmström and Kaplan (2001) argue, the "pace of economic change has accelerated" since the late 1970s. Since 1978, some of the most important industries in the United States—e.g., airlines, broadcasting, entertainment, natural gas, trucking, banks and thrifts, utilities, and telecommunications—have undergone massive deregulations. The 1980s and 90s also witnessed important technological innovations, notably in the information technology and telecom sectors, that have radically altered the industrial landscape. Both these developments—key forces behind what Jensen (1993) calls "Modern Industrial Revolution"—have increased the pressure on firms to adapt their business strategies at an increasingly faster pace.

In the light of these developments, Proposition 5 suggests that—along with the "increasing pace of economic change"—CEO pay should also increase, which is consistent with evidence by Hall and Liebman (1998) that the mean value of CEO stock option grants has increased almost sevenfold between 1980 and 1994. Similarly, Bebchuk and Grinstein (2005) find that the average CEO pay among S&P 500 firms has increased almost threefold between 1993 and 2003. As for severance pay, Walker (2005) points to a surge in (contractual) severance pay, while Lefanowicz, Robinson, and Smith (2000) and Bebchuk, Cohen, and Ferrell (2004) find that both the usage and magnitude of golden parachutes has increased during the 1980s and 90s.[27]

Cross-sectionally, Proposition 5 implies that CEO pay should be higher in unstable industries. There is evidence that CEO pay is lower in regulated industries which, one might argue, are more stable (Murphy, 1999). In a similar fashion, Crawford, Ezzell, and Miles (1995) and Hubbard and Palia (1995) conclude that deregulation appears to have contributed to the rise of CEO pay in the banking industry.

An extreme case of a strategy change is the firm's shutdown. Proposition 5 implies that the best way to induce managers to accept the necessary shutdown of their firms is a combination of severance pay and high-powered on-the-job pay, such as bonus schemes and option grants. Consistent with this argument, Mehran, Norgler, and Schwartz (1997) provide empirical evidence that option grants have a positive effect on the likelihood of voluntary liquidation. The argu-

---

[27]Our argument for the rise in CEO pay over the past decades is different from Dow and Raposo (2005a), who also link this trend to firms' business environments becoming more unstable over time. In their model, CEOs receive higher pay because the range of possible business strategies has increased, and shareholders want to induce CEOs not to pursue (overtly radical) strategy changes. For alternative arguments, see, e.g., Almazan and Suarez (2003), Bebchuk and Fried (2004), Murphy and Zábojník (2004), and Gabaix and Landier (2006).

ment is also consistent with Dial and Murphy's (1995) clinical study of General Dynamic's partial liquidation, where stock option programs helped to overcome executives' resistance against liquidating plants, even if this meant that the executives had to sacrifice their own jobs.

Firm Size

Another determinant of the CEO's informational rents in our model is firm size. For any given signal $\theta < \theta_{FB}$, the cost of forgoing an efficient strategy change is higher for larger firms. To mitigate these higher costs, larger firms should have a higher cutoff signal θ^*, which in turn implies that CEOs of larger firms should receive both higher severance pay and higher on-the-job pay. To examine the relation between firm size and CEO pay, we scale both V and s by a factor $\alpha > 0$. We obtain the following result.

Proposition 6. *CEOs of larger firms should earn more informational rents—and receive both higher severance pay and higher on-the-job pay—than CEOs of smaller firms.*

Proof. See Appendix.

Consistent with Proposition 6, Conyon (1997), Schwab and Thomas (2004), Bebchuk and Grinstein (2005), and Rusticus (2006) all find that CEO pay increases with firm size.

Corporate Governance

The (binding) effort constraint (4) suggests that the magnitude of CEO pay depends on two factors. First, the CEO must be compensated for forgoing his private benefits of B. Second, on top of this compensation, the CEO must be left an informational rent equal to W to induce truthtelling. We now examine how these two factors interact.

Suppose B increases. The direct effect, which follows from (4), is that the CEO's on-the-job pay must increase to preserve his incentives to put in high effort. But this makes it only more attractive for the CEO to stay in office at any given signal $\theta \in \Theta$, pushing the cutoff signal θ^* down. To counteract the decrease in θ^*, the board must increase the CEO's severance pay, which implies that it must (once again) increase the CEO's on-the-job pay. As Proposition 7 shows, it is not optimal to push θ^* all the way back to its original level, however, implying that an increase in B results in a lower cutoff signal, less informational rents, and lower CEO pay.

The size of B is likely to depend on the firm's corporate governance, which imposes a constraint on the CEO's ability to extract private benefits. In this regard, B is inversely related

to the quality of the firm's corporate governance: it is small in firms with good corporate governance and large in firms with bad corporate governance. We have the following result.

Proposition 7. *CEOs of firms with good corporate governance should earn less informational rents—and receive both lower severance pay and lower on-the-job pay—than CEOs of firms with bad corporate governance.*

Proof. See Appendix.

5   Conclusion

We consider the joint optimal design of CEOs' severance pay and on-the-job pay in a model in which the CEO has interim private information about the likely success of his strategy. To safeguard his own position, the CEO may withhold information from the board indicating that a strategy change is desirable. While offering the CEO severance pay induces truthful revelation in low-signal states, it leaves the CEOs informational rents. When designing the CEO's compensation package, the board thus faces a tradeoff. On the one hand, it wants to minimize the range of signals at which the firm forgoes an efficient strategy change. On the other hand, it wants to minimize the CEO's informational rents. Using an optimal contracting approach, we find that the uniquely optimal CEO compensation package is a combination of severance pay and high-powered on-the-job pay that shifts all of the CEO's pay into the highest cash-flow states, such as bonus schemes and option grants. Moreover, despite the CEO having private information about the likely success of his strategy, the optimal menu takes a simple form: it consists of fixed severance pay and a single on-the-job pay scheme. That is, neither the CEO's severance pay nor his on-the-job pay are contingent on his reported signal.

Our model shows that CEOs' severance pay and on-the-job pay are endogenously tied together, implying that there is a delicate balance between the two instruments. In particular, they must move in the same direction. Our model also implies that CEOs of firms in volatile business environments should earn more informational rents—and receive both higher severance pay and higher on-the-job pay—than CEOs of firms in stable environments. Finally, our model suggests that along with the increase in strategic uncertainty over the past decades, CEOs' informational rents—and thus their severance pay and on-the-job pay—should have increased.

6 Appendix

Proof of Proposition 1. The fact that $W > 0$ follows from the argument in the main text. It remains to prove that it is uniquely optimal to give the CEO an on-the-job pay scheme of the form $w(s) = 0$ if $s < \hat{s}$ and $w(s) = s$ if $s \geq \hat{s}$ for some $\hat{s} \in (\underline{s}, \overline{s})$.

We argue to a contradiction. Suppose it was optimal to implement a given cutoff signal θ^* with a different on-the-job pay scheme $\widetilde{w}(s)$, and denote the corresponding severance pay by $\widetilde{W}$. We show that there then exists some on-the-job pay scheme $w(s)$ such that (i) the incentive constraint (4) remains binding and (ii) we can still implement θ^*—though now with a lower severance pay W. That is, in a slight abuse of notation, we show that the new on-the-job pay scheme satisfies $\theta^*(w, W) = \theta^*(\widetilde{w}, \widetilde{W}) = \theta^*$ and $W < \widetilde{W}$, which by inspection of (7), contradicts the optimality of $\widetilde{w}(s)$.

We proceed in two steps. We first choose $\overline{W} = \widetilde{W}$ and $\overline{w}(s)$ with $\overline{w}(s) = 0$ for $s < \hat{s}'$ and $\overline{w}(s) = s$ for $s \geq \hat{s}'$ such that $\theta^*(\overline{w}, \overline{W}) = \theta^*$. That is, defining $d(s) := \widetilde{w}(s) - \overline{w}(s)$, we have that

$$\int_S d(s) g_{\theta^*}(s) ds = 0. \tag{9}$$

Given the construction of $\overline{w}(s)$, there exists a value $\widetilde{s} \in (\underline{s}, \overline{s})$ such that $d(s) \geq 0$ for all $s < \widetilde{s}$ and $d(s) \leq 0$ for all $s > \widetilde{s}$, where both inequalities are strict over sets of positive measure. Take now any signal $\widehat{\theta} > \theta^*$. By MLRP of $G_\theta(s)$ and (9), it then holds that

$$\int_S d(s) g_{\widehat{\theta}}(s) ds = \int_{\underline{s}}^{\widetilde{s}} d(s) g_{\theta^*}(s) \frac{g_{\widehat{\theta}}(s)}{g_{\theta^*}(s)} ds + \int_{\widetilde{s}}^{\overline{s}} d(s) g_{\theta^*}(s) \frac{g_{\widehat{\theta}}(s)}{g_{\theta^*}(s)} ds \tag{10}$$

$$< \frac{g_{\widehat{\theta}}(\widetilde{s})}{g_{\theta^*}(\widetilde{s})} \int_S d(s) g_{\theta^*}(s) ds = 0,$$

which implies the incentive constraint (4) is slack under $\overline{w}(x)$ and $\overline{W}$.

In a second step, we can now construct the asserted pay scheme with $w(s) = 0$ for $s < \hat{s}$ and $w(s) = s$ for $s \geq \hat{s}$ and $W < \overline{W} = \widetilde{W}$. In order to do this, we continuously increase the threshold $\hat{s}'$ in $\overline{w}(s)$ and decrease $\overline{W}$, while still satisfying $\theta^*(\overline{w}, \overline{W}) = \theta^*$, until (4) again binds. The fact that this is possible follows from continuity of all payoffs in $\hat{s}'$ and the fact that (4) is violated as $\hat{s}' \to \overline{s}$. Q.E.D.

Proof of Proposition 2. By the argument in the main text, the manipulation problem adds one additional constraint to the maximization problem: the slope of $w(s)$ must not exceed γ. We argue to a contradiction and assume we want to implement a given cutoff signal θ^* with

different on-the-job pay scheme $\widetilde{w}(s)$. Like in the Proof of Proposition 1, we can then again construct $\overline{w}(s)$ with $\overline{w}(s) = 0$ for $s < \widehat{s}'$ and $\overline{w}(s) = \gamma(s - s)$ for $s \geq \widehat{s}'$ such that (9) is satisfied. As the slope of $\widetilde{w}(s)$ cannot exceed γ, there then exists a value $\widetilde{s} \in (\underline{s}, \overline{s})$ such that $d(s) \geq 0$ for all $s < \widetilde{s}$ and $d(s) \leq 0$ for all $s > \widetilde{s}$, where both inequalities are strict over sets of positive measure. The rest of the argument is identical to that in the Proof of Proposition 1. Q.E.D.

Proof of Corollary 1. By Proposition 2, an increase in γ increases the slope of the optimal on-the-job pay scheme to the right of the threshold $\widehat{s}$.[28] It remains to prove that as γ increases, implementing a given cutoff signal θ^* requires a lower amount of severance pay.

To prove this, we totally differentiate (3), which pins down θ^*, and also the constraint (6) to obtain (holding θ^* fixed)

$$\frac{dW}{d\gamma} = \frac{\int_{\theta^*}^{\overline{\theta}} \left[\int_{\widehat{s}}^{\overline{s}} \left[[1 - G_\theta(\widehat{s})] [1 - G_{\theta^*}(s)] - [1 - G_{\theta^*}(\widehat{s})] [1 - G_\theta(s)] \right] ds \right] f(\theta) d\theta}{\int_{\theta^*}^{\overline{\theta}} [G_{\theta^*}(\widehat{s}) - G_\theta(\widehat{s})] f(\theta) d\theta}. \tag{11}$$

The denominator of (11) is positive as $G_\theta(s)$ satisfies FOSD, which is implied by MLRP. To see that the numerator is negative, note that $[1 - G_\theta(s)] / [1 - G_{\theta^*}(s)]$ is strictly increasing in s for all $\theta > \theta^*$, which is again implied by MLRP.[29] Q.E.D.

Proof of Proposition 3. Totally differentiating (3), which pins down θ^*, and the constraint (6) while substituting the optimal on-the-job compensation scheme from Proposition 2 yields

$$\frac{d\theta^*}{dW} = -\frac{1}{\gamma} \frac{\int_{\theta^*}^{\overline{\theta}} [G_{\theta^*}(\widehat{s}) - G_\theta(\widehat{s})] f(\theta) d\theta}{\left[\int_{\widehat{s}}^{\overline{s}} \frac{dG_{\theta^*}(\widehat{s})}{d\theta^*} ds \right] \left[\int_{\theta^*}^{\overline{\theta}} [1 - G_\theta(\widehat{s})] f(\theta) d\theta \right]}. \tag{12}$$

To evaluate the sign of (12), note that MLRP implies that $G_\theta(s)$ is decreasing in θ for all $s \in (\underline{s}, \overline{s})$, implying that $d\theta^*/dW > 0$.

Consider next the claim that $\theta^* < \theta_{FB}$. Totally differentiating (3) and (6) while substituting the optimal on-the-job compensation scheme from Proposition 2 yields

$$\frac{d\widehat{s}}{dW} = -\frac{1}{\gamma} \frac{1 - F(\theta^*)}{\int_{\theta^*}^{\overline{\theta}} [1 - G_\theta(\widehat{s})] f(\theta) d\theta}. \tag{13}$$

---

[28] Moreover, holding either θ^* or W fixed, to keep the incentive constraint (4) binding, the threshold $\widehat{s}$ must shift to the right as γ increases.

[29] As $G_\theta(s)$ is differentiable, this is equivalent to requiring that $g_\theta(s)/[1 - G_\theta(s)]$ is strictly decreasing in θ for any given $s \in (\underline{s}, \overline{s})$. This is the Monotone Hazard Rate Property (MHRP), which is implied by MLRP. To obtain this expression, we use the fact that $E[w(s) \mid \theta] = \gamma \int_{\widehat{s}}^{\overline{s}} [1 - G_\theta(s)] ds$, which follows from partial integration.

Next, differentiating (7) with respect to W and substituting $d\widehat{s}/dW$ from (13), we obtain

$$-\frac{d\theta^*}{dW} f(\theta^*) \left[E[s \mid \theta^*] - V \right] - 1, \tag{14}$$

which yields the first-order condition

$$E[s \mid \theta^*] - V = \frac{-F(\theta^*)}{f(\theta^*)(d\theta^*/dW)}. \tag{15}$$

Given that $d\theta^*/dW > 0$ from (12), we have at an interior solution that $E[s \mid \theta^*] - V < 0$ and therefore that $\theta^* < \theta_{FB}$. Q.E.D.

Proof of Proposition 4. By the argument in Section 2.2, we can restrict consideration to on-the-job pay schemes $w(s, \theta)$ that are strictly increasing in s on a set of positive measure. In conjunction with the fact that $G_\theta(s)$ satisfies MLRP, truthtelling then implies that $\Theta_- = [\underline{\theta}, \theta^*)$ and $\Theta_+ = [\theta^*, \overline{\theta}]$ with $E[w(s, \theta^*) \mid \theta^*] = W$. The following auxiliary result follows now immediately from the Proof of Proposition 1.

Claim 1. *Take two different feasible on-the-job pay schemes $\widetilde{w}(s)$ and $\widehat{w}(s)$ such that $\widehat{w}(s) = 0$ for $s < \widehat{s}$ and $\widehat{w}(s) = s$ for $s \geq \widehat{s}$. If $E[\widehat{w}(s) \mid \theta'] \geq E[\widetilde{w}(s) \mid \theta']$ for some $\theta' < \overline{\theta}$, then $E[\widehat{w}(s) \mid \theta'] > E[\widetilde{w}(s) \mid \theta']$ for all $\theta > \theta'$.*

To complete the proof, we must distinguish between two cases. If $w(s, \theta^*)$ satisfies $w(s, \theta^*) = 0$ for $s < \widehat{s}$ and $w(s, \theta^*) = s$ for $s \geq \widehat{s}$, Claim 1 and truthtelling imply that the same on-the-job pay scheme is also chosen for all $\theta \geq \theta^*$. That is, the optimal menu is degenerate with $w(s, \theta) = w(s, \theta^*)$. Suppose next that $w(s, \theta^*)$ takes a different form as above. As in the Proof of Proposition 1, we can then construct a single on-the-job pay scheme $\widehat{w}(s)$ satisfying $\widehat{w}(s) = 0$ for $s < \widehat{s}$ and $\widehat{w}(s) = s$ for $s \geq \widehat{s}$ such that the same cutoff signal θ^* is implemented while the effort constraint is relaxed. This follows from the fact that $E[\widehat{w}(s) \mid \theta] > E[w(s, \theta) \mid \theta]$ for all $\theta > \theta^*$, which in turn follows from Claim 1 and the truthtelling requirement for the original menu. As in Proposition 1, we can finally adjust the new (single) on-the-job pay scheme $\widehat{w}(s)$ so as to implement θ^* with a lower severance pay. Q.E.D.

Proof of Proposition 5. We show that the optimal choice of W is strictly increasing in V, which by (12) implies that the corresponding optimal choice of θ^* is also strictly increasing. Implicit differentiation of the first-order condition for W in (15) gives

$$\frac{dW}{dV} = -\frac{f(\theta^*)(d\theta^*/dW)}{SOC} > 0,$$

24

where $SOC < 0$ must hold at an interior optimum. Q.E.D.

Proof of Proposition 6. We show that the optimal choice of W is strictly increasing in α, which by (12) implies that the corresponding optimal choice of θ^* is also strictly increasing. Note first that the first-order condition (15) now transforms to

$$\alpha \left[E[s \mid \theta^*] - V \right] = \frac{-F(\theta^*)}{f(\theta^*)(d\theta^*/dW)}, \tag{16}$$

where we can again substitute $d\theta^*/dW$ from (12). Implicit differentiation of (16) gives

$$\frac{dW}{d\alpha} = \frac{E[s \mid \theta^*] - V}{SOC} > 0,$$

where we used the result that at the optimum it holds that $\theta^* < \theta_{FB}$, while $SOC < 0$ must hold at an interior optimum. Q.E.D.

Proof of Proposition 7. We show first that in order to implement a given cutoff signal θ^*, the higher is B the higher must also be the severance pay W. We totally differentiate (3), which determines θ^*, and the constraint (6) to obtain

$$\frac{dW}{dB} = \frac{1 - G_{\theta^*}(\widehat{s})}{\int_{\theta^*}^{\overline{\theta}} \left[G_{\theta^*}(\widehat{s}) - G_\theta(\widehat{s}) \right] f(\theta) d\theta} > 0, \tag{17}$$

where the sign follows again from MLRP of $G_\theta(s)$, which implies FOSD.

Take now some value $B = \widehat{B}$. The optimal compensation package specifies an amount of severance pay $W = \widehat{W}$ and some on-the-job pay scheme $w(s) = \widehat{w}(s)$, which is in turn characterized by a unique threshold (or "strike price") $\widehat{s} = \widehat{s}'$. Denote the corresponding cutoff signal by $\theta^* = \widehat{\theta}^*$. If $B = \widetilde{B} > \widehat{B}$, we know from (17) that in order to implement the same cutoff signal $\theta^* = \widehat{\theta}^*$, the severance pay must increase: $W = \widetilde{W} > \widehat{W}$. To still satisfy (6), the CEO's expected on-the-job pay must also increase, i.e., the new threshold $\widehat{s} = \widehat{s}''$ must satisfy $\widehat{s}'' < \widehat{s}'$.

Consider next the derivative (14) of the board's objective function. By construction, we have for $B = \widehat{B}$, $\widehat{s} = \widehat{s}'$, $W = \widehat{W}$, and $\theta^* = \widehat{\theta}^*$ that the derivative is just zero. (This is just the first-order condition.) We now want to evaluate the sign of the derivative when we substitute $B = \widetilde{B}$, $\widehat{s} = \widehat{s}''$, $W = \widetilde{W}$, and $\theta^* = \widehat{\theta}^*$, i.e., we want to evaluate the sign of the derivative at the point where with higher private benefits the same cutoff signal is implemented, albeit with higher severance pay and a higher expected on-the-job pay. More precisely, we want to show that the derivative (14) is then negative. By inspection of (14), this is the case if at $\theta^* = \widehat{\theta}^*$

the derivative $d\theta^*/dW$ is strictly lower when $B = \widetilde{B}$ and thus $W = \widetilde{W}$ and $\widehat{s} = \widehat{s}''$. Given that $\widehat{s}'' < \widehat{s}'$, this in turn holds if the derivative (12) is strictly increasing in $\widehat{s}$. To show that this is the case, we rearrange (12) to obtain

$$\frac{d\theta^*}{dW} = \frac{1}{\gamma} \left(\frac{-1}{\int_{\widehat{s}}^{\overline{s}} \frac{dG_{\theta^*}(s)}{d\theta^*}ds} \right) \left(\frac{\int_{\theta^*}^{\overline{\theta}} \left[G_{\theta^*}(\widehat{s}) - G_{\theta}(\widehat{s}) \right] f(\theta)d\theta}{\left[\int_{\theta^*}^{\overline{\theta}} \left[1 - G_{\theta}(\widehat{s}) \right] f(\theta)d\theta \right]} \right). \tag{18}$$

The first expression in parentheses is positive by $dG_{\theta^*}(\widehat{s})/d\theta^* < 0$, which is implied by FOSD and thus by MLRP, strictly increasing in $\widehat{s}$. Next, after some transformations, we have that the sign of the derivative of the last term in (18) with respect to $\widehat{s}$ is given by the expression

$$\int_{\theta^*}^{\overline{\theta}} \left[g_{\theta^*}(\widehat{s}) \left[1 - G_{\theta}(\widehat{s}) \right] - g_{\theta}(\widehat{s}) \left[1 - G_{\theta^*}(\widehat{s}) \right] \right] f(\theta)d\theta > 0. \tag{19}$$

The fact that (19) is also strictly positive follows again from MLRP, by which $g_{\theta}(s)/[1 - G_{\theta}(s)]$ must be strictly decreasing in θ for all $s \in (\underline{s}, \overline{s})$. (Recall that MLRP implies MHRP.) Hence, we have shown that given $B = \widetilde{B}$, if we evaluate (14) at the value $W = \widetilde{W}$ where $\theta^* = \widetilde{\theta}^*$, then the derivative is strictly negative. Given strict quasiconcavity of the objective function and the fact that θ^* (and thus, in particular, $\widetilde{\theta}^*$) is interior, we thus have that for $B = \widetilde{B}$ the optimal severance pay is strictly lower than $W = \widetilde{W}$. But this finally implies that under the optimal compensation package there is more entrenchment if $B = \widetilde{B}$ than if $B = \widehat{B} < \widetilde{B}$. Q.E.D.

7   References

Almazan, A., and J. Suarez, 2003, Entrenchment and severance pay in optimal governance structures, Journal of Finance 58, 519-548.

Bagwell, L. S., and J. Zechner, 1993, Influence costs and capital structure, Journal of Finance 48, 975-1008.

Bebchuk, L. A., and J. M. Fried, 2004, Pay without performance. Harvard University Press: Cambridge, MA.

Bebchuk, L. A., and Y. Grinstein, 2005, The growth of executive pay, Oxford Review of Economic Policy 21, 283-303.

Bertrand, M., and A. Schoar, 2003, Managing with style: The effect of managers on firm policies, Quarterly Journal of Economics 118, 1169-1208.

Carpenter, J. N., 2000, Does option compensation increase managerial risk appetite? Journal of Finance 55, 2311-2331.

Comin, D., and T. Philippon, 2005, The rise in firm-level volatility: Causes and consequences, NBER Macroeconomic Annual.

Conyon, M. J., 1997, Corporate governance and executive compensation, International Journal of Industrial Organization 15, 493-509.

Crawford, A. J., J. R. Ezzell, and J. A. Miles, 1995, Bank CEO pay-performance relations and the effects of deregulation, Journal of Business, 68, 231-256.

Dial, J., and K. J. Murphy, 1995, Incentives, downsizing, and value creation in General Dynamics, Journal of Financial Economics 37, 261-314.

Dittmann, I., and E. Maug, 2006, Lower salaries and no options? On the optimal structure of executive pay, forthcoming, Journal of Finance.

Dow, J., and C. C. Raposo, 2005a, CEO compensation, change, and corporate strategy, Journal of Finance 60, 2701-2727.

Dow, J., and C. C. Raposo, 2005b, On the optimality of entrenched lame ducks: An economic model of leadership and organizational change, Mimeo, London Business School.

DTI, 2003, Rewards for failure: Directors' remuneration—contracts, performance and severance, Department of Trade and Industry, London.

Edlin, A. S., and J. E. Stiglitz, 1995, Discouraging rivals: Managerial rent-seeking and economic inefficiencies, American Economic Review 85, 1301-1312.

Eisfeldt, A. L., and A. A. Rampini, 2006, Managerial incentives, capital reallocation, and the business cycle, forthcoming, Journal of Financial Economics.

Finkelstein, S., and D. C. Hambrick, 1990, Top-management team tenure and organizational outcomes: The moderating role of managerial discretion, Administrative Science Quarterly 35, 484-503.

Fudenberg, D., and J. Tirole, 1992. Game Theory. Cambridge: MIT Press.

Fulghieri, P., and L. S. Hodrick, 2005, Synergies and internal agency conflicts: The double-edged sword of mergers, forthcoming, Journal of Economics and Management Strategy.

Gabaix, X., and A. Landier, 2006, Why has CEO pay increased so much? Mimeo, New York University.

Geletkanycz, M. A., 1997, The salience of 'culture's consequences': The effects of cultural values on top executive commitment to the *status quo*, Strategic Management Journal 18, 615-634.

Geletkanycz, M. A., and S. S. Black, 2001, Bound by the past? Experience-based effects on commitment to the strategic *status quo*, Journal of Management 27, 3-21.

Hall, B. J., and J. B. Liebman, 1998, Are CEOs really paid like bureaucrats? Quarterly Journal of Economics 113, 653-691.

Harris, E. G., 1990, Antitakeover measures, golden parachutes, and target firm shareholder welfare, Rand Journal of Economics 21, 614-625.

Hermalin, B. E., and M. S. Weisbach, 1998, Endogenously chosen boards of directors and their monitoring of the CEO, American Economic Review 88, 96-118.

Hofer, C. W., 1980, Turnaround strategies, Journal of Business Strategy 1, 19-31.

Holmström, B. R., 1979, Moral hazard and observability, Bell Journal of Economics 10, 74-91.

Holmström, B. R., and S. N. Kaplan, 2001, Corporate covernance and merger activity in the United States: Making sense of the 1980s and 1990s, Journal of Economic Perspectives 15, 121-144.

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.

Innes, R. D., 1990, Limited liability and incentive contracting with ex-ante action choices, Journal of Economic Theory 52, 45-67.

Jensen, M. C., 1993, The modern industrial revolution, exit, and the failure of internal control systems, Journal of Finance 48, 831-880.

Jenter, D., 2002, Executive compensation, incentives, and risk, Mimeo, Massachussetts Institute of Technology.

Lambert, R. A., and D. F. Larcker, 1985, Golden parachutes, executive decision-making, and shareholder wealth, Journal of Accounting and Economics 7, 179-203.

Lefanowicz, C. E., J. R. Robinson, and R. Smith, 2000, Golden parachutes and managerial incentives in corporate acquisitions: Evidence from the 1980s and 1990s, Journal of Corporate Finance 6, 215-239.

Mehran, H., G. E. Nogler, and K. B. Schwartz, 1997, CEO incentive plans and corporate liquidation policy, Journal of Financial Economics 50, 319-349.

Milgrom, P. R., 1981, Good news and bad news: Representation theorems and applications, Bell Journal of Economics 12, 380-391.

Miller, D., and P. H. Friesen, 1980, Archetypes of organizational transition, Administrative Science Quarterly 25, 268-299.

Murhpy, K. J., and J. Zábojník, 2004, Managerial capital and the market for CEOs, Mimeo, University of Southern California.

Nachman, D. C., and T. H. Noe, 1994, Optimal design of securities under asymmetric information, Review of Financial Studies 7, 1-44.

Ross, S. A., 2004, Compensation, incentives, and the duality of risk aversion and riskiness, Journal of Finance 59, 207-225.

Rotemberg, J. J., and G. Saloner, 2000, Visionaries, managers, and strategic direction, Rand Journal of Economics 31, 693-716.

Rusticus, T. O., 2006, Executive severance agreements, Mimeo, Wharton School.

Schwab, S. J., and R. S. Thomas, 2004, What do CEOs bargain for? An empirical study of key legal components of CEO contracts, Mimeo, Cornell Law School.

Shapiro, C., and J. E. Stiglitz, 1984, Equilibrium unemployment as a worker discipline device, American Economic Review 75, 433-444.

Shleifer, A., and R. W. Vishny, 1989, Management entrenchment: The Case of manager-specific investments, Journal of Financial Economics 25, 123-139.

Smith C., and R. M. Stulz, 1985, The determinants of firms' hedging policies, Journal of Financial and Quantitative Analysis 20, 391-405.

Tushman, M. L., B. Virany, and E. Romanelli, 1985, Executive succession, strategic reorientations, and organizational evolution, Technology in Society 7, 297-313.

Tushman, M. L., W. H. Newman, and E. Romanelli, 1986, Convergence and upheaval: Managing the unsteady pace of organizational evolution, California Management Review 29, 29-44.

Walker, D. I., 2005, The manager's share, Mimeo, Boston University School of Law.

White, M. C., M. Smith, and T. Barnett, 1997, CEO succession: Overcoming forces of inertia, Human Relations 50, 805-828.

Wiersema, M. F., and K. A. Bantel, 1992, Top management team demography and corporate strategic change, Academy of Management Journal 35, 91-121.

Yellen, J. L., 1994, Efficiency wage models of unemployment, American Economic Review Papers and Proceedings 74, 200-205.