

CEO Replacement under Private Information*

Roman Inderst[†] Holger M. Mueller[‡]

October 2009

Abstract

This paper examines the optimal CEO compensation and replacement policy when the CEO is privately informed about the firm's continuation value under his leadership. Ex-ante moral hazard implies that the CEO must receive ex-post quasi rents, which endogenously biases him towards continuation. Our model shows that to induce "bad" CEOs to quit it may be best to make continuation costly (through steep incentive pay) rather than simply rewarding quitting (through severance pay). Incentive pay makes continuation attractive for "good" CEOs, who can expect high future on-the-job pay, but unattractive for "bad" CEOs, who may instead prefer their outside option payoff. Our model generates novel empirical implications that jointly relate CEO compensation and turnover to corporate governance, firm size, cash-flow risk, and the informativeness of performance measurement.

*We are grateful to Paolo Fulghieri (the editor), three anonymous referees, Andres Almazan, Dirk Jenter, Wei Jiang, Lasse Pedersen, Thomas Phillipon, Javier Suarez, David Yermack, and seminar participants at Stanford, Berkeley, Wharton, NYU, USC, LBS, LSE, CEMFI, HEC, the European Summer Symposium in Financial Markets in Gerzensee, and the NBER Corporate Finance Meeting in Cambridge for helpful comments and suggestions. This is a substantially revised version of an earlier draft entitled "Keeping the Board in the Dark: CEO Compensation and Entrenchment."

[†]Goethe University Frankfurt, CEPR, and ECGI. E-mail: inderst@finance.uni-frankfurt.de.

[‡]NYU Stern School of Business, NBER, CEPR, and ECGI. Email: hmueller@stern.nyu.edu.

1 Introduction

One of the basic tasks of any board is to replace “bad” CEOs. However, if CEOs are privately informed about their potential to create value for the firm going forward, then this task can be difficult: Replacement must be incentive compatible from the CEO’s perspective.¹ This paper shows that to induce “bad” CEOs to quit it may be best to make continuation costly (through steep incentive pay) rather than simply rewarding quitting (through severance pay). Incentive pay makes continuation attractive for “good” CEOs, who can expect high future on-the-job pay, but unattractive for “bad” CEOs, who may instead prefer their outside option payoff.

In our model, the CEO privately observes an interim signal that is informative about the quality of the match between him and the firm and thus about the likely firm value under his continued leadership.² A low signal may indicate a poor match, but it may also indicate that the CEO has shirked. Hence, while severance pay makes quitting attractive for “bad” CEOs—i.e., those who turn out to be a poor match—it also rewards failure, thus making it more costly to induce the CEO to exert effort. To preserve the CEO’s effort incentives, a dollar of severance pay must be matched by a dollar of expected on-the-job pay, thus raising the CEO’s overall expected pay by one dollar. As a result, each dollar of severance pay constitutes a dollar of rent for the CEO. In contrast, steep incentive pay—by tying the CEO’s continuation payoff to his (match-specific) “type” rather than simply rewarding quitting—does not leave the CEO any rents.

More generally, to induce the CEO to exert effort he must be promised a quasi rent (in the form of generous on-the-job pay) in case he continues. This endogenously biases the CEO towards continuation—in contrast to models in which continuation preferences derive from exogenously specified private benefits of control—and makes the problem of eliciting his private information non-trivial. Our model shows that it is best to award this quasi rent through steep incentive pay, not base pay. Incentive pay ensures that the CEO’s

¹The notion that the CEO has private information that the board needs to make an efficient replacement decision is not new. For instance, Jensen (1993, p. 864) notes that “the CEO most always determines the agenda and the information given to the board. This limitation on information hinders the ability of even highly talented board members to contribute effectively to the monitoring and evaluation of the CEO.”

²Our notion of match quality borrows from Hermalin and Weisbach (1998) and Hermalin (2005). In particular, Hermalin and Weisbach argue that what matters for the board’s replacement decision is not so much the CEO’s ability per se but the quality of the match with the firm. Empirical support for the hypothesis that the match quality affects CEO replacement is provided by Allgood and Farrell (2003).

expected on-the-job pay is high precisely when the match quality—and thus the firm value under his continued leadership—is high, while it is low precisely when the match quality is low, thereby aligning the CEO’s continuation preferences with those of the firm at the interim stage. Our theory of CEO incentive pay generates many empirical implications, some of which are orthogonal to those of classical agency theory (e.g., Holmström, 1979). Most notably, the optimal incentive pay in our model is steeper when the underlying performance measure is noisier or when the firm’s cash flows are riskier.

While severance pay is costly—it leaves the CEO valuable rents—it may nevertheless be part of the optimal contract. Precisely, this is the case when the CEO’s outside option payoff is sufficiently low so that even under steep incentive pay a “bad” CEO prefers to continue. In that case, the board faces a tradeoff between a more efficient replacement policy—i.e., one that induces “more” (out of a continuum of types) “bad” CEOs to quit—and minimizing the CEO’s rents. While the optimal contract may or may not feature severance pay, it is never optimal to set severance pay so high as to implement the first-best replacement policy. Hence, at the optimal solution some “bad” CEOs are not replaced, implying there is CEO entrenchment. One implication of our model is that a regulatory cap on severance pay—while it would limit CEO rent extraction—may lead to more entrenchment and lower firm value. Our model can also distinguish between severance pay that is agreed upon *ex ante* and “golden handshakes” that are additionally granted at the time of the CEO’s departure. Interestingly, if a regulatory cap applied only to “golden handshakes” then firms would benefit in our model as it would help them overcome a commitment problem *vis-à-vis* the CEO.

A key feature of our model is that incentive pay and severance pay must move together in equilibrium as otherwise there would be too high a reward for failure. This implication is different from Bebchuk and Fried’s (2004) rent-extraction argument, which states that severance pay is merely a substitute form of stealth compensation for (more visible) incentive pay. Another difference is that our model predicts that firms with weaker corporate governance should have more high-powered CEO incentive pay. In contrast, if CEOs could set their own pay, then we would expect that they primarily increase their base pay, not their incentive pay.

Our model generates novel predictions that jointly relate CEO turnover and incentive pay to observable primitives. Depending on the variation in question, CEO turnover and

incentive pay may or may not go hand in hand. For instance, CEO turnover and incentive pay move in opposite directions if the variation is with respect to primitives that directly affect the firm’s agency problem, such as corporate governance, cash-flow risk, and the informativeness of performance measurement. In contrast, CEO turnover and incentive pay move in the same direction if the variation is with respect to primitives that determine the efficiency of CEO replacement.

In an extension of our model, the board can observe a public signal that is informative about the match quality between the firm and the CEO and thus about the firm’s continuation value. The optimal CEO compensation and replacement policy depends on the public signal in the sense that more favorable signals are associated with lower CEO turnover but also with steeper incentive pay.

Our model is closely related to Levitt and Snyder (1997), who also analyze the tension between providing ex-ante effort incentives and incentives for truth-telling at the interim stage. They consider a similar principal-agent model where the agent must also first exert effort, then he receives private information, and then the principal must decide whether to continue the project. Like in our model the agent must be rewarded for coming forward with bad news, thus allowing the principal to stop the project. Apart from having a different focus, our model extends these insights by providing novel comparative statics results linking CEO turnover and incentive pay to corporate governance, firm size, cash-flow risk, and the informativeness of performance measurement.

Eisfeldt and Rampini (2008) also study a model of “information-based entrenchment.” In their model, a manager must be rewarded for coming forward with bad news about the productivity of assets under his control, thus allowing a more efficient asset re-allocation.³ This “reward” takes the form of a bonus, which plays a similar role as severance pay does in our model. Our model extends these insights by studying information-based entrenchment in more detail, focusing especially on the implications for on-the-job (incentive) pay (instead of severance pay) and on the interaction between on-the-job pay and severance pay, while providing novel comparative statics results.⁴

³Adams and Ferreira (2007) also study information exchange between the board and the CEO under conflicting preferences but do not consider either CEO compensation or replacement.

⁴In Eisfeldt and Rampini (2008), the manager’s desire to become entrenched arises from exogenous private benefits of control, whereas in our model the CEO’s desire to become entrenched arises endogenously from the promise of quasi rents in the form of generous on-the-job pay, which is necessary to induce him to exert effort. This naturally leads to a difference in focus since, unlike private benefits of control, the

Lambert and Larcker (1985), Knoeber (1986), and Harris (1990) are all models in which severance pay (or golden parachutes) mitigates managerial entrenchment. In contrast, in our model it is best to mitigate entrenchment through incentive pay, while severance pay may or may not be part of the optimal contract. More recently, Almazan and Suarez (2003) consider the role of severance pay for renegotiations between the CEO and the board. While severance pay may provide stronger ex-ante effort incentives than incentive pay does, CEO replacement is always (first-best) efficient in their model given that there is no asymmetric information between the CEO and the board.

The rest of this paper is organized as follows. Section 2 lays out the basic model. Section 3 sets up the board’s problem. Section 4 derives the optimal solution. Section 5 provides empirical implications. Section 6 considers an extension in which the board can observe a public signal that is informative about the CEO’s “type.” Section 7 distinguishes between severance pay that is agreed upon ex ante and “golden handshakes” that are additionally granted at the time of the CEO’s departure. Section 8 discusses regulatory caps on severance pay. Section 9 offers concluding remarks. All proofs are in Appendix A. Appendix B extends the optimal contract design to a continuum of firm values.

2 A Model of Endogenous CEO Entrenchment

There are four dates: $t = 0$, $t = 0.5$, $t = 1$, and $t = 2$. At $t = 0$ the firm hires a new CEO. While the CEO is the best available candidate at the time of the hiring, there is uncertainty as to whether he remains a good match for the firm going forward. This uncertainty is captured by the state of nature $\theta \in \Theta := [\underline{\theta}, \bar{\theta}]$, which is realized at some interim date $t = 1$, and which characterizes the match quality between the CEO and the firm. The state of nature is only observed by the CEO and thus, in particular, not by the firm’s board, which is assumed to act in the firm’s interest. Together with a simple problem of moral hazard to be introduced below, the CEO’s private information at the interim stage endogenously creates scope for entrenchment.⁵ The CEO is risk neutral and his utility from being employed elsewhere is $u_0 > 0$.

CEO’s on-the-job pay can be optimally designed by the firm’s board.

⁵One could also think of a richer strategy space for the CEO, e.g., one in which—after privately observing the state of nature—the CEO entrenches himself by undertaking an irreversible action that makes it prohibitively costly to replace him as in, e.g., Shleifer and Vishny (1989), Bagwell and Zechner (1993), and Fulghieri and Hodrick (2006).

Technology and Beliefs

The firm value under the CEO's continued leadership is denoted by $s \in S := \{\underline{s}, \bar{s}\}$. The probability of high firm value, $p(\theta) := \Pr(s = \bar{s} \mid \theta)$, depends on the match quality between the CEO and the firm and thus on the state of nature θ . We assume that $p(\theta)$ is continuously differentiable and strictly increasing in θ . The expected firm value conditional on θ is denoted by $v(\theta) := \underline{s} + p(\theta)\Delta_s$, where $\Delta_s := \bar{s} - \underline{s}$.

While the true firm value under the CEO's leadership may not be readily observable, the CEO's performance can be evaluated at $t = 2$ based on the performance measure $y \in \{\underline{y}, \bar{y}\}$. This measure is noisy in the sense that $y = \bar{y}$ coincides with $s = \bar{s}$ only with probability $\mu > 1/2$ while $y = \underline{y}$ coincides with $s = \underline{s}$ only with probability $\mu > 1/2$. The case where $\mu = 1$ is equivalent to the case where the true firm value is itself observable. An (inverse) measure of the informativeness of the performance measure is given by

$$\rho := \frac{1 - \mu}{2\mu - 1}, \quad (1)$$

which is zero if the performance measure is perfectly informative ($\mu = 1$) and which goes to infinity as the performance measure becomes completely uninformative ($\mu \rightarrow 1/2$). It should be noted that setting $\mu = 1$ does not cause any problems; all our results hold if the performance measure is perfectly informative.

Realistically, the firm value under the CEO's continued leadership also depends on how dedicated the CEO is to his job. At $t = 0.5$, after the CEO is hired but before the state of nature θ is realized, the CEO can either work hard or shirk. If the CEO shirks then $\theta = \underline{\theta}$ is realized with certainty at $t = 1$. Shirking yields non-pecuniary private benefits $B > 0$. In contrast, if the CEO works hard then θ is realized with density $f(\theta) > 0$ for all $\theta \in \Theta$ and distribution function $F(\theta)$. Let $\hat{p} := \int_{\underline{\theta}}^{\bar{\theta}} p(\theta)f(\theta)d\theta$ denote the expected probability of high firm value if the CEO works hard. As we will explain below, the sole purpose of this simple moral hazard problem is to endogenize the CEO's continuation preferences at the interim stage.

CEO Replacement and Compensation

At $t = 1$, after the CEO privately observes the state of nature, there is the possibility of (early) CEO replacement. While the CEO realistically has private information about the (match-specific) firm value under his continued leadership, $v(\theta)$, we assume that the firm value under a potential replacement CEO, V , is common knowledge. The latter is

gross of the cost u_0 of hiring a replacement CEO. Accordingly, it is first-best optimal to replace the incumbent CEO if and only if $V > v(\theta)$.⁶ For brevity we restrict consideration to the case where $p(\underline{\theta}) < (V - \underline{s})/\Delta_s < p(\bar{\theta})$, which together with continuity and strict monotonicity of $p(\theta)$ implies the existence of a unique interior cutoff value $\theta_{FB} \in (\underline{\theta}, \bar{\theta})$ satisfying

$$p(\theta_{FB}) = \frac{V - \underline{s}}{\Delta_s}, \quad (2)$$

such that it is first-best optimal to replace the CEO if and only if $\theta < \theta_{FB}$.

As the CEO privately observes the state of nature, the question of whether the first best is attainable depends on the compensation contract. As we will argue in more detail below, we can without loss of generality restrict attention to simple contracts specifying a single on-the-job pay scheme $w(y)$ and a single severance payment W if the CEO is replaced.⁷ Note that we do not impose any financial constraints on the part of the firm as to what it can pay the CEO.⁸ Also, as the performance measure y can only take on two possible values it is convenient to write $\underline{w} := w(\underline{y})$ and $\bar{w} := w(\bar{y})$, where $\Delta_w := \bar{w} - \underline{w}$ denotes the steepness (“incentive slope”) of the CEO’s on-the-job pay scheme. In a slight abuse of terminology we will sometimes refer to Δ_w simply as the CEO’s “incentive pay.” Given this definition, we can write the CEO’s expected on-the-job pay (conditional on θ) as $\hat{w}(\theta) := \underline{w} + \pi(\theta)\Delta_w$, where $\pi(\theta) := p(\theta)\mu + (1 - p(\theta))(1 - \mu)$ represents the total probability with which the performance measure takes on the value $y = \bar{y}$. We impose a standard limited liability constraint requiring that payments to the CEO cannot be negative. As it will always hold that $\Delta_w \geq 0$ and $W \geq 0$ at the optimal solution, it is sufficient to impose this requirement only for the CEO’s base wage, i.e., $\underline{w} \geq 0$.

Discussion

Our theory of CEO compensation and replacement relies crucially on the *interaction* between the CEO’s moral hazard problem and his private information problem at the interim stage. In particular, the optimal CEO on-the-job pay scheme is not driven by

⁶The total first-best payoff if the CEO is replaced is given by the sum of $V - u_0$ (for the firm) and u_0 (for the CEO, who realizes u_0 if he is replaced). Note that as θ only provides information about the match quality with the CEO’s *current* employer, a low value of θ does not compromise the CEO’s outside option payoff from being employed elsewhere.

⁷Precisely, within the class of *deterministic* mechanisms we can ignore menus that condition payments to the CEO directly on the (truthfully revealed) state of nature θ , i.e., $w(\theta, y)$ and $W(\theta)$.

⁸In this regard, note that as long as the performance measure is noisy ($\mu < 1$) it may be the case that $y = \bar{y}$ even though $s = \underline{s}$.

either one of the two incentive problems alone but by their interaction. If only one of the two incentive problems were present the first best could be trivially attained with a simple fixed-wage contract. To illustrate this, consider first the CEO's moral hazard problem in isolation. Specifically, suppose θ is verifiable so that it is (trivially) optimal to implement the first-best replacement policy $\theta^* = \theta_{FB}$. Absent any private information problem at the interim stage it is straightforward to show that severance pay is strictly suboptimal in this case, implying that $W = 0$. As the CEO is replaced for sure if he shirks—given that shirking results in $\theta = \underline{\theta} < \theta_{FB}$ —his payoff from shirking is $u_0 + B$. The CEO thus refrains from shirking if and only if

$$\int_{\theta_{FB}}^{\bar{\theta}} \hat{w}(\theta) f(\theta) d\theta + F(\theta_{FB}) u_0 \geq u_0 + B,$$

which can be easily satisfied, for instance, with a fixed-wage contract

$$w = u_0 + \frac{B}{1 - F(\theta_{FB})} \tag{3}$$

that just compensates the CEO for his shirking benefits B . Note that while this particular contract is optimal if θ is verifiable it is not uniquely optimal. In fact, there exists an infinite number of optimal compensation schemes in this case.

Consider next the CEO's private information problem in isolation. Suppose the CEO always works hard so that the only problem is to make him indifferent between lying and truthtelling at the interim stage. It is then again optimal to implement the first-best replacement policy θ_{FB} , which can be (trivially) accomplished with a fixed-wage contract $w = u_0$ that makes the CEO indifferent between replacement and continuation for all θ . In contrast, if there is an additional moral hazard problem the CEO's truthtelling problem at the interim stage becomes non-trivial: To induce the CEO to work hard he must be promised a quasi rent if he continues, which biases him towards continuation and rules out the optimality of fixed-wage contracts of the sort described above.

3 The Board's Problem

To keep the analysis simple we shall initially *assume* that the board wants to induce the CEO to work hard. We will later derive a sufficient condition for when this is indeed optimal.

CEO Replacement

If the CEO is replaced he receives u_0 from being employed elsewhere plus (possibly) severance pay W .⁹ In contrast, if the CEO continues his expected on-the-job pay (conditional on θ) is $\widehat{w}(\theta)$, implying that replacement is incentive compatible from the CEO's perspective if and only if

$$\widehat{w}(\theta) \leq W + u_0. \quad (4)$$

As will be clear from the analysis below, at the optimal solution it must hold that $\Delta_w > 0$, implying that $\widehat{w}(\theta)$ must be strictly increasing in θ . Given that $\widehat{w}(\theta)$ is continuous in θ , condition (4) thus gives rise to a unique cutoff value θ^* such that the CEO is replaced if and only if $\theta < \theta^*$. In case of an interior solution this cutoff value is uniquely characterized by the CEO's indifference condition

$$\widehat{w}(\theta^*) = W + u_0, \quad (5)$$

which can be written as

$$\pi(\theta^*) = \frac{u_0 - \underline{w} + W}{\Delta_w}. \quad (6)$$

Comparing the second-best cutoff value θ^* from (6) with the first-best cutoff value θ_{FB} from (2) will be key to our analysis. Note that—as any particular realization of θ is a zero-probability event (that is, on the equilibrium path where the CEO works hard)—we can capture the boundary cases where the CEO is either always or never replaced by setting $\theta^* = \bar{\theta}$ and $\theta^* = \underline{\theta}$, respectively.

Moral Hazard

In case of an interior solution $\theta^* \in (\underline{\theta}, \bar{\theta})$ we obtain that the CEO works hard if

$$\int_{\theta^*}^{\bar{\theta}} \widehat{w}(\theta) f(\theta) d\theta + F(\theta^*)(u_0 + W) \geq u_0 + W + B, \quad (7)$$

which can be written as

$$\int_{\theta^*}^{\bar{\theta}} (\widehat{w}(\theta) - W - u_0) f(\theta) d\theta \geq B. \quad (8)$$

⁹For brevity, we refrain from setting up a more general mechanism design problem. As we will explain below, the solution provided here is optimal within the class of *deterministic* mechanisms, i.e., mechanisms that deterministically specify a compensation scheme and replacement policy for any given “message” $\hat{\theta}$ sent by the CEO.

Condition (8), which we henceforth refer to as the “non-shirking” condition, is intuitive. It states that to induce the CEO to work hard there must be a sufficient wedge between his expected on-the-job pay $\widehat{w}(\theta)$ and his payoff from being replaced, $W + u_0$.

It is straightforward to show that (8) must bind. Inserting Δ_w from (6) into the binding constraint (8) we obtain for an interior solution $\theta^* \in (\underline{\theta}, \bar{\theta})$ the requirement that

$$\int_{\theta^*}^{\bar{\theta}} \left[\frac{p(\theta) - p(\theta^*)}{\rho + p(\theta^*)} \right] f(\theta) d\theta = \frac{B}{u_0 - \underline{w} + W}, \quad (9)$$

where the left-hand side is strictly decreasing and continuous in θ^* . Hence, if a solution to (9) exists then it is unique. Moreover, the left-hand side of (9) goes to zero as $\theta^* \rightarrow \bar{\theta}$ while the right-hand side is strictly positive by (6) and $\Delta_w > 0$. Consequently, for an interior solution to exist it must hold that

$$\frac{\widehat{p} - p(\underline{\theta})}{\rho + p(\underline{\theta})} > \frac{B}{u_0 - \underline{w} + W}, \quad (10)$$

where $\widehat{p} := \int_{\underline{\theta}}^{\bar{\theta}} p(\theta) f(\theta) d\theta$. To ensure that an interior solution exists we assume that (10) holds for $\underline{w} = 0$ and $W = 0$, in which case it also holds for all $W > 0$.

Consider next the boundary cases $\theta^* = \bar{\theta}$ and $\theta^* = \underline{\theta}$. A solution where $\theta^* = \bar{\theta}$ (CEO is always replaced) cannot be optimal as the CEO would then always shirk. In contrast, a solution where $\theta^* = \underline{\theta}$ (CEO is never replaced) implies that the CEO is not even replaced if he shirks. In this (uninteresting) case the board’s problem reduces to a standard moral hazard problem as the state of nature θ is irrelevant. For brevity we shall initially assume that it is optimal to implement a replacement policy $\theta^* > \underline{\theta}$ under which the CEO is replaced with positive probability. (Condition (10) only ensures that such a replacement policy is feasible, not that it is optimal.) We will later provide a sufficient condition for when this is indeed optimal.

Board’s Objective Function

Firm profits are given by

$$\Pi := \int_{\theta^*}^{\bar{\theta}} [v(\theta) - \widehat{w}(\theta)] f(\theta) d\theta + F(\theta^*) (V - u_0 - W). \quad (11)$$

Inserting the binding “non-shirking” constraint (8) into (11) we obtain

$$\Pi := \int_{\theta^*}^{\bar{\theta}} v(\theta) f(\theta) d\theta + F(\theta^*) V - (u_0 + B + W). \quad (12)$$

Accordingly, the board maximizes (12) subject to the CEO’s indifference condition (6), which states that θ^* must be incentive compatible from the CEO’s perspective, and the non-negativity constraint $\underline{w} \geq 0$.

Inspection of (12) yields two important insights. First, given that the “non-shirking” constraint must bind the firm becomes the residual claimant. The board will consequently want to make a replacement decision that is as efficient as possible, i.e., implement a cutoff value θ^* that is as close as possible to θ_{FB} . (Holding W fixed and without imposing (6) and $\underline{w} \geq 0$ the board’s objective function (12) is maximized at $\theta^* = \theta_{FB}$.) Second, a dollar of severance pay constitutes a dollar of rent for the CEO, which reduces firm profits. (The partial derivative of (12) with respect to W is equal to one.)¹⁰ Intuitively, this is because the CEO receives severance pay not only if he works hard and the match quality is low, but also if he shirks given that shirking results in his sure replacement.

Severance Pay and CEO Replacement

The “non-shirking” condition (8) endogenously creates scope for entrenchment as it implies that the CEO must receive an ex-post quasi rent if he continues. Below we will show under what conditions there is indeed entrenchment in equilibrium. Before doing that, however, we will show that severance pay alone does not guarantee a reduction in entrenchment. In fact, it may have absolutely no effect on the cutoff value θ^* (though it always increases the CEO’s rents) unless it is accompanied by a simultaneous increase in the CEO’s incentive pay.

In our model the CEO’s severance pay W and his expected on-the-job pay $\widehat{w}(\theta)$ are tightly linked through the “non-shirking” condition (8). If the board were to raise W without also raising $\widehat{w}(\theta)$ then this constraint would be violated and the CEO would prefer to shirk. Hence, a comparative statics analysis of the CEO’s indifference condition (6) in W alone (to gauge the effect of an increase in W on θ^*) while holding the CEO’s on-the-job pay fixed is illegitimate. However, if the board must accompany an increase in severance pay by a simultaneous increase in the CEO’s on-the-job pay to prevent him from shirking, then it becomes relatively *more* attractive for the CEO to continue, running counter to the goal of reducing entrenchment. As it turns out, whether or not severance pay reduces

¹⁰Another way to illustrate that a dollar of severance pay constitutes a dollar of rent for the CEO is by looking at the (untransformed) “non-shirking” constraint (7). As this constraint binds, W is equal to the difference between the CEO’s ex-ante expected payoff from being employed (left-hand side) and the sum of his outside option payoff and his private benefits from shirking, $u_0 + B$.

entrenchment depends on how the accompanying increase in the CEO’s on-the-job pay is structured. For instance, suppose the board increases the CEO’s severance pay and his on-the-job pay in a way that leaves his incentive pay Δ_w unchanged. This implies that, to satisfy the (binding) “non-shirking” constraint (8), the board must increase W , \underline{w} , and \bar{w} all by the same amount. But it then follows immediately from the CEO’s indifference condition (6) that the cutoff value θ^* remains unchanged, implying that entrenchment is not reduced. On the other hand, the firm gets the biggest “bang for the buck”—i.e., the biggest reduction in entrenchment for a given increase in severance pay—if the increase in severance pay is accompanied by an increase in the CEO’s incentive pay Δ_w while the base wage \underline{w} remains unchanged (implying W and \bar{w} increase but not \underline{w}). This ensures that the CEO’s expected on-the-job pay $\widehat{w}(\theta)$ increases primarily at high values of θ , implying that at low values of θ it increases by less than W , thus raising the cutoff value θ^* and reducing entrenchment. We thus have:

Lemma 1 *The firm’s board can implement any desired cutoff value θ^* by increasing the CEO’s severance pay W and simultaneously increasing his incentive pay Δ_w while leaving the base wage \underline{w} unchanged. Formally, implicit differentiation of (9) shows that at any interior solution $\theta^* \in (\underline{\theta}, \bar{\theta})$ the marginal effect of an increase in W on θ^* , holding \underline{w} fixed, is given by*

$$\frac{d\theta^*}{dW} = \frac{1}{p'(\theta^*)} \frac{B}{(u_0 - \underline{w} + W)^2} \frac{[\rho + p(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} [\rho + p(\theta)] f(\theta) d\theta} > 0.$$

4 Optimal CEO Compensation and Replacement

The CEO’s base wage \underline{w} and his incentive pay Δ_w are jointly determined by the CEO’s indifference condition (6) and the binding “non-shirking” constraint (8). Inserting (6) into (8) and noting that $\pi(\theta) = p(\theta)(2\mu - 1) + 1 - \mu$ and $1 - \mu = \rho(2\mu - 1)$ we obtain

$$\Delta_w = \frac{1}{2\mu - 1} \frac{B}{\int_{\theta^*}^{\bar{\theta}} [p(\theta) - p(\theta^*)] f(\theta) d\theta} \quad (13)$$

and

$$\underline{w} = W + u_0 - \frac{B[\rho + p(\theta^*)]}{\int_{\theta^*}^{\bar{\theta}} [p(\theta) - p(\theta^*)] f(\theta) d\theta}. \quad (14)$$

This characterization holds for a *given* cutoff value θ^* and a *given* severance payment W . Eventually both θ^* and W will be jointly determined in equilibrium. For this we must

distinguish between different cases depending on whether the first-best replacement policy $\theta^* = \theta_{FB}$ can be implemented without severance pay—in which case it will indeed be optimal to implement $\theta^* = \theta_{FB}$ —or whether the first-best replacement policy cannot be implemented without severance pay—in which case it will be optimal to implement an inefficient replacement policy $\theta^* < \theta_{FB}$.

First Best

We first characterize under what condition the first-best replacement policy can be implemented without severance pay, and thus without leaving the CEO any rents. To characterize the optimal solution with $\theta^* = \theta_{FB}$ and $W = 0$ it is convenient to define

$$\Delta V_{FB} := \int_{\theta_{FB}}^{\bar{\theta}} (v(\theta) - V)f(\theta)d\theta, \quad (15)$$

which represents the value created under the incumbent CEO (versus a potential replacement CEO) provided he is replaced efficiently. (Recall from Section 2 that $v(\theta) > V$ if and only if $\theta > \theta_{FB}$.) Inserting $\theta^* = \theta_{FB}$ and $W = 0$ into (13) and (14) and noting that $v(\theta_{FB}) = V$ (from condition (2)) and $1 - \mu = \rho(2\mu - 1)$ we obtain

$$\Delta_w = \frac{B}{\Delta V_{FB}} \frac{\Delta_s}{2\mu - 1} \quad (16)$$

and

$$\underline{w} = u_0 - \frac{B}{\Delta V_{FB}} \Delta_s (p(\theta_{FB}) + \rho). \quad (17)$$

The characterization of the CEO's incentive pay in (16) ensures—besides inducing him to work hard—that the CEO's and the firm's interests are perfectly aligned at the replacement stage. It links the steepness of the CEO's incentive pay, Δ_w , to the value created under the (incumbent) CEO, ΔV_{FB} , his shirking benefits, B , the firm's own “upside potential,” Δ_s , and the informativeness of the performance measure, μ . If we ignore the latter for a moment by setting $\mu = 1$, equation (16) simplifies to

$$\frac{\Delta_w}{\Delta_s} = \frac{B}{\Delta V_{FB}}.$$

Hence, for the CEO's and the firm's interests to be perfectly aligned at the replacement stage the relation between the steepness of the CEO's incentive pay and the firm's own “upside potential” must mirror the relation between the CEO's cost of (non-)shirking and the value created under the CEO.

Implementing the first-best replacement policy $\theta^* = \theta_{FB}$ without severance pay may require steep incentive pay. Equation (16) shows that the required incentive pay Δ_w is steeper the larger is the CEO's cost of (non-)shirking relative to the value created under the CEO, $B/\Delta V_{FB}$, the larger is the firm's own "upside potential," Δ_s , and the noisier is the performance measure (i.e., the smaller is μ). If the required incentive pay is too steep it may no longer be feasible to implement the first-best replacement policy without severance pay. This can be seen from inserting (16) into (17), which yields

$$\underline{w} = u_0 - \Delta_w(2\mu - 1)(p(\theta_{FB}) + \rho).$$

By inspection, if Δ_w becomes too large the non-negativity constraint $\underline{w} \geq 0$ is violated (recall that $\mu > 1/2$), which implies the compensation scheme in (16)-(17) is no longer feasible. Accordingly, the first-best replacement policy can be implemented without severance pay—and thus without leaving the CEO any rents—if and only if

$$\frac{\Delta_s B}{\Delta V_{FB}}(p(\theta_{FB}) + \rho) \leq u_0. \quad (18)$$

Condition (18), which is key to our analysis, has a natural interpretation. We have already shown that this condition is more difficult to satisfy the steeper is the incentive pay $\Delta_w = \Delta_s B/\Delta V_{FB}$ that is required to implement the first-best replacement policy without severance pay. By inspection, condition (18) is also more difficult to satisfy the larger is $p(\theta_{FB})$ (and thus θ_{FB}) and the smaller is u_0 . A larger first-best cutoff value θ_{FB} implies that there are more states of nature in which the CEO must be induced to give up his (future) quasi rents. In the absence of severance pay the only reason why the CEO would ever be willing to give up his quasi rents is to realize his outside option payoff u_0 from being employed elsewhere. Hence, if either θ_{FB} is too large or u_0 is too small it will not be possible to implement the first-best replacement policy without severance pay. Finally, condition (18) is also more difficult to satisfy the larger is ρ , i.e., the less informative is the performance measure (cf., equation (1)).

Second Best

In a first-best world where condition (18) holds there is no tradeoff: The board can always implement the first-best replacement policy $\theta^* = \theta_{FB}$ without leaving the CEO any rents. However, if condition (18) does not hold the board faces a tradeoff between

implementing a replacement policy that is as efficient as possible—i.e., implementing a cutoff value θ^* that is as close as possible to θ_{FB} —and minimizing the CEO’s rents.

We first derive a result that is intuitive in light of our above discussion: If condition (18) does not hold then it is optimal to set the CEO’s base wage as low as possible (i.e., $\underline{w} = 0$) as otherwise firm profits could be increased by lowering the CEO’s base wage and offering him instead steeper incentive pay (i.e., larger Δ_w). Offering the CEO steeper incentive pay shifts more of his expected on-the-job pay $\widehat{w}(\theta)$ from low into high states of nature θ , thus making it less attractive for him to continue in low states and permitting the board to implement a higher cutoff value θ^* with a given amount of severance pay or, likewise, implement a given cutoff value with less severance pay and thus less rents for the CEO. Either way firm profits are higher.

Lemma 2 *If condition (18) does not hold, so that it is not possible to implement the first-best cutoff value $\theta^* = \theta_{FB}$ without severance pay, then it is optimal to set the CEO’s base wage as low as possible, i.e., $\underline{w} = 0$.*

Proof. See Appendix.

Setting $\underline{w} = 0$ in equation (14) yields a uniquely optimal second-best cutoff value $\theta^* \in (\underline{\theta}, \bar{\theta})$ for a given choice of severance pay. Inserting θ^* into equation (13) in turn yields a uniquely optimal choice of incentive pay Δ_w . Thus, the only unknown remains the optimal choice of severance pay W .

By Lemma 1 the board can always implement a higher cutoff value by increasing the CEO’s severance pay. Hence, even if condition (18) does not hold the board could in principle still implement the first-best replacement policy $\theta^* = \theta_{FB}$ by offering the CEO sufficiently large severance pay. However, this will not be optimal. Totally differentiating the board’s objective function (12) we obtain

$$\frac{d\Pi}{dW} = f(\theta^*) [V - v(\theta^*)] \frac{d\theta^*}{dW} - 1, \quad (19)$$

where $d\theta^*/dW > 0$ is given by Lemma 1 (with $\underline{w} = 0$). Since $V \leq v(\theta)$ if and only if $\theta \geq \theta_{FB}$ this implies that $d\Pi/dW < 0$ for all $\theta^* \geq \theta_{FB}$, which in turn implies that $\theta^* < \theta_{FB}$. Thus, even though the board could in principle still implement the first-best replacement policy by offering the CEO sufficiently large severance pay, doing so will not

be optimal as it would require leaving him too much rents. Naturally it will then also not be optimal to implement a higher cutoff value $\theta^* > \theta_{FB}$, which is less efficient and would require leaving the CEO even more rents. Thus, in the second-best case where condition (18) does not hold the optimal replacement policy is inefficient—precisely, it is too lenient—and there is entrenchment in equilibrium.

Besides showing that $\theta^* < \theta_{FB}$ we can derive a lower bound for the optimal cutoff value θ^* . By Lemma 1 the optimal cutoff value is strictly increasing in W . Therefore, the lowest optimal cutoff value is the value of θ^* that solves equation (14) for $W = 0$ and $\underline{w} = 0$ (by Lemma 2). Denote this cutoff value by θ_0^* , where $\theta_0^* > \underline{\theta}$ follows from our assumption that condition (10) holds for $\underline{w} = 0$ and $W = 0$. If $\theta^* = \theta_0^*$ is chosen the CEO receives no severance pay and thus no rents. At the same time, CEO turnover is lower—and CEO entrenchment is higher—than under any higher optimal cutoff value $\theta^* > \theta_0^*$. For example, it is optimal to choose $\theta^* = \theta_0^*$ if the derivative $d\Pi/dW$ in (19) is negative for all $W > 0$ (and therefore for all $\theta^* > \theta_0^*$).

The following proposition summarizes our results.

Proposition 1 *The optimal CEO compensation and replacement policy is as follows:*

- i) If condition (18) holds the first-best replacement policy $\theta^* = \theta_{FB}$ can be implemented without severance pay and thus without leaving the CEO any rents.*
- ii) If condition (18) does not hold the board faces a tradeoff between implementing a replacement policy that is as efficient as possible and minimizing the CEO's rents. The second-best optimal cutoff value satisfies $\theta^* \in [\theta_0^*, \theta_{FB})$, where $\theta^* = \theta_0^* > \underline{\theta}$ implies no rents for the CEO while any higher optimal cutoff value $\theta^* > \theta_0^*$ implies a more efficient replacement policy—and thus less entrenchment—but also positive rents for the CEO. It is never optimal to implement the first-best replacement policy $\theta^* = \theta_{FB}$, implying there is always entrenchment in equilibrium.*

In Case i) of Proposition 1 the uniquely optimal choice of severance pay is $W = 0$ while the uniquely optimal choice of incentive pay Δ_w and base wage \underline{w} is determined by equations (16) and (17), respectively. In Case ii) the uniquely optimal base wage is $\underline{w} = 0$ (see Lemma 2) while the uniquely optimal choice of incentive pay Δ_w is determined by equation (13) after inserting the uniquely optimal cutoff value θ^* , which in turn is determined by equation (14) (with $\underline{w} = 0$) for a given optimal choice of severance pay

W . If we *assume* that the board’s objective function (12) is strictly quasiconcave in W we can also uniquely determine the optimal choice of severance pay. In that case, the optimal solution is $W = 0$ (implying that $\theta^* = \theta_0^*$) if and only if the derivative $d\Pi/dW$ in (19) (where $d\theta^*/dW$ is given by Lemma 1 with $\underline{w} = 0$) is nonpositive at $W = 0$ and thus at $\theta^* = \theta_0^*$. Otherwise the optimal severance pay is positive (implying that $\theta^* > \theta_0^*$) and uniquely determined by the first-order condition $d\Pi/dW = 0$ in (19) together with θ^* from equation (14) with $\underline{w} = 0$.

It remains to provide a sufficient condition under which it is optimal to induce the CEO to work hard. For this we shall derive a lower bound on firm profits in case the CEO works hard by considering a replacement policy under which the CEO is never replaced, i.e., $\theta^* = \underline{\theta}$. A sufficient condition ensuring that it is optimal to induce the CEO to work hard is then that this lower bound on firm profits exceeds firm profits in case the CEO shirks. If the CEO is never replaced the board’s problem reduces to a standard moral hazard problem. The CEO then works hard if $(2\mu - 1)\hat{p}\Delta_w \geq B$ while condition (10) (with $W = 0$) implies that the CEO receives no rents in this moral hazard problem alone, which in turn implies that firm profits are given by $\underline{s} + \hat{p}\Delta_s - B - u_0$. Given that firm profits if the CEO shirks are $V - u_0$, a sufficient condition ensuring that it is optimal to induce the CEO to work hard is

$$\underline{s} + \hat{p}\Delta_s \geq V + B. \tag{20}$$

Discussion of Optimal Contract Design

In our model incentive pay is optimal because it ensures that the CEO’s expected on-the-job pay is low precisely when the expected firm value under his continued leadership is low, thus making it unattractive for him to continue in states where replacement is efficient. This argument is different from those in standard moral hazard models, such as Innes (1990). While in our model the CEO must also be induced to work hard, the “non-shirking” condition (8) does not determine the functional form of the CEO’s on-the-job pay. Rather, it only determines the “wedge” between the CEO’s *expected* on-the-job pay and his payoff from being replaced. In fact, as we have emphasized before, if the only problem was moral hazard (e.g., because the state of nature θ is verifiable) then there would exist infinitely many optimal on-the-job pay schemes. Instead, the optimal on-the-job pay scheme in our model derives from the interaction between the CEO’s ex-ante moral

hazard problem and his private information problem at the interim stage.¹¹

We have restricted attention to simple mechanisms that specify a single on-the-job pay scheme $w(s)$ and a fixed severance payment W . Within the class of deterministic mechanisms this restriction is without loss of generality. For instance, it is obvious that the CEO must receive the same severance pay $W(\theta) = W$ in all states $\theta < \theta^*$ where he is replaced. That said, it is straightforward to design an incentive-compatible menu of on-the-job pay schemes such that the CEO prefers different on-the-job pay schemes in different continuation states $\theta \geq \theta^*$. However, given that the choice between continuation and replacement is only binary, there is no benefit from offering such a menu here. Put differently, any finer partitioning of the “replacement region” $\theta < \theta^*$ or the “continuation region” $\theta \geq \theta^*$ is of no value.¹² On the other hand, our restriction to deterministic mechanisms is not without loss of generality. As has already been noted by Levitt and Snyder (1997, footnote 10), who consider a similar setting as we do, if the firm is financially unconstrained so that it can make arbitrarily large payments then in this type of setting a stochastic mechanism can attain the first best in the limit.

5 Comparative Statics Analysis

Proposition 1 has three possible outcomes. Either the first-best cutoff value $\theta^* = \theta_{FB}$ can be implemented without severance pay or the optimal cutoff value satisfies $\theta^* < \theta_{FB}$ while severance pay is either zero or positive. When we analyze in what follows how a change in primitives affects CEO turnover (i.e., the optimal cutoff value θ^*) our results always hold weakly in all cases but only strictly in some cases. This is most obvious when the first-best cutoff value $\theta^* = \theta_{FB}$ is obtained both before and after the change in primitives. We assume that the board’s objective function (12) is strictly quasiconcave in W to ensure that all contractual variables, including the optimal cutoff value θ^* , change continuously

¹¹In Appendix B, where we extend our model to a continuum of firm values, we will further elaborate on how the assumptions that we must then invoke (in particular, the Monotone Likelihood Ratio Property) link our paper also technically to seminal papers of optimal contract design, such as Innes (1990) and Nachman and Noe (1994).

¹²It can be shown that in the second-best case a (non-degenerate) menu is even *strictly suboptimal* as it makes the CEO’s expected on-the-job pay “flatter” (as a function of θ) relative to the case where a single contract is being offered. See also Levitt and Snyder (1997), who consider a similar setting as we do and who also reach the conclusion that both the payoff scheme under continuation and the payment to the agent under cancellation are optimally made independent of the agent’s interim message (see Proposition 7 and especially Step 2 of the corresponding proof on p. 659).

as a function of the model's primitives.

Proposition 2 *The model predicts more high-powered CEO incentive pay (higher Δ_w) and lower CEO turnover (lower θ^*), and thus more entrenchment, if:*

- i) the CEO's shirking benefits are higher (higher B);*
- ii) the firm's cash flows are riskier (mean-preserving spread in s);*
- iii) the performance measure is noisier (lower μ).*

Proof. See Appendix.

Corporate Governance

Case i) of Proposition 2 considers an increase in the CEO's shirking benefits B . Remember that the shirking problem is important as it creates a conflict of interest between the CEO and the firm at the interim stage. Without this shirking problem only a single agency problem would remain, namely, that of eliciting the CEO's private information at the interim stage. As we explained earlier, absent any conflict of interest at this stage the truth-telling problem can be trivially resolved.

The intuition for Case i) is straightforward. As the CEO's shirking benefits increase he must be promised a higher ex-post quasi rent to prevent him from shirking. This biases him more towards continuation, thus putting downward pressure on the optimal cutoff value θ^* . The optimal way to promise the CEO a higher ex-post quasi rent is through steeper incentive pay (higher Δ_w) as this minimizes the increase in expected on-the-job pay in low states of nature where the match value is relatively low. If condition (18) holds after the increase in B (in which case it also holds before) then the first-best optimal replacement policy $\theta^* = \theta_{FB}$ can be preserved. In contrast, if condition (18) does not hold, which is more likely the higher is B , then an increase in B leads to a decrease in θ^* and thus to more entrenchment.

Empirically we might associate higher values of B with weaker corporate governance. Our result then predicts that firms with weaker corporate governance should have lower CEO turnover and steeper CEO incentive pay. Consistent with this prediction several empirical studies find a positive relation between CEO turnover and various measures of corporate governance, such as outside blockholders (Denis, Denis, and Sarin, 1997),

takeover pressure (Hadlock and Lumer, 1997; Mikkelsen and Partch, 1997), and the G-index (Fisman, Khurana, and Rhodes-Kropf, 2005).¹³ Moreover, and also consistent with our result, Fahlenbrach (2009) finds a negative relation between CEO pay-for-performance sensitivity and various corporate governance measures, such as board independence and institutional ownership, while Bertrand and Mullainathan (1999) find that CEO pay-for-performance sensitivity increased following the passage of state anti-takeover legislation in the U.S., which is widely believed to have weakened corporate governance.

Note that our prediction differs from Hermalin (2005), who argues that stronger corporate governance (in the form of more diligent board monitoring) results in higher CEO compensation because CEOs must be compensated for the increased disutility from being monitored. In contrast, in our model stronger corporate governance leads to lower ex-post quasi rents for the CEO and thus lower CEO compensation.

Risk

Standard agency theory predicts a negative relation between risk and incentives. In contrast, Case ii) of Proposition 2 predicts a positive relation. Intuitively, holding the CEO's compensation scheme fixed while increasing cash-flow risk in the sense of a mean-preserving spread (MPS) in firm value, s , the “non-shirking” condition (8) becomes violated. Similar to above the CEO must then be promised a higher ex-post quasi rent, which is optimally done through steeper incentive pay (higher Δ_w), resulting in a decrease in the optimal cutoff value θ^* and thus to more entrenchment.

In an influential paper, Prendergast (2002) notes that the empirical evidence regarding the relation between risk and incentives is mixed. While some empirical studies find a negative relation as predicted by standard agency theory (e.g., Lambert and Larcker, 1987; Aggarwal and Samwick, 1999a), others find a positive relation as predicted by both our model and Prendergast's model (e.g., Demsetz and Lehn, 1985; Core and Guay, 1999; Oyer and Shaefer; 2005).¹⁴ As for the relation between risk and CEO turnover, the empirical evidence is scarce. A notable exception is a recent study by Bushman, Dai, and Wang (2009), who find that CEO turnover is positively related to the idiosyncratic portion of a firm's stock return volatility but negatively related to risk that is beyond the CEO's control

¹³Also consistent with our result is the finding by Huson, Parrino, and Starks (2002) that improvements in corporate governance over the past decades have led to a significant increase in forced CEO turnover.

¹⁴Yet other studies find no statistically significant relation between risk and incentives (e.g., Garen, 1994; Bushman, Indjejikian, and Smith, 1996).

as measured by the portion of a firm’s stock return volatility that is due to market- and industry-wide returns.

Informativeness of Performance Measurement

Case iii) of Proposition 2 argues that less informative performance measurement (lower μ) leads to steeper CEO incentive pay and more entrenchment. Intuitively, for a given steepness of the CEO’s incentive pay, Δ_w , a decrease in the informativeness of the performance measure, μ , makes the CEO’s expected on-the-job pay $\hat{w}(\theta)$ “flatter”. To keep the CEO’s interests aligned with those of the firm at the interim stage, it is thus necessary to notch up the steepness of his incentive pay. This can be most easily seen in the characterization of the first-best incentive slope in (16), which states that $\Delta_w(2\mu - 1)$ must equal $\Delta_s B / \Delta V_{FB}$. Hence, as μ decreases the incentive slope Δ_w must increase. If μ becomes too small, so that the first best is no longer attainable, then a decrease in μ additionally leads to a decrease in the optimal cutoff value θ^* and thus to more entrenchment.

Appealing to standard agency theory (e.g., Holmström, 1979), DeFond and Park (1999) hypothesize that performance measurement is more informative in competitive industries where information on competitors’ performance can be used for relative performance evaluation. Consistent with our prediction, they find a positive relation between CEO turnover and industry competition and thus, by implication, between CEO turnover and the informativeness of performance measurement. Moreover, and also consistent with our prediction, Aggarwal and Samwick (1999b) find a positive relation between CEO pay-for-performance sensitivity and industry competition and thus, by implication, the informativeness of performance measurement.¹⁵

In Proposition 2, CEO turnover and incentive pay move in opposite directions following a change in the model’s primitives. Note that all of these primitives directly affect the firm’s agency problem: The CEO’s cost of shirking, cash-flow risk, and the informativeness of performance measurement. As we will show next, CEO turnover and incentive pay move in the same direction if the change in primitives is with respect to the value created under a replacement CEO.

¹⁵Aggarwal and Samwick (1999b) provide an alternative explanation for their empirical finding based on firms’ strategic interaction in the product market.

Proposition 3 *The model predicts more high-powered CEO incentive pay (higher Δ_w) and higher CEO turnover (higher θ^*) if the firm value under a potential replacement CEO is higher (higher V).*

Proof. See Appendix.

Proposition 3 considers an increase in the (match-specific) firm value under a potential replacement CEO, V . It is important to emphasize again the notion of match specificity. If the CEO is replaced because the (match-specific) firm value is higher under another CEO then this should not affect the incumbent CEO's outside option payoff u_0 . The replacement merely signals that the incumbent CEO is no longer the best possible match for this *particular* firm, but it does not signal anything about the value he could create at another firm. By the same logic, an increase in V should not affect the expected firm value under the incumbent CEO if he continues.

An increase in V leads to an increase in the first-best cutoff value θ_{FB} (see equation (2)). This is in contrast to Proposition 2, where a change in primitives only affected the firm's agency problem but not the first best. As θ_{FB} increases it becomes optimal to also increase the second-best cutoff value θ^* , thereby increasing the likelihood that the CEO is replaced.¹⁶ This in turn requires that conditional on *not* being replaced the CEO must receive a higher ex-post quasi rent, or else he would prefer to shirk. But this only strenghtens the CEO's bias towards continuation, putting downward pressure on the optimal cutoff value θ^* , contrary to what the firm aims to accomplish. By our previous arguments it is then optimal to grant the CEO steeper incentive pay, i.e., higher Δ_w , and possibly also higher severance pay.

Proposition 3 does not immediately lend itself to a testable prediction as V —the firm value under a potential replacement CEO—is likely to be unobservable in practice. (An observable takeover bid might provide some idea though.) However, given that an increase in V implies an increase in θ_{FB} we can look for observable characteristics that—while outside the CEO's control—might proxy for a higher likelihood that replacement is (first-best) efficient. For instance, firms may have growth opportunities whose arrival is ex-ante uncertain and which require a fundamental change in strategy that cannot be realized

¹⁶Note that Proposition 3 is silent about changes in CEO entrenchment. Since θ_{FB} and θ^* both increase, CEO entrenchment could be higher or lower after an increase in V .

with the incumbent CEO. In this case, Proposition 3 would predict that firms with substantial growth opportunities should have more CEO turnover and steeper CEO incentive pay. Consistent with this prediction, several empirical studies find that CEO pay-for-performance sensitivity and the use of stock-option grants are greater in firms with higher market-to-book ratios, which is a common proxy for growth opportunities (e.g., Smith and Watts, 1992; Gaver and Gaver, 1993; Core and Guay, 1999; Fahlenbrach, 2009).

Proposition 4 *The model predicts higher CEO turnover (higher θ^*), and thus less entrenchment, as firm size increases.*

Proof. See Appendix.

In Proposition 4 an increase in firm size refers to scaling up both V and s by some common factor α . Unlike an increase in V alone, scaling up firm size does not affect the first-best cutoff value θ_{FB} . Also unaffected are the CEO’s incentives to shirk and his continuation preferences at the interim stage.¹⁷ Scaling up firm size does, however, affect the board’s tradeoff between alleviating CEO entrenchment and minimizing his rents. Intuitively, there is “more at stake” at larger firms, implying that entrenchment is more costly. Hence, scaling up firm size (higher α) increases the optimal cutoff value θ^* , thus increasing CEO turnover.¹⁸ Consistent with this prediction, Lausten (2002) and Bushman, Dai, and Wang (2009) both find a positive relation between CEO turnover and firm size, while Mikkelsen and Partch (1997) find no significant relation.

6 Board Information

To model the CEO’s informational advantage in the simplest possible way we assumed that only the CEO—but not the firm’s board—can observe the state of nature θ . While we continue to assume that the CEO privately observes θ we now additionally assume that

¹⁷Scaling up the CEO’s shirking benefits B along with V and S renders the implications for CEO turnover ambiguous, as is evident from a comparison of Proposition 4 and Case i) of Proposition 2, which make opposite predictions regarding CEO turnover. Also, while a change in B affects condition (18) and thus whether the first best is feasible, the same is not true for a change in α .

¹⁸While it can be shown that Δ_w is weakly increasing in α , this does not provide a satisfactory measure of how the CEO’s incentives change given that Δ_s is scaled up at the same time. (In the empirical literature a common proxy for CEO incentives is the pay-for-performance sensitivity Δ_w/Δ_s).

prior to the replacement decision a publicly observable signal z is realized that is informative about θ . The question is how CEO compensation, and especially CEO turnover, depend on the public signal z .

The signal z is informative about θ in the sense that it is generated by the distribution function $G(z | \theta)$ with density $g(z | \theta)$ and support $z \in Z := [\underline{z}, \bar{z}]$. We assume that $G(z | \theta)$ satisfies the Monotone Likelihood Ratio Property (MLRP), implying that higher values of z are associated with higher likely values of θ . The ex-ante density of z is given by $h(z) := \int_{\Theta} g(z | \theta) f(\theta) d\theta$, while it follows from Bayes' rule that the posterior density satisfies $\hat{h}(\theta | z) := f(\theta)g(z | \theta)/h(z)$, where the posterior distribution $\hat{H}(\theta | z)$ satisfies MLRP.

For brevity we restrict attention to the case where the performance measure is fully informative so that $\mu = 1$ and therefore $\rho = 0$. Given that the signal is publicly observable the board can offer a menu of compensation schemes $\{\underline{w}(z), \Delta_w(z), W(z)\}$, implying that for each z there is a corresponding optimal cutoff value θ_z^* solving

$$\hat{w}(\theta_z^*, z) = W(z) + u_0, \quad (21)$$

provided the solution is interior. If all optimal cutoff values are interior we obtain in perfect analogy to (8) the (binding) “non-shirking” condition

$$\int_Z \left[\int_{\theta_z^*}^{\bar{\theta}} [\hat{w}(\theta, z) - W(z) - u_0] \hat{h}(\theta | z) d\theta \right] h(z) dz = B, \quad (22)$$

while firm profits are now given by

$$\Pi := \int_Z \left[\int_{\theta_z^*}^{\bar{\theta}} [v(\theta) - \hat{w}(\theta, z)] \hat{h}(\theta | z) d\theta + [1 - \hat{H}(\theta_z^* | z)] [V - u_0 - W(z)] \right] h(z) dz. \quad (23)$$

Before we proceed let us briefly comment on the different cases that can arise here. Note first that in order to attain the first best it must hold that $\theta_z^* = \theta_{FB}$ together with $W(z) = 0$ for all z . Moreover, by our previous arguments, in the (marginal) case where the first best is *just* feasible it must hold that $\underline{w}(z) = 0$ for all z . To see under what conditions it is possible to attain the first best we can thus substitute $\underline{w}(z) = 0$, which together with $\theta_z^* = \theta_{FB}$ and $W(z) = 0$ implies that $\Delta_w(z) = \Delta_w$. As the menu $\{\underline{w}(z), \Delta_w(z), W(z)\}$ is thus degenerate, the condition under which the first best is feasible is still characterized by (18). To focus on the interesting case where the optimal replacement policy depends

non-trivially on the signal—as opposed to the first-best case where $\theta_z^* = \theta_{FB}$ for all z —we assume in what follows that condition (18) does not hold. Note also that condition (20) is still sufficient to ensure that it is optimal to induce the CEO to work hard.¹⁹

It is immediate that under the optimal mechanism it must hold that $\theta_z^* < \theta_{FB}$ and therefore $\theta_z^* < \bar{\theta}$ for all z .²⁰ Expressions (22) and (23) presume, however, that all optimal θ_z^* are interior and thus that additionally $\theta_z^* > \underline{\theta}$ holds for all z , implying that the CEO is always replaced with positive probability. In what follows we assume that it is optimal to ensure that the CEO is always replaced with positive probability.²¹ The board’s problem is then to choose a menu $\{\underline{w}(z), \Delta_w(z), W(z)\}$ that maximizes firm profits Π , as given by (23), subject to the “non-shirking” condition (22), the non-negativity constraints $\underline{w}(z) \geq 0$ and $W(z) \geq 0$, and the requirement that θ_z^* must be incentive compatible from the CEO’s perspective as given by the CEO’s indifference condition (21).²²

Proposition 5 *Suppose that prior to the board’s replacement decision a publicly observable signal z is realized that is informative about the state of nature θ in the sense of MLRP. If condition (18) does not hold, so that it is not possible to implement the first-best cutoff value $\theta_z^* = \theta_{FB}$ without severance pay, then higher signals z are associated with:*

- i) lower optimal cutoff values $\theta_z^* < \theta_{FB}$, implying lower CEO turnover;*
- ii) steeper CEO incentive pay (higher $\Delta_w(z)$) while the base pay is $\underline{w}(z) = 0$ for all z ;*
- iii) higher severance pay (higher $W(z)$).*

Proof. See Appendix.

A higher signal z makes it less likely that the CEO is replaced, for two reasons. First, for a given optimal cutoff value θ_z^* a higher signal makes it more likely (by MLRP of the posterior distribution $\hat{H}(\theta | z)$) that $\theta \geq \theta_z^*$. Second, the higher is the signal the lower is the optimal cutoff value θ_z^* . In addition to making it less likely that the CEO is replaced a higher signal is also good news for the CEO because—although he can directly observe

¹⁹Recall that condition (20) was derived by considering a replacement policy under which the CEO is never replaced. By construction, condition (20) therefore does not depend on the public signal z .

²⁰That $\theta_z^* < \theta_{FB}$ must hold is implied by equation (32) in the Proof of Proposition (5). Recall that in our basic model the fact that $\theta_z^* < \bar{\theta}$ must hold was implied by incentive compatibility as otherwise the “non-shirking” condition (8) would be violated.

²¹This is true, for instance, if the CEO is replaced with positive probability under the highest possible signal $z = \bar{z}$, i.e. $\theta_z^* > \underline{\theta}$.

²²Recall that in our basic model it was not necessary to impose a non-negativity constraint $W \geq 0$ as it was always the case that $W \geq 0$ at the optimal solution.

θ —it implies that he receives both higher severance pay and steeper incentive pay and thus, given that $\underline{w}(z) = 0$ for all z , also higher on-the-job pay.

The intuition for Proposition 5 is straightforward. To satisfy the “non-shirking” condition (22) the CEO must be promised an ex-post quasi rent, which biases him towards continuation. As in our basic model—except now for each possible signal z —the board optimally trades off the value reduction due to CEO entrenchment against the costs of reducing entrenchment by offering the CEO severance pay, thus leaving him rents. At the first-best cutoff value $\theta_z^* = \theta_{FB}$ the first-order effect of a marginal reduction in θ_z^* on firm profits is again zero while the cost savings are positive, implying again that $\theta_z^* < \theta_{FB}$, except now for all z . Thus, there is again entrenchment in equilibrium. Also, recall that in our basic model steep incentive pay was optimal because it maximizes the CEO’s expected on-the-job pay at high values of θ where the expected firm value under his leadership is high. By the same logic, it is now optimal to give the CEO steeper incentive pay after higher signals z . As higher signals imply higher likely values of θ , this again maximizes the CEO’s expected on-the-job pay precisely when the expected firm value under his continued leadership is high.²³

Proposition 5 predicts lower CEO turnover after a positive signal about the CEO’s performance, which is consistent with empirical studies showing that poorly performing CEOs (based on either accounting or stock price measures of performance) are more likely to be replaced (e.g., Coughlan and Schmidt, 1985; Weisbach, 1988; Huson, Parrino, and Starks, 2001).

Importantly, the fact that the CEO receives higher on-the-job pay after higher signals does not constitute a reward for more effort. Recall that on the equilibrium path (where he works hard) the CEO has absolutely no control over the state of nature θ and thus over the signal z . Thus, whether the signal is high or low reflects pure luck on the CEO’s part. Bertrand and Mullainathan (2001) and Garvey and Milbourn (2006) find evidence for such a “reward for luck” but attribute it to managerial skimming. Our model offers an alternative explanation based on optimal contracting.

²³By giving the CEO steeper incentive pay after higher signals the board can achieve a better alignment between the CEO’s expected on-the-job pay and the state of nature θ relative to the basic model where the CEO (optimally) receives the same pay scheme for all θ . By implication, expected firm profits are strictly higher compared to the basic model.

7 Renegotiations and “Golden Handshakes”

In our model the CEO’s compensation scheme is chosen at the time when he is hired. Suppose now instead that after the CEO has exerted effort but before the state of nature θ is realized the CEO and the board could renegotiate the original contract. Based on the renegotiated contract—which gives rise to a new optimal cutoff value θ^* —the CEO then plays a message game as before. The point is that at the interim stage severance pay is more effective at reducing entrenchment than it is at the ex-ante stage. This is because at the interim stage the CEO’s effort is already sunk, implying that an increase in severance pay need no longer be matched by a simultaneous increase in the CEO’s on-the-job pay to satisfy the “non-shirking” condition (8). That severance pay is more effective at the interim stage can be formally seen from implicit differentiation of the CEO’s indifference condition (5),

$$\frac{d\theta^*}{dW} = \frac{1}{\pi'(\theta^*)\Delta_w}, \quad (24)$$

where the incentive slope Δ_w is held fixed as it no longer needs to adjust to a change in W . In contrast, at the ex-ante stage, where Δ_w must adjust to a change in W to satisfy the “non-shirking” condition (8), implicit differentiation of (5) yields

$$\frac{d\theta^*}{dW} = \frac{1}{\pi'(\theta^*)\Delta_w + \pi(\theta^*)(d\Delta_w/d\theta^*)}. \quad (25)$$

Given that $d\Delta_w/d\theta^* > 0$ the right-hand side in (25) is smaller than the right-hand side in (24). Not only is severance pay more effective at the interim stage, but it is also less costly. From an ex-ante perspective a dollar of severance pay constitutes a dollar of rent for the CEO, as was shown in Section 3. In contrast, at the interim stage each additional dollar of severance pay is only paid with probability $F(\theta^*)$, thus further strengthening the board’s incentives to offer the CEO additional severance pay (“golden handshake”) at the interim stage.

Accordingly, whenever the original contract specifies positive severance pay, then it is not renegotiation proof. The board always wants to notch up the CEO’s severance pay later on, which furthermore implies that the original contract violates the “non-shirking” condition (8). To restore incentive compatibility the board must consequently also increase the CEO’s expected on-the-job pay.²⁴ If the original contract specifies no severance pay then

²⁴Strictly speaking, the board is indifferent between offering a renegotiation-proof contract at the ex-

it may or may not be optimal for the board to offer the CEO a “golden handshake” given that doing so is (still) costly to the firm.

Proposition 6 *Suppose condition (18) does not hold, so that it is not possible to implement the first-best cutoff value $\theta^* = \theta_{FB}$ without severance pay. If the original contract specifies positive severance pay and the board cannot commit not to renegotiate, then it will award the CEO an additional “golden handshake” at the interim stage, where the CEO’s effort is already sunk. In the absence of commitment the CEO thus receives both higher severance pay and higher on-the-job pay. Moreover, the optimal cutoff value θ^* is higher—implying CEO entrenchment is lower—though the optimal replacement policy remains inefficient, i.e., it still holds that $\theta^* < \theta_{FB}$.*

Proof. See Appendix.

In the absence of commitment there is an *additional* benefit from offering the CEO incentive pay: It makes awarding a “golden handshake” at the interim stage less effective, as can be seen from (24) where $d\theta^*/dW$ is decreasing in Δ_w . Consequently, offering incentive pay ex ante mitigates the board’s commitment problem at the interim stage. Intuitively, when put on a steeper incentive scheme the CEO’s expected on-the-job pay $\widehat{w}(\theta)$ becomes steeper. At the interim stage this implies that either a larger increase in severance pay is needed to accomplish a given increase in the optimal cutoff value θ^* —thereby making a “golden handshake” more costly—or a given increase in severance pay induces a smaller increase in θ^* —thereby making a “golden handshake” less effective.

8 Policy Implications

Shareholder activists frequently rally against seemingly excessive severance pay for top executives. A common argument is that severance pay rewards failure instead of success. For instance, the U.K. Combined Code of Corporate Governance recommends contract durations and notice periods of at most one year for CEOs in order to restrict severance

ante stage and awarding the CEO a “golden handshake” by increasing his severance pay at the interim stage. Thus, in the absence of commitment the optimal CEO compensation scheme might feature either higher severance pay in the original (renegotiation-proof) contract or the prospect of an additional “golden handshake” at the interim stage.

costs.²⁵ Our model is consistent with the notion that granting CEOs severance pay risks rewarding them for failure, thereby reducing their incentives to work hard. In our model, each dollar of severance pay constitutes a dollar of rent for the CEO.

However, severance pay also rewards CEOs for giving up entrenchment, at least when coupled with high-powered on-the-job incentive pay (see Lemma 1). In our model, policy intervention that imposes a binding cap on severance pay would thus arguably reduce CEO pay—as well as CEO rent extraction—but the associated increase in entrenchment would reduce social welfare and firm profits. This is markedly different if policy intervention were to target only “golden handshakes,” i.e., severance pay that is not part of the original contract but that is additionally granted at the time of the CEO’s departure. In our model, firms would benefit from a binding regulatory cap on such “golden handshakes” as it would mitigate their commitment problem vis-à-vis the CEO at the interim stage. However, social welfare would be lower given that “golden handshakes” help to mitigate CEO entrenchment.²⁶

Corollary 1 *A binding regulatory cap on contractually stipulated severance pay leads to lower CEO pay and less rent extraction by the CEO but also more entrenchment. Overall, such a cap reduces both firm profits and social welfare. In contrast, a binding regulatory cap on “golden handshakes” that are granted at the time of the CEO’s departure also leads to lower CEO pay and more entrenchment, and thus lower social welfare, but higher firm profits.*

The notion that the ability to commit not to renegotiate severance pay increases firm profits extends to the case where CEO compensation, including severance pay, can be made contingent on some verifiable event that affects the match value. Instead, we would conjecture that the flexibility to renegotiate contracts might be of value if the event is observable but not verifiable. The policy implications of Corollary 1 should be viewed with this caveat in mind.

²⁵In response to public pressure, the U.K. government recently initiated an inquiry about “rewards for failure” (DTI 2003). The U.K., like many other countries, has also witnessed a recent wave of shareholder activism against large severance packages involving companies such as WPP, Sainsbury’s, and GlaxoSmithKline.

²⁶The divergence between firm profits and social welfare follows from the fact that rents to the CEO reduce firm profits but not social welfare. From a social welfare perspective all that matters is whether CEO entrenchment is reduced.

9 Conclusion

This paper derives joint implications for the optimal CEO compensation and replacement policy based on a model of “information-based entrenchment.” In our model, the CEO’s desire to become entrenched is endogenous and does not derive from exogenously specified private benefits of control. Rather, it derives from the optimal compensation scheme inducing the CEO to work hard, which must promise him an ex-post quasi rent (in the form of generous on-the-job pay) in case he continues. This biases the CEO towards continuation which, together with his private information at the interim stage, drives a wedge between efficient CEO replacement and actual CEO turnover.

High-powered incentive pay, and possibly also severance pay, can mitigate CEO entrenchment. Incentive pay ensures that the CEO’s expected on-the-job pay is high precisely when the firm value under his continued leadership is high, thus aligning the CEO’s continuation preferences with those of the firm at the replacement stage. The role of severance pay in our model is more nuanced, as an increase in severance pay must be accompanied by a simultaneous increase in the CEO’s on-the-job pay. Otherwise, there would be too high a reward for failure and the CEO would shirk. As a result, each dollar of severance pay constitutes a dollar of rent for the CEO. Importantly, however, whether severance pay can mitigate entrenchment depends on the structure of the CEO’s on-the-job pay. As our model shows, the firm gets the biggest “bang for the buck” (i.e., the biggest reduction in entrenchment) if an increase in severance pay is accompanied by a simultaneous increase in incentive pay, not base pay.

Our model abstracts from many real-world features that are likely to affect CEO turnover and compensation in practice. For instance, the board’s role may not be confined to monitoring (and replacing) the CEO but it may additionally include giving the CEO valuable advice, as in Adams and Ferreira (2007). Also, our model assumes that the board acts in the firm’s best interest when designing the optimal CEO compensation and replacement policy. Alternatively, it has been argued that boards are captured by the CEO, maximizing the CEO’s utility instead of firm profits (e.g., Bebchuk and Fried, 2004). Realistically, the truth will probably lie somewhere in between these two polar views. Extending our model along these lines might provide a fruitful avenue for future research.

Appendix A: Proofs

Proof of Lemma 2. We argue to a contradiction and assume that, while condition (18) does not hold, it is optimal to set $\underline{w} > 0$.

Suppose first that $W > 0$. By (9) and $\underline{w} > 0$ we can decrease both \underline{w} and W —while increasing Δ_w to satisfy the “non-shirking” condition (8)—such that the right-hand side in (9) remains constant, implying that θ^* remains unchanged. As this means that the same cutoff value θ^* can be implemented with less severance pay W while condition (8) remains satisfied, firm profits Π must be strictly higher, contradiction.

Suppose next that $W = 0$. Note from (9) that a decrease in \underline{w} —combined with an increase in Δ_w to satisfy the “non-shirking” condition (8)—leads to a strictly higher cutoff value θ^* . (Recall that the left-hand side of (9) is strictly decreasing in θ^* .) Given that $W = 0$, while condition (18) does not hold, the higher cutoff value θ^* still satisfies $\theta^* < \theta_{FB}$, while firm profits Π are again strictly higher, contradiction. **Q.E.D.**

Proof of Proposition 2. We consider each of the three assertions i)-iii) in turn.

Assertion i) considers a marginal change in B . If condition (18) holds, so that the first best is feasible both before and after the change in B , then the result regarding Δ_w follows immediately from (16), while $\theta^* = \theta_{FB}$ remains unchanged. Note that condition (18) is relaxed as B decreases. If condition (18) does not hold, so that the first best is not feasible, then we must distinguish between two cases. If $W = 0$ it follows immediately from (9) that $\theta^* < \theta_{FB}$ is strictly decreasing in B . As for Δ_w , note from (6) that for a given θ^* and $W = 0$ (and $\underline{w} = 0$ by Lemma 2) we have that

$$\Delta_w = \frac{u_0}{(1 - \mu) + p(\theta^*)(2\mu - 1)}, \quad (26)$$

which is strictly decreasing in θ^* and thus also in $p(\theta^*)$. Recall next that when the firm just starts to pay severance pay, i.e., when W changes from zero to $W > 0$, then to satisfy (8) this requires that it must also raise Δ_w , which by Lemma 1 implies that θ^* increases. We now show that when taking into account the optimal adjustment of the CEO’s contract, then following an increase in B the optimal cutoff value θ^* must decrease. For this we make use of the postulated strict quasiconcavity of the board’s objective function in W . Specifically, we now evaluate the derivative $d\Pi/dW$ in (19) at a *given* cutoff value θ^* . We already know that, given any two values $B'' > B'$, the former requires strictly higher

severance pay, $W'' > W'$. After substituting for $d\theta^*/dW$ from Lemma 1 together with $\underline{w} = 0$ from Lemma 2 we thus obtain that the firm's benefit from raising θ^* is strictly higher with (B', W') than with (B'', W'') if and only if

$$\frac{B'}{(u_0 + W')^2} > \frac{B''}{(u_0 + W'')^2}. \quad (27)$$

If condition (27) holds the (*true*) optimal cutoff value must be strictly lower under $B = B''$ than under $B = B'$. To see that (27) indeed holds, note that for a given θ^* rearranging (9) (with $\underline{w} = 0$) yields

$$\frac{B'}{u_0 + W'} = \frac{B''}{u_0 + W''}, \quad (28)$$

implying that condition (27) transforms to $W'' > W'$, which is true.

Assertion ii) considers a mean-preserving spread (MPS) in firm value, s . To apply the MPS uniformly for all $\theta \in \Theta$ we consider a (marginal) change from the firm value distribution $(p(\theta), \Delta_s)$ to the distribution $(\tilde{p}(\theta), \tilde{\Delta}_s)$ such that $p(\theta)\Delta_s = \tilde{p}(\theta)\tilde{\Delta}_s$ for all θ . It thus holds that $\tilde{p}(\theta) = \gamma p(\theta)$, where $\gamma := \Delta_s/\tilde{\Delta}_s < 1$. If the first best is feasible both before and after the MPS the assertion regarding Δ_w follows immediately from the characterization in (16). Also, condition (18) becomes harder to satisfy after the MPS (i.e., after substituting Δ_s with $\tilde{\Delta}_s > \Delta_s$ while noting that ΔV_{FB} remains unchanged). Next, in the second-best case with $W = 0$ we obtain from (9) (with $\underline{w} = 0$) that after the MPS the optimal cutoff value, which we shall denote by $\tilde{\theta}^*$, is determined by

$$\int_{\tilde{\theta}^*}^{\bar{\theta}} \left[\frac{p(\theta) - p(\tilde{\theta}^*)}{\frac{\rho}{\gamma} + p(\tilde{\theta}^*)} \right] f(\theta) d\theta = \frac{B}{u_0}. \quad (29)$$

Compared to the case before the MPS, which is simply obtained from substituting $\gamma = 1$ and using θ^* , we thus have from (29) that $\tilde{\theta}^* < \theta^*$. That Δ_w must increase follows now immediately from (26) after noting that, when evaluated at the respective cutoff values, it holds that $\tilde{p}(\tilde{\theta}^*) < p(\theta^*)$ given that $\tilde{p}(\theta) < p(\theta)$ and $\tilde{\theta}^* < \theta^*$. Finally, as for the second-best case with $W > 0$ recall that a switch from $W = 0$ to $W > 0$ requires that the firm must also raise Δ_w . We now finally show in analogy to assertion i) that following a (marginal) MPS the firm's incentives to raise θ^* through a higher W are muted. Precisely, this is done by showing that at some *given* cutoff value θ^* the derivative $d\Pi/dW$ in (19) is lower after the MPS. For this, denote the respective levels of severance pay that are required to

implement θ^* by $\widetilde{W} > W$. At the given θ^* it needs to be shown, in analogy to (27), that

$$\begin{aligned} & \frac{1}{p'(\theta^*)} \frac{B}{(u_0 + W)^2} \frac{[\rho + p(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} [\rho + p(\theta)] f(\theta) d\theta} \\ & > \frac{1}{\widetilde{p}'(\theta^*)} \frac{B}{(u_0 + \widetilde{W})^2} \frac{[\rho + \widetilde{p}(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} [\rho + \widetilde{p}(\theta)] f(\theta) d\theta}. \end{aligned}$$

Substituting for $\widetilde{p}(\theta) = \gamma p(\theta)$ and using (9) this inequality holds if

$$\frac{[p(\theta) - p(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} [\rho + p(\theta)] f(\theta) d\theta} > \frac{[p(\theta) - p(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} \left[\frac{\rho}{\gamma} + p(\theta) \right] f(\theta) d\theta},$$

which is true as $\gamma < 1$.

Assertion iii) considers a marginal change in μ . The analysis if $W = 0$ (both first- and second-best case) is analogous to that in assertions i) and ii). We thus confine ourselves to showing that in the second-best case with $W > 0$ the incentives to raise θ^* increase as μ increases. To show this, take $\mu'' < \mu'$ and thus $\rho'' > \rho'$. Suppose again that the same cutoff value θ^* is implemented in both cases by choosing the respective levels of severance pay $W'' > W'$ accordingly. Then the derivative $d\Pi/dW$ in (19), after using Lemma 1 together with (9), is strictly lower with (ρ'', W'', θ^*) than with (ρ', W', θ^*) if

$$\frac{[p(\theta) - p(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} [\rho' + p(\theta)] f(\theta) d\theta} > \frac{[p(\theta) - p(\theta^*)]^2}{\int_{\theta^*}^{\bar{\theta}} [\rho'' + p(\theta)] f(\theta) d\theta},$$

which is true as $\rho'' > \rho'$. **Q.E.D.**

Proof of Proposition 3. If the first best is feasible both before and after the change in V then following an increase in V —and thus a reduction in ΔV_{FB} (as defined in (15))—it holds from (16) that Δ_w strictly increases. Note next that condition (18) is relaxed when V is smaller. As for the second-best case with $W = 0$ note that a marginal change in V does not affect either θ^* or the optimal compensation scheme, as is evident from (9). Finally, in the second-best case with $W > 0$, which is more likely the higher is V , it follows from the first-order condition $d\Pi/dW = 0$ (using (19)) that the optimal W , and thus also the corresponding optimal θ^* , are strictly increasing in V , resulting in a higher Δ_w (see (13)). (Note here that $d\theta^*/dW$ is unaffected by the change in V given strict quasiconcavity of the board's objective function.) **Q.E.D.**

Proof of Proposition 4. Note first that condition (18) is unaffected as both s and V are scaled up by α . (In particular, note with respect to condition (18) that both ΔV_{FB} and the term in brackets on the left-hand side are scaled up by α .) In analogy to the Proof of Proposition 3 the optimal cutoff value θ^* remains unaffected if the second-best case with $W = 0$ holds both before and after the change in α . As for the second-best case with $W > 0$, note first that this case becomes more likely if, for a given $\theta_0^* < \theta_{FB}$, we increase α and thereby the difference $V - v(\theta_0^*) > 0$. As in the Proof of Proposition 3 the assertion then follows again from the first-order condition $d\Pi/dW = 0$, while using (19), implying a higher value of W and thus also θ^* as α increases. **Q.E.D.**

Proof of Proposition 5. As we focus on the case where condition (18) does not hold one can again show that it is uniquely optimal to set $\underline{w}(z) = 0$ for all z . (For brevity we omit the proof, which is analogous to that in Lemma 2.) For a given signal z it thus remains to determine $\Delta_w(z)$ and $W(z)$. Given that θ_z^* is interior (see main text) this uniquely specifies θ_z^* from (21), which implies we can instead consider the dual problem of choosing $W(z)$ and θ_z^* , while we obtain $\Delta_w(z)$ from

$$\Delta_w(z) = \frac{W(z) + u_0}{p(\theta_z^*)}, \quad (30)$$

which follows from (21) after substituting for $\widehat{w}(\theta_z^*, z)$. Setting up the board's problem, after substituting for $\widehat{w}(\theta, z)$ and using (30), the Lagrangian is

$$\begin{aligned} & \int_Z \left[\int_{\theta_z^*}^{\bar{\theta}} \left[v(\theta) - V - (W(z) + u_0) \left(\frac{p(\theta)}{p(\theta_z^*)} - 1 \right) \right] \widehat{h}(\theta | z) d\theta - W(z) \right] h(z) dz \quad (31) \\ & + \lambda \left[\int_Z \left[(W(z) + u_0) \int_{\theta_z^*}^{\bar{\theta}} \left[\frac{p(\theta)}{p(\theta_z^*)} - 1 \right] \widehat{h}(\theta | z) d\theta \right] h(z) dz - B \right] + V - u_0. \end{aligned}$$

Pointwise maximization with respect to θ_z^* (recall that θ_z^* is interior) yields, after substituting for $\widehat{h}(\theta | z)$, the first-order condition

$$[V - v(\theta_z^*)] f(\theta_z^*) g(z | \theta_z^*) - (\lambda - 1) (W(z) + u_0) \frac{p'(\theta_z^*)}{[p(\theta_z^*)]^2} \int_{\theta_z^*}^{\bar{\theta}} p(\theta) f(\theta) g(z | \theta) d\theta = 0. \quad (32)$$

Likewise, pointwise maximization with respect to $W(z)$ yields, given an interior solution $W(z) > 0$, the first-order condition

$$(\lambda - 1) \int_{\theta_z^*}^{\bar{\theta}} \left[\frac{p(\theta)}{p(\theta_z^*)} - 1 \right] \widehat{h}(\theta | z) d\theta - 1 = 0. \quad (33)$$

(See below for the case where $W(z) = 0$.)

We are interested in the comparative statics of $(\theta_z^*, W(z))$ in z . For this purpose consider first the case where before and after a marginal change in z the optimal values satisfy the respective first-order conditions (32) and (33). For expositional felicity it is convenient to denote these conditions, in a slight abuse of notation, by $a_{\theta^*}(\theta^*, W, z) = 0$ (for (32)) and $b_W(\theta^*, W, z) = 0$ (for (33)). In both cases the subscripts indicate that we consider first derivatives of the Lagrangian. Solving the system of total derivatives

$$\begin{pmatrix} a_{\theta^*\theta^*} & a_{\theta^*W} \\ b_{W\theta^*} & b_{WW} \end{pmatrix} \begin{pmatrix} d\theta^* \\ dW \end{pmatrix} = -dz \begin{pmatrix} a_{\theta^*z} \\ b_{Wz} \end{pmatrix}$$

we obtain from Cramer's rule, with determinant $D := a_{\theta^*\theta^*}b_{WW} - b_{W\theta^*}a_{\theta^*W}$, that

$$\frac{d\theta^*}{dz} = -\frac{D_{\theta^*}}{D}, \text{ where } D_{\theta^*} := a_{\theta^*z}b_{WW} - b_{Wz}a_{\theta^*W},$$

and

$$\frac{dW}{dz} = -\frac{D_W}{D}, \text{ where } D_W := a_{\theta^*\theta^*}b_{Wz} - b_{W\theta^*}a_{\theta^*z}.$$

To sign these derivatives consider first the function $a_{\theta^*}(\theta^*, W, z)$ evaluated at some pair (θ^*, W) where $a_{\theta^*} = 0$. We argue that at this point it holds for the corresponding second derivative that $a_{\theta^*z} < 0$. This follows after dividing the left-hand side of (32) by $g(z | \theta^*)$ and observing from MLRP of $G(z | \theta)$ that $g(z | \theta)/g(z | \theta^*)$ is strictly increasing in z for all $\theta > \theta^*$. (Note also that $\lambda > 1$ from (33).) Next, it is immediate that $a_{\theta^*W} < 0$. Turning to the function $b_W(\theta^*, W, z)$ evaluated at some pair (θ^*, W) where $b_W = 0$, note first that the posterior $\hat{H}(\theta | z)$ satisfies MLRP. From this we have that $b_{Wz} > 0$ while it is immediate that $b_{W\theta^*} < 0$.

Given strict concavity of the firm's problem, the (Hessian) matrix of second-order derivatives of the objective function is negative semi-definite, which implies that $a_{\theta^*\theta^*} < 0$ and $b_{WW} < 0$, next to $D > 0$. From this, together with the preceding observations on the second-order derivatives, we have, first, that $D_{\theta^*} > 0$ and $D_W < 0$ and, second, that indeed $d\theta^*/dz < 0$ and $dW/dz > 0$. Note further that from monotonicity of $p(\theta)$ we have from (30) that also $d\Delta_w/dz > 0$. Finally, while the preceding analysis assumed that $W(z) > 0$, given the signs of the derivatives of the first-order conditions $a_{\theta^*} = 0$ and $b_W = 0$ it also extends (weakly) to the case where $W(z) = 0$. (Precisely, we have for any $z'' > z'$ that $W(z') > 0$ implies $W(z'') > 0$, while it may be that $W(z') = 0$ but $W(z'') > 0$.) **Q.E.D.**

Proof of Proposition 6. At the interim stage a marginal change in W affects firm profits Π according to the derivative

$$f(\theta^*) [V - v(\theta^*)] \frac{d\theta^*}{dW} - F(\theta^*), \quad (34)$$

while the corresponding derivative at the ex-ante stage is given by (19). Hence, given that $F(\theta^*) < 1$, in order to show that an increase in W raises firm profits by strictly more at the interim stage, it suffices to show that $d\theta^*/dW$ is equal or greater at the interim stage. This was shown in the main text. Note also that from (34) the board would not want to increase W all the way up to $\theta^* = \theta_{FB}$.

One option for the board is to offer in $t = 0$ a renegotiation proof contract. The board's program is then to maximize firm profits Π subject to (5), which pins down θ^* , the "non-shirking" constraint (8), the constraint that $\underline{w} \geq 0$, and the constraint that, in addition, the optimal contract is renegotiation proof. Using the shorter expression (24) for $d\theta^*/dW$, for a given triplet $(\underline{w}, \Delta_w, W)$, it must thus additionally hold that

$$\Delta_w \geq \frac{f(\theta^*)}{\pi'(\theta^*)F(\theta^*)} [V - v(\theta^*)]. \quad (35)$$

We already know that a given cutoff value θ^* can be obtained with less severance pay when the CEO's expected on-the-job pay is made as steep as possible. By (35) a switch to such a steeper on-the-job pay scheme that implements the same cutoff value also relaxes the renegotiation-proofness constraint. **Q.E.D.**

Appendix B: Continuum of Firm Values

The following analysis extends our basic argument to a continuum of firm values $s \in S := [\underline{s}, \bar{s}]$. Each θ gives rise to a conditional distribution function over cash flows, $G_\theta(s)$, with associated density $g_\theta(s)$, where $G_\theta(s)$ is continuously differentiable in both s and θ . We assume that θ is informative in the sense of the Monotone Likelihood Ratio Property (MLRP), implying that $g_{\theta''}(s)/g_{\theta'}(s)$ is strictly increasing in s for all $\theta'' > \theta' > \underline{\theta}$. As previously we define $v(\theta) := \int_S s g_\theta(s) ds$, which is continuous and strictly increasing in θ .

For brevity we restrict consideration to the case where the performance measure y is perfectly informative, such that $y = s$.²⁷ We can thus write the CEO's on-the-job pay

²⁷The following argument can be extended to the case where s gives rise to a non-degenerate distribution, say $H_s(y)$, which also satisfies MLRP.

$w(y)$ directly as a function of s , i.e., $w(s)$. We also require that $w(s) \leq s$.²⁸ As a final requirement, which helps to simplify the analysis but does not bind at the optimum, we stipulate that $w(s)$ must be non-decreasing.

As in our main analysis we derive the solution for the case where the following conditions hold: i) It is optimal to induce the CEO to work hard, ii) the first-best cutoff value is interior, $\underline{\theta} < \theta_{FB} < \bar{\theta}$, and iii) CEO replacement occurs with positive probability in equilibrium, $\underline{\theta} < \theta^* < \bar{\theta}$. For brevity, and in contrast to our main analysis, we shall not derive (sufficient) conditions under which all these conditions hold.

Note first that the CEO's indifference condition (5) and the "non-shirking" condition (8) still apply, while the board's objective function is also the same. For brevity we focus on the interesting case where the first best is not feasible (precisely, where the first-best cutoff value θ_{FB} cannot be implemented without severance pay). The precise condition for this will be derived at the end of this proof. Our key result is then as follows.

Proposition A1. *If the first best is not feasible the uniquely optimal compensation scheme specifies $w(s) = 0$ for $s < \hat{s}$ and $w(s) = s$ if $s \geq \hat{s}$, where $\hat{s} \in (\underline{s}, \bar{s})$.*

The second-best optimal on-the-job pay scheme is a "live-or-die" contract that pays the CEO as much as possible, namely, $w(s) = s$, for all sufficiently high cash flows and as little as possible, namely, $w(s) = 0$, for all lower cash flows. Before we provide a proof of this result note that the underlying argument is the same as in the case with only two firm-value realizations: The optimal on-the-job pay scheme shifts as much as possible of the CEO's on-the-job pay into high firm values s , thus making his expected on-the-job pay $\hat{w}(\theta)$ steeper and pushing down the optimal cutoff value θ^* (or, likewise, allowing to implement a given cutoff value at lower cost, W). While this argument differs from existing models of optimal contract (or security) design, it shares with some of these models the fact that it relies on MLRP (or similar assumptions on the distribution of cash flows). In Innes (1990), for example, in order to reduce agency costs, the principal wants to maximize the expected payoff differential between low and high effort, which by MLRP is achieved by paying the agent only in high cash-flow states. A similar logic also applies to problems with ex-ante private information (in contrast to our case with interim private information). For example, Nachman and Noe (1994) show that high-value firms prefer

²⁸As will be clear from the optimal solution, we would otherwise have an open-set problem.

the pooling (financing) contract that allows them to retain cash flows for high cash-flow realizations rather than low cash-flow realizations. By maximizing the expected cash-flow differential between high- and low-value firms this reduces the differential with respect to the investor's stake, thus minimizing the high-value firms' underpricing problem.

To prove Proposition A1 we argue to a contradiction. Suppose it was optimal to implement a given θ^* with a different on-the-job pay scheme $\tilde{w}(s)$. Denote the corresponding severance payment by \tilde{W} . We show that there exists some on-the-job pay scheme $w(s)$ such that i) the “non-shirking” constraint (8) continues to bind and ii) we can still implement the same θ^* , though now with lower severance pay W . (The argument for the case where $\tilde{W} = 0$ holds is analogous; see also the Proof of Lemma 2.) That is, in a slight abuse of notation, we show that the new on-the-job pay scheme satisfies $\theta^*(w, W) = \theta^*(\tilde{w}, \tilde{W}) = \theta^*$ and $W < \tilde{W}$, contradicting the optimality of $\tilde{w}(s)$.

We proceed in two steps. We first choose $\overline{W} = \tilde{W}$ 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 $\theta^*(\overline{w}, \overline{W}) = \theta^*$. Defining $z(s) := \tilde{w}(s) - \overline{w}(s)$ we thus obtain

$$\int_{\underline{s}}^{\overline{s}} z(s) g_{\theta^*}(s) ds = 0. \quad (36)$$

Given the construction of $\overline{w}(s)$, there exists a value $\tilde{s} \in (\underline{s}, \overline{s})$ such that $z(s) \geq 0$ for all $s < \tilde{s}$ and $z(s) \leq 0$ for all $s > \tilde{s}$, where both inequalities are strict over sets of positive measure. Take now any $\hat{\theta} > \theta^*$. By MLRP of $G_{\theta}(s)$ and (36) it then holds that

$$\begin{aligned} \int_{\underline{s}}^{\overline{s}} z(s) g_{\hat{\theta}}(s) ds &= \int_{\underline{s}}^{\tilde{s}} z(s) g_{\theta^*}(s) \frac{g_{\hat{\theta}}(s)}{g_{\theta^*}(s)} ds + \int_{\tilde{s}}^{\overline{s}} z(s) g_{\theta^*}(s) \frac{g_{\hat{\theta}}(s)}{g_{\theta^*}(s)} ds \\ &< \frac{g_{\hat{\theta}}(\tilde{s})}{g_{\theta^*}(\tilde{s})} \int_{\underline{s}}^{\overline{s}} z(s) g_{\theta^*}(s) ds = 0, \end{aligned} \quad (37)$$

which implies that the “non-shirking” constraint (8) is slack under $\overline{w}(s)$ and \overline{W} .

In a second step we can construct the asserted on-the-job pay scheme with $w(s) = 0$ for $s < \hat{s}$ and $w(s) = s$ for $s \geq \hat{s}$ and $W < \overline{W} = \tilde{W}$. 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 (8) again binds. Here, the existence of a solution to the corresponding equation system (namely, the binding “non-shirking” constraint (8) and the CEO's indifference condition (5)) is ensured by continuity of all payoffs in \hat{s}' as well as \overline{W} and the fact that (8) is violated as $\hat{s}' \rightarrow \overline{s}$.²⁹

²⁹The equation system may have more than one solution. In this case, the firm strictly prefers the solution with the lowest W (and thus the lowest rents for the CEO).

Using the obtained characterization of the second-best contract, we can finally derive the condition for when the first best is not feasible. For this we obtain first from (5) a unique cutoff \hat{s} such that the contract specified in Proposition A1 (i.e., $w(s) = 0$ for $s < \hat{s}$ and $w(s) = s$ if $s \geq \hat{s}$) satisfies $\hat{w}(\theta_{FB}) = u_0$. Using this contract the first best is then feasible if and only if, after substituting into (8), it holds that

$$\int_{\theta_{FB}}^{\bar{\theta}} [\hat{w}(\theta) - \hat{w}(\theta_{FB})] f(\theta) d\theta \geq B.$$

References

- Adams, R., and D. Ferreira, 2007, A theory of friendly boards, *Journal of Finance* 62, 217-250.
- Aggarwal, R. K., and A. A. Samwick, 1999a, The other side of the tradeoff: The impact of risk on executive compensation, *Journal of Political Economy* 107, 65-105.
- Aggarwal, R. K., and A. A. Samwick, 1999b, Executive compensation, strategic competition, and relative performance evaluation: Theory and evidence, *Journal of Finance* 54, 1999-2043.
- Allgood, S., and K. A. Farrell, 2003, The match between CEO and firm, *Journal of Business* 76, 317-341.
- 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.
- Bertrand, M., and S. Mullainathan, 1999, Corporate governance and executive pay: Evidence from takeover legislation, mimeo, MIT.
- Bertrand, M., and S. Mullainathan, 2001, Are CEOs rewarded for luck? The ones without principals are, *Quarterly Journal of Economics* 116, 901-932.
- Bushman, R. M., R. J. Indjejikian, and A. Smith, 1996, CEO compensation: The role of individual performance evaluation, *Journal of Accounting and Economics* 21, 161-193.
- Bushman, R. M., Z. Dai, and X. Wang, 2009, Risk and CEO turnover, forthcoming, *Journal of Financial Economics*.

- Core, J., and W. Guay, 1999, The use of equity grants to manage optimal equity incentive levels, *Journal of Accounting and Economics* 28, 151-184.
- Coughlan, A. T., and R. M. Schmidt, 1985, Executive compensation, management turnover, and firm performance: An empirical investigation, *Journal of Accounting and Economics* 7, 43-66.
- DeFond, M. L., and C. W. Park, 1999, The effect of competition on CEO turnover, *Journal of Accounting and Economics* 27, 35-56.
- Demsetz, H., and K. Lehn, 1985, The structure of corporate ownership: Causes and consequences, *Journal of Political Economy* 93, 1155-1177.
- Denis, D. D., D. K. Denis, and A. Sarin, 1997, Ownership structure and top executive turnover, *Journal of Financial Economics* 45, 193-221.
- DTI, 2003, Rewards for failure: Directors' remuneration—contracts, performance and severance, Department of Trade and Industry, London.
- Eisfeldt, A. L., and A. A. Rampini, 2008, Managerial incentives, capital reallocation, and the business cycle, *Journal of Financial Economics* 87, 177-199.
- Engel, E., R.M. Hayes, and X. Wang, 2003, CEO turnover and properties of accounting information, *Journal of Accounting and Economics* 36, 197-226.
- Fahlenbrach, R., 2009, Shareholder rights, boards, and CEO compensation, *Review of Finance* 13, 81-113.
- Fisman, R. J., R. Khurana, and M. Rhodes-Kropf, 2005, Governance and CEO turnover: Do something or do the right thing?, mimeo, Columbia University.
- Fulghieri, P., and L. S. Hodrick, 2006, Synergies and internal agency conflicts: The double-edged sword of mergers, *Journal of Economics and Management Strategy* 15, 549-576.
- Garen, J. E., 1994, Executive compensation and principal-agent theory, *Journal of Political Economy* 102, 1175-1199.
- Garvey, G., and T. T. Milbourn, 2006, Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad, *Journal of Financial Economics* 82, 197-225.
- Gaver, J. J., and K. M. Gaver, 1993, Additional evidence on the association between the investment opportunity set and corporate financing, dividend, and compensation policies, *Journal of Accounting and Economics* 16, 125-160.
- Hadlock, C. J., and G. B. Lumer, 1997, Compensation, turnover and top management incentives: Historical evidence, *Journal of Business* 70, 153-187.
- Harris, E. G., 1990, Antitakeover measures, golden parachutes, and target firm shareholder welfare, *Rand Journal of Economics* 21, 614-625.

- Hermalin, B. E., 2005, Trends in corporate governance, *Journal of Finance* 60, Page 2351-2384.
- 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.
- Holmström, B. R., 1979, Moral hazard and observability, *Bell Journal of Economics* 10, 74-91.
- Huson, M. R., R. Parrino, and L. T. Starks, 2001, Internal monitoring mechanisms and CEO turnover: A long-term perspective, *Journal of Finance* 56, 2265-2297.
- 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.
- Knoeber, C. R., 1986, Golden parachutes, shark repellents, and hostile tender offers, *American Economic Review* 76, 155-167.
- Lambert, R. A., and D. F. Larcker, 1985, Golden parachutes, executive decision-making, and shareholder wealth, *Journal of Accounting and Economics* 7, 179-203.
- Lambert, R. A., and D. F. Larcker, 1987, An analysis of the use of accounting and market measures of performance in executive compensation contracts, *Journal of Accounting Research* 25, 85-125.
- Lausten, M., 2002, CEO turnover, firm performance and corporate governance: Empirical evidence on Danish firms, *International Journal of Industrial Organization* 20, 391-414.
- Levitt, S. D., and C. M. Snyder, 1997, Is no news bad news? Information transmission and the role of “early warning” in the principal-agent model, *RAND Journal of Economics* 28, 641-661.
- Mikkelsen, W. H., and M. M. Partch, 1997, The decline of takeovers and disciplinary managerial turnover, *Journal of Financial Economics* 44, 205-228.
- Nachman, D. C., and T. H. Noe, 1994, Optimal design of securities under asymmetric information, *Review of Financial Studies* 7, 1-44.
- Oyer, P., and S. Shaefer, 2005, Why do some firms give stock options to all employees? An empirical examination of alternative theories, *Journal of Financial Economics* 76, 99-133.
- Prendergast, C., 2002, The tenuous trade-off between risk and incentives, *Journal of Political Economy* 110, 1071-1102.
- Shleifer, A., and R. W. Vishny, 1989, Management entrenchment: The case of manager-specific investments, *Journal of Financial Economics* 25, 123-139.

Smith, C. W. Jr., and R. L. Watts, 1992, The investment opportunity set and corporate financing, dividend, and compensation policies, *Journal of Financial Economics* 32, 263-192.

Weisbach, M. S., 1988, Outside directors and CEO turnover, *Journal of Financial Economics* 20, 431-460.