Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 15-016

"Managerial Turnover and Entrenchment"

by

Zenan Wu and Xi Weng

http://ssrn.com/abstract=2581552

Managerial Turnover and Entrenchment*

Zenan Wu † Xi Weng ‡

Abstract

We consider a two-period model in which the success of the firm depends on the effort of a first-period manager (the incumbent) and the ability of a second-period manager. At the end of the first period, the board receives a noisy signal of the incumbent manager's ability and decides whether to retain or replace the incumbent manager. We show that the information technology the board has to assess the incumbent manager's ability is an important determinant of the optimal contract and replacement policy. The contract must balance providing incentives for the incumbent manager to exert effort and ensuring that the second-period manager is of high ability. We show that severance pay in the contract serves as a costly commitment device to induce effort. Unlike existing models, we identify conditions on the information structure under which both entrenchment and anti-entrenchment emerge in the optimal contract.

Keywords: entrenchment, managerial turnover, contracting, information order.

JEL classification: $D86$, $J33$, $M52$.

*We are deeply indebted to Hanming Fang and Aislinn Bohren for their guidance and continuous support. We are grateful to David Dillenberger and Andrew Postlewaite for their invaluable advice. We would like to thank Navin Kartik, SangMok Lee, George Mailath, Steven Matthews, Lucian Taylor and seminar participants at University of Pennsylvania for helpful comments and suggestions. Any remaining errors are ours.

†Department of Economics, University of Pennsylvania. zenan@sas.upenn.edu

‡Guanghua School of Management, Peking University. wengxi125@gsm.pku.edu.cn

1 Introduction

Designing compensation schemes in managerial contracts and deciding whether to replace a manager, such as a CEO, are important aspects of firm organization. These decisions are linked through the severance agreement, a key component of the contracts between a board and a manager. The severance agreement specifies payments to the manager upon his forced departure. Approximately 50% of the CEO compensation contracts implemented between 1994 and 1999 involved some form of severance agreement (Rusticus, 2006). The percentage of S&P firms that included a severance agreement in their CEO compensation contracts increased from 20% in 1993 to more than 55% in 2007 (Huang, 2011). In general, a contract with a severance agreement adds an explicit cost to the board's retention decision and makes replacement more difficult relative to a compensation contract without such an agreement.

A widely held belief is that CEOs are replaced too infrequently, or entrenched.[1] Entrenchment may arise for many reasons. For example, it may be an instance of governance failure in the form of a captive board of directors (Inderst and Mueller, 2010; Shleifer and Vishny, 1989; Hermalin and Weisbach, 1998) or a way to mitigate a moral hazard problem (Almazan and Suarez, 2003; Casamatta and Guembel, 2010; Manso, 2011). Taylor (2010) makes the first attempt to measure the cost of entrenchment using a structural model of CEO turnover and finds suggestive evidence of the opposite. In particular, he finds that boards in large firms fire CEOs with higher frequency than is optimal. We refer to this phenomenon as anti-entrenchment. This finding cannot be rationalized by the existing models on CEO turnover and thus calls for a new model to better understand the determinants of managerial turnover.

This paper investigates how optimal design of the severance agreement influences managerial entrenchment. A manager is said to be entrenched if the board retains an incumbent manager who has an expected ability lower than that of a replacement manager. Anti-entrenchment occurs when the board fires some managers with higher than average expected ability. We propose a two-period principal-agent model of manage-

[1] Although evidence shows forced CEO turnover is increasing over time and indicates boards are using more aggressive replacement policies, it is widely believed that CEOs are rarely fired and thus are entrenched. For instance, Kaplan and Minton (2012) find that board-driven turnover increased steadily from 10.93% (1992 – 1999) to 12.47% (2000 – 2007) using data from publicly traded Fortune 500 companies.

rial turnover and identify conditions that predict the emergence of entrenchment and anti-entrenchment. Formally, we consider a setup in which the first-period manager is incentivized by a contract that contains performance-related pay and severance pay. The firm's success depends on the initial manager's effort and the second-period manager's ability. Thus, the board faces an ability selection problem and a moral hazard problem. After the initial manager exerts effort, the board observes a non-contractible signal regarding his ability. The board can fire the initial manager by paying the severance pay specified in the contract and hire a replacement manager. Since the board's information about the initial manager's ability is non-contractible, it lacks commitment power and cannot write a contract that specifies a retention decision contingent on the signal. Severance pay is used as a costly device to provide commitment to not firing the initial manager. By committing to a high severance pay, the board ensures a low expected profit for itself after replacement, which leads to a less aggressive replacement policy. The board's optimal replacement policy balances incentive provision, manager selection and commitment.

Our main result characterizes the optimal replacement policy and shows how it depends on the precision of the signal of the manager's ability. When this monitoring technology is noisy, entrenchment is optimal. In such a scenario, the board places higher priority on motivating the incumbent manager to exert effort rather than on maximizing the manager's ability. Setting an aggressive replacement policy will fire the incumbent of high ability too often and dis-incentivize the incumbent to exert effort, while saving little on severance pay. As a result, a contract that induces entrenchment is optimal for the board.

Anti-entrenchment is optimal when the board's monitoring technology is sufficiently informative. The board is reluctant to provide commitment. On the one hand, a contract that favors the incumbent manager does not increase effort by much because of the low probability of replacement when the incumbent is of high ability. On the other hand, an aggressive replacement policy helps the board avoid paying the performance-related pay to the incumbent manager and can increase the firms profit. Thus, anti-entrenchment is optimal for the board. To the best of our knowledge, we are the first to study the interaction between the board's monitoring technology and managerial turnover, and to show that a contract with anti-entrenchment is sometimes optimal.

Our model can be applied to a variety of real-world settings. For example, the model can be used to analyze the turnover of founder CEOs in venture-capital-backed companies where the venture capitalist is a large shareholder and engages in active monitoring. It could also be used to analyze the contracts between head coaches and professional sports teams.

Related Literature: This paper belongs to the literature on the principal-agent model with replacement.[2] One strand of research views entrenchment as a potential source of inefficiency that the board aims to mitigate. Consequently, anti-entrenchment cannot be observed. Inderst and Mueller (2010) solve the optimal contract for the incumbent manager who holds private information on the firm's future performance and can avoid replacement by concealing bad information. Consequently, the optimal contract is designed to induce the incumbent to voluntarily step down when evidence suggests low expected profit under his management. Similarly, entrenchment occurs if the incumbent can make manager-specific investments to create cost of replacement to the board (Shleifer and Vishny, 1989) or if there exist close ties between the board and manager (Hermalin and Weisbach, 1998).

Another strand of research views entrenchment as a feature of the optimal contract (board structure) that helps overcome the moral hazard problem. Manso (2011) shows that tolerance for early failure (entrenchment) can be part of the optimal incentive scheme when motivating a manager to pursue more innovative business strategies is important to the board. Casamatta and Guembel (2010) study the optimal contract for the incumbent manager with reputational concern. In their model, entrenchment is optimal because the incumbent manager would like to see his strategy succeed and is less costly to motivate than the replacement manager. Almazan and Suarez (2003) study the optimal board structure for incentivizing the incumbent manager. They show that it can be optimal for shareholders to relinquish some power and choose a weak board, where the incumbent can veto his departure, rather than a strong board, where the board can fire the incumbent at will. In the same spirit, Laux (2008) studies the optimal degree of board independence for shareholders. He shows that some lack of independence can increase shareholder value. In these papers, boards (shareholders) provide better job security to the incumbent by making dismissal more difficult to

[2]See Laux (2014) for a comprehensive survey of the theoretical models on this topic.

induce more effort. Our paper contributes to the existing literature by pointing out that despite all the incentive-providing merits of entrenchment, the cost of incentivizing can be high when the board's monitoring technology is sufficiently informative.

In terms of modeling, the paper is most similar to Taylor and Yildirim (2011). They study the benefits and costs of different review policies and identify conditions under which the principal commits not to utilize the agent's information and chooses blind review as optimal policy. We apply their model to analyze managerial turnover by adding a contract stage to endogenize the agent's payoff and allowing the principal to replace the agent in the interim stage.

The remainder of the paper is organized as follows. Section 2 describes the model. Section 3 defines entrenchment and anti-entrenchment and characterizes the optimal contract. Section 4 studies the impact of informativeness on optimal replacement policy. Section 5 discusses extensions of the model. Section 6 concludes. All proofs are in the Appendix.

2 Model

There are two periods $t = 1, 2$ and an initial contract stage.

Contract stage. The board (principal), hires a manager (agent) from a pool with unknown ability $\theta_i \in \{0, 1\}$ to work for the firm with common prior $\Pr(\theta_i = 1) = \frac{1}{2}$.[3] The ability is unknown to both sides. The board offers a contract to the manager. We describe the contract details below.

Both the board and the managers are risk-neutral. Moreover, we assume that managers are protected by limited liability.[4] Finally, we assume the value of the outside option to the manager is 0. This assumption guarantees that the individual rationality (IR) constraint never binds and simplifies the analysis.

***Period* 1.** The manager exerts effort to create a project of quality q with cost $C(q) = \frac{1}{2}q^2$. Simultaneously, the board receives a signal $s \in \mathcal{S}$ of the manager's ability and decides whether to replace the incumbent manager. If the incumbent manager is

[3] The analysis is unchanged for a different prior of θ_i.

[4] This assumption is necessary because it excludes the possibility that the board sells the whole firm to the manager in order to provide the greatest possible incentive in the optimal contract.

fired, a replacement manager is hired and has ability θ_r randomly drawn from the same pool of managers.[5]

***Period* 2.** The manager who stays in office implements the project with no additional effort and payoffs are realized. Implementation is assumed to be costless and depends only on manager's ability.[6] To formalize this idea, we assume that the expected quality of the project is equal to $q\tilde{\theta}$, where q is the incumbent manager's choice of how much effort to exert and $\tilde{\theta}$ is the ability of the manager who stays in office at the beginning of period 2. With probability $q\tilde{\theta}$, the project is of high quality and yields outcome $y = 1$. With complementary probability $1 - q\tilde{\theta}$, the project is of low quality and yields outcome $y = 0$. After payoffs are realized, the incumbent manager receives payment according to the contract signed in period 0 and the game comes to an end.

In the optimal contract, the wage for low output is 0. A contract is defined by the tuple (w, k), where w is the wage rate when $y = 1$ and k is the severance pay to the incumbent manager if he is fired. By the limited liability assumption, $w \geq 0$ and $k \geq 0$.

Information structure. The board receives a noisy signal $s \in \mathcal{S}$ about incumbent manager's ability θ_i. s is drawn from distribution with cdf $F_{\theta_i}(\cdot)$ and pdf $f_{\theta_i}(\cdot)$ for $\theta_i \in \{0, 1\}$. Without loss of generality, we assume $\mathcal{S} = [0, 1]$ and normalize $s = \frac{1}{2}F_1(s) + \frac{1}{2}F_0(s)$ for $s \in [0, 1]$.[7] The two conditional density functions $\{f_1(s), f_0(s)\}$ suffice to define an information structure under such normalization. Three assumptions are imposed on the information structure.

Assumption 1 *The monotone likelihood ratio property (MLRP): $\frac{f_1(s)}{f_0(s)}$ is strictly increasing in s for $s \in [0, 1]$.*

For binary states, the MLRP assumption is without loss of generality because signals can always be relabeled according to likelihood ratio to satisfy this assumption.

[5]The project generation process can also be interpreted as a project selection process as in Casamatta and Guembel (2010). Assume some unknown state of the world $\eta \in [0, 1]$ is randomly drawn, and a manager is hired to select a project $a \in [0, 1]$ to match the underlying state. The quality of the project is 1 if $a = \eta$ and 0 otherwise. The manager incurs cost $C(q)$ to receive a signal ν of the true state. With probability q, the manager identifies η, that is, $\nu = \eta$, and with probability $1 - q$, ν is pure noise. Given q, the expected quality of the selected project is q. These two specifications lead to the same model.

[6]This assumption is relaxed in Section 5.2.

[7]This assumption is without loss of generality due to the fact any information structure can be normalized via integral probability transformation. See Appendix B for more details.

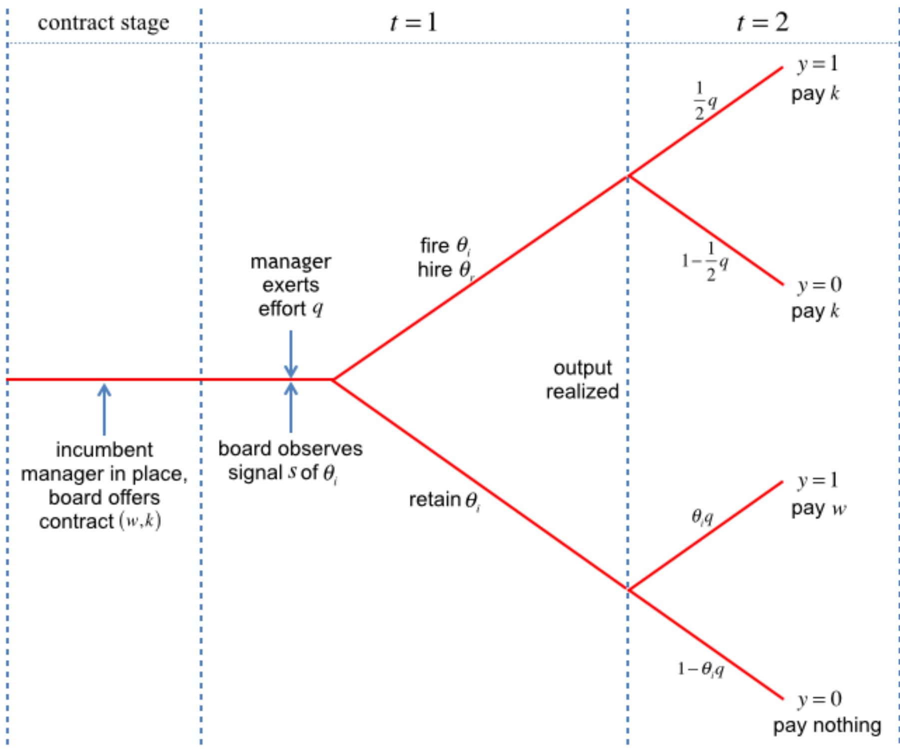

Figure 1: Timeline

Assumption 2 *Perfectly informative at extreme signals:* $\lim_{s\to 0} \frac{f_1(s)}{f_0(s)} = 0$ *and* $\lim_{s\to 1} \frac{f_1(s)}{f_0(s)} = +\infty$.

Assumption 2 guarantees that support of the posterior belief is always $[0,1]$. The last assumption imposed on the information structure is symmetry. This assumption allows us to define the first best replacement policy on the signal space.

Assumption 3 $f_1(s) = f_0(1-s)$ *for all* $s \in [0,1]$.

By Assumption 3, $f_1(\frac{1}{2}) = f_0(\frac{1}{2})$. Thus the likelihood ratio at $s = \frac{1}{2}$ is always 1 and the Bayesian update of the incumbent manager's ability at $\frac{1}{2}$ is equal to the prior.

Finally we introduce an index $\alpha \in (0,\infty)$ to parameterize the information structure. We assume that $f_{\theta_i}(s;\alpha)$ is continuous in s and α for $\theta_i \in \{0,1\}$ and define the information structures for the two extreme values of α as follows.

Assumption 4 (Completely informative/uninformative information structure)

1. *The information structure becomes completely uninformative when* $\alpha \to 0$, *i.e.,* $\lim_{\alpha\to 0}[f_0(s;\alpha) - f_1(s;\alpha)] = 0$ *for* $s \in (0,1)$.

2. *The information structure becomes completely informative when $\alpha \to \infty$, i.e., $\lim_{\alpha\to\infty} f_1(s;\alpha) = 0$ for $s \in [0, \frac{1}{2})$ and $\lim_{\alpha\to\infty} f_0(s;\alpha) = 0$ for $s \in (\frac{1}{2}, 1]$.*[8]

When the information structure becomes completely uninformative ($\alpha \to 0$), the two conditional density functions are the same. When the information structure becomes completely informative ($\alpha \to \infty$), the board will not observe a signal below $\frac{1}{2}$ when the incumbent manager is of high ability and a signal above $\frac{1}{2}$ when the incumbent manager's ability is low.

3 The Optimal Contract

3.1 The benchmark case: contractible effort

We first pin down the socially optimal replacement policy. By Assumption 1, the socially optimal replacement policy is a cutoff rule. Denote $\hat{s}$ as the signal cutoff.

Lemma 1 (First best cutoff) *Suppose the board can contract on effort q of the incumbent manager. Then the optimal effort is $q^{\mathcal{FB}} = \frac{1}{2} + \frac{1}{4}\left[F_0(\hat{s}^{\mathcal{FB}}) - F_1(\hat{s}^{\mathcal{FB}})\right]$, where the optimal replacement cutoff $\hat{s}^{\mathcal{FB}} = \frac{1}{2}$.*

When effort is contractible, the board is able to optimize effort and selection separately. Thus, there is no tradeoff between the moral hazard problem and the selection problem. It is optimal to replace the incumbent manager when the posterior belief about the incumbent's ability falls below the expected value of the pool and retain the incumbent otherwise. By Assumption 3, the likelihood ratio $\frac{f_1(s)}{f_0(s)}$ at $s = \frac{1}{2}$ is always equal to 1. Consequently, the Bayesian update of the incumbent manager's ability is always equal to the prior independent of the informativeness α of the information structure. Consequently, the socially optimal cutoff $\hat{s}^{\mathcal{FB}} = \frac{1}{2}$ for all α.

Given the first best cutoff, we can now define entrenchment. Denote (w^*, k^*) as the optimal contract to the board. Let $(\hat{s}^*, q^*)$ be the equilibrium replacement cutoff and effort of the continuation game induced by the optimal contract.

[8]Both completely informative and uninformative information structures are defined using pointwise convergence.

Definition 1 *We define* ***entrenchment*** *as a cutoff* $\hat{s}^* < \frac{1}{2}$ *and* ***anti-entrenchment*** *as* $\hat{s}^* > \frac{1}{2}$.

For the case where $\hat{s}^* = \frac{1}{2}$, we say that neither entrenchment nor anti-entrenchment is observed. The replacement policy coincides with the socially optimal policy. When $\hat{s}^* < \frac{1}{2}$, the replacement policy favors the incumbent manager: the board could have improved implementation by replacing the incumbent. Similarly, the replacement policy is considered aggressive and places the incumbent manager at a disadvantage when $\hat{s}^* > \frac{1}{2}$.

3.2 Characterizing the Optimal Contract

In this section, we solve the equilibrium outcome when effort is non-contractible. The board can only commit to the wage w and severance pay k in the contract. We are interested in the cutoff $\hat{s}^*$ induced by the optimal contract.

3.2.1 Incentives under fixed contract (w, k)

A contract (w, k) induces a simultaneous move game. We first solve the sub-game in period 1, fixing contract (w, k). The incumbent manager's effort q and the board's replacement policy $\hat{s}$ will be determined in a Cournot-Nash equilibrium.

For contract (w, k), the incumbent manager's best response to cutoff $\hat{s}$ is effort q that maximizes:

$$\max_q \frac{1}{2}\big[1 - F_1(\hat{s})\big]qw + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]k - C(q).$$

$$\Rightarrow q(\hat{s}; w, k) = \frac{1}{2}\big[1 - F_1(\hat{s})\big]w. \tag{1}$$

The board can provide incentive on effort by increasing wage w or lowering equilibrium cutoff $\hat{s}$. For a fixed contract (w, k) the board's best response to the incumbent manager's effort level q is cutoff $\hat{s}$ that maximizes:

$$\max_{\hat{s}} \frac{1}{2}\big[1 - F_1(\hat{s})\big]q(1 - w) + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\big(\frac{1}{2}q - k\big).$$

$$\Rightarrow \hat{s}(q; w, k) \text{ solves } \frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}q(1 - w) = \frac{1}{2}q - k. \tag{2}$$

Because a higher cutoff implies higher posterior belief about the incumbent manager's ability, the board chooses a cutoff such that the expected profit created by the marginal incumbent manager is equal to the expected profit under replacement in equilibrium.

Given contract (w,k), the optimal cutoff and effort $\big(\hat{s}(w,k), q(w,k)\big)$ are pinned down by equations (1) and (2). We can calculate the corresponding contract (w,k) that induces any tuple $(\hat{s}, q)$ as follows,

$$w(\hat{s}, q) = \frac{q}{\frac{1}{2}\big[1 - F_1(\hat{s})\big]} \tag{3}$$

and

$$k(\hat{s}, q) = \frac{1}{2}q - \frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})} q\big[1 - w(\hat{s}, q)\big].^{9} \tag{4}$$

3.2.2 Derive the optimal contract for fixed replacement policy

The board chooses contract (w,k) to maximize expected profit:

$$\max_{\{w,k\}} \frac{1}{2}\big[1 - F_1(\hat{s})\big] q(1-w) + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\big(\frac{1}{2}q - k\big)$$

s.t.

$$q = \frac{1}{2}\big[1 - F_1(\hat{s})\big] w$$

and

$$\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})} q(1-w) = \frac{1}{2}q - k.$$

Equivalently, the board is maximizing expected profit over $(\hat{s}, q)$, with $w(\hat{s}, q)$ and $k(\hat{s}, q)$ as determined in equations (3) and (4). Substituting equations (3) and (4) into the board's profit function yields expected profit as a function of $(\hat{s}, q)$,

$$q\left[1 - \frac{q}{\frac{1}{2}\big[1 - F_1(\hat{s})\big]}\right]\left\{\frac{1}{2}\big[1 - F_1(\hat{s})\big] + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\right\}.$$

It can be verified that $q = \frac{1}{4}\big[1 - F_1(\hat{s})\big]$ under the optimal contract. Consequently, $w^* = \frac{1}{2}$. We summarize the results of the previous pages in a lemma.

[9] The non-negativity assumption on k is not always satisfied for all $\hat{s}$ and q. We ignore this limited liability constraint for the moment and solve the unconstrained problem. This is not a big concern since it can be proved later that the optimal wage is $w^* = \frac{1}{2}$ and k is non-negative for all $\hat{s} \in [0,1]$.

Lemma 2 *Fixing $\hat{s}$, the board maximizes expected profit by offering a contract,*

$$w = \frac{1}{2}$$

and

$$k(\hat{s}) = \frac{1}{4}\big[1 - F_1(\hat{s})\big]\left[\frac{1}{2} - \frac{1}{2}\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\right].$$

Moreover, in equilibrium, the incumbent manager chooses effort

$$q(\hat{s}) = \frac{1}{4}\big[1 - F_1(\hat{s})\big].$$

By Lemma 2, $k(\hat{s})$ is decreasing in the equilibrium cutoff $\hat{s}$. By committing to a higher severance pay, the board chooses a lower replacement cutoff in equilibrium and is able to induce more effort. The expected profit can be rewritten in terms of $\hat{s}$ alone:

$$\pi(\hat{s}) := \frac{1}{8}\big[1 - F_1(\hat{s})\big]\left\{\frac{1}{2}\big[1 - F_1(\hat{s})\big] + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\right\}.$$

The optimal cutoff depends on the informativeness of the information structure. Rewrite the expected profit as follows:

$$\pi(\hat{s}) = \frac{1}{8}\underbrace{\Big[1 - F_1(\hat{s})\Big]}_{\text{incentive effect}}\Bigg\{\underbrace{\left[\frac{1}{2}\big[1 - F_1(\hat{s})\big] + \frac{1}{4}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\right]}_{\text{selection effect}}$$
$$+ \underbrace{\frac{1}{2}\Big[F_1(\hat{s}) + F_0(\hat{s})\Big]\left(\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})} - \frac{1}{2}\right)}_{\text{commitment effect}}\Bigg\}.$$

Three effects play a role in determining the optimal cutoff. Because the outcome depends on the expected ability of the manager in period 2, the board faces a selection problem. This is captured by $\left[\frac{1}{2}\big[1 - F_1(\hat{s})\big] + \frac{1}{4}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\right]$, which is called the *selection effect*. This is the expected ability of the manager in period 2. Increasing $\hat{s}$ will increase the expected ability of the manager in office when $\hat{s} < \frac{1}{2}$ and decrease the expected ability when $\hat{s} \geq \frac{1}{2}$. To optimize selection independently, the board would choose $\hat{s} = \frac{1}{2}$.

Because the outcome also depends on the effort choice of the incumbent manager, the board faces a moral hazard problem and needs to incentivize the incumbent. This is captured by $\left[1 - F_1(\hat{s})\right]$, which is referred to as the *incentive effect*. As the equilibrium replacement cutoff $\hat{s}$ increases, the incumbent manager expects a lower retaining probability in equilibrium and exerts less effort accordingly. The board provides more job security to better incentivize the incumbent manager in response. By this effect alone, the board sets $\hat{s} = 0$.

If the selection effect and the incentive effect were the only effects, a cutoff below $\frac{1}{2}$ is optimal to the board and entrenchment emerges under optimal contract. However, because the signal is non-contractible, board lacks commitment power on replacement policy. Severance pay serves as a costly commitment device that helps make replacement of the incumbent less likely. As the severance pay increases, it lowers the expected payoff of replacement, which creates a stronger incentive for the board to not replace the incumbent. In equilibrium the expected profit of replacement is equal to the expected profit created by the marginal incumbent manager. When board lowers the cutoff ($\hat{s} < \frac{1}{2}$) to provide more incentive on effort, it has to increase severance pay to make the equilibrium replacement policy credible. This generates a net loss compared to the first best replacement policy. It is captured by $\frac{1}{2}\left[F_1(\hat{s}) + F_0(\hat{s})\right]\left(\frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} - \frac{1}{2}\right)$, which is referred to as the *commitment effect*. Compared to the first best cutoff $\hat{s} = \frac{1}{2}$, the board obtains a net commitment gain by providing less commitment and designing a contract that induces cutoff above $\frac{1}{2}$. Similarly, the board suffers a commitment loss by committing to a cutoff that is below $\frac{1}{2}$. The net commitment effect is shown by $\left(\frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} - \frac{1}{2}\right)$. Multiplied by the probability of replacement, this yields the total net commitment gain/loss. By this effect alone, the board sets $\hat{s} = 1$.

If incentive effect dominates commitment effect, entrenchment is optimal to the board. Otherwise, anti-entrenchment is optimal.

4 The Optimal Replacement Policy

In this section, we study how the optimal replacement policy varies depending on the informativeness of the board's monitoring technology.

4.1 Replacement at limiting distribution

Proposition 1 *Suppose* $\{f_1(\cdot;\alpha), f_0(\cdot;\alpha)\}$ *satisfies Assumptions 1 - 4. Then there exist* $\overline{\alpha}$ *and* $\underline{\alpha}$ *such that,*

1. $\hat{s}^*(\alpha) > \frac{1}{2}$ *for* $\alpha > \overline{\alpha}$*;*

2. $\hat{s}^*(\alpha) < \frac{1}{2}$ *for* $\alpha < \underline{\alpha}$*.*

When the information structure is noisy, providing incentives is more profitable than obtaining more commitment. Choosing $\hat{s} > \frac{1}{2}$ reduces severance pay by a small amount because the Bayesian update around $\hat{s} = \frac{1}{2}$ changes very slowly and the expected ability of the incumbent manager at the cutoff is close to the expected ability at a cutoff of $\frac{1}{2}$. On the other hand, choosing $\hat{s} > \frac{1}{2}$ reduces the incumbent manager's incentive to exert effort. Consequently, it is optimal for the board to design a contract that leads to entrenchment.

When the board's monitoring technology is sufficiently informative, the benefit of commitment dominates and choosing $\hat{s} < \frac{1}{2}$ is not optimal for the board. Since the probability of firing a high ability manager is very small for all signals below $\frac{1}{2}$, lowering the equilibrium replacement cutoff does not have a large effect on the incumbent's effort. On the other hand, it is easy to obtain commitment gain. The expected ability of the manager in the right neighborhood of $\frac{1}{2}$ is very close to 1 when the information structure is sufficiently informative. That is, the board can largely reduce the severance pay by choosing a cutoff slightly above $\frac{1}{2}$. Thus, anti-entrenchment is optimal to the board.

4.2 Optimal replacement and informativeness

Proposition 1 does not characterize the equilibrium replacement policy for intermediate α. To do this, it is necessary to introduce an information order.

4.2.1 Distribution of posterior beliefs

Denote $p = \varphi(s)$ as the posterior belief of θ after observing signal s. Then $\varphi(s) = \frac{f_1(s)}{f_1(s)+f_0(s)}$. By Assumption 1, $\varphi(s)$ is strictly increasing in s. By Assumption 2, the support of p is $[0,1]$. Denote $g(p)$ as the corresponding density function. Since $\mathbb{E}(\mathbb{E}(\theta|s)) = \frac{1}{2}$, the only constraint we impose on $g(\cdot)$ is that $\int_0^1 pg(p)dp = \frac{1}{2}$.

Given an information structure $\left\{f_1(\cdot), f_0(\cdot)\right\}$, the density function of posterior belief p can be calculated as follows:

$$g(p) = \frac{1}{2}\left[f_1(\varphi^{-1}(p)) + f_0(\varphi^{-1}(p))\right]\frac{d\varphi^{-1}(p)}{dp}.$$

Lemma 3 *For any density function $g(\cdot)$ with support $[0,1]$ that satisfies $\int_0^1 pg(p)dp = \frac{1}{2}$, there exists a unique information structure $\left\{f_1(\cdot), f_0(\cdot)\right\}$ that induces $g(\cdot)$.*

By Lemma 3, there exists a one-to-one mapping between $g(\cdot)$ and information structure $\left\{f_1(\cdot), f_0(\cdot)\right\}$. Thus working on the information structure $\left\{f_1(\cdot), f_0(\cdot)\right\}$ is equivalent to working on distribution of the posterior belief $g(\cdot)$. Consequently, we can define information order on $g(\cdot)$. By Assumption 3 on $\left\{f_1(\cdot), f_0(\cdot)\right\}$, $g(p) = g(1-p)$ and $G(p) = 1 - G(1-p)$ for $p \in [0,1]$. Thus, it suffices to order different information structures based on $G(p)$ for $p \in [0, \frac{1}{2}]$.

4.2.2 The ρ-concave order

We use ρ-concavity to define the informativeness of the information structure.[10] To the best of our knowledge, this is the first paper that defines information order using ρ-concavity.

Given $G(\cdot)$, define local ρ-concavity at p as,

$$\rho(p) := 1 - \frac{G(p)g'(p)}{g^2(p)}.$$

By definition, $\rho(p)$ is the power of $G(\cdot)$ such that the second order Taylor expansion at p drops out. Thus, $\rho(p)$ is a measure of the concavity of $G(\cdot)$ at point p. Log-concavity is equivalent to $\rho(p) \geq 0$ and concavity is equivalent to $\rho(p) \geq 1$. We focus on the distributions such that $\rho(p) \in (0, \infty)$. This assumption is a necessary condition to guarantee the initial condition $G(0) = 0$ is satisfied.[11]

[10] For more applications of ρ-concavity in economics, see Mares and Swinkels (2014) on auction theory; Anderson and Renault (2003), Weyl and Fabinger (2013) on industry organization.

[11] Imposing this non-negativity assumption on $\rho(\cdot)$ is without loss of generality: a completely uninformative information structure can still be defined under this constraint.

Definition 2 (ρ-concave order) *$G_1(p)$ is said to be more informative than $G_2(p)$ in the ρ-concave order if $\rho(p|G_1) > \rho(p|G_2)$ for all $p \in [0, \frac{1}{2}]$.*

By definition, $G_1(p)$ is more informative than $G_2(p)$ if $G_1(p)$ is everywhere more concave than $G_2(p)$ measured by local ρ-concavity. The ρ-concave order is a stronger condition than the rotation order and Blackwell's order: if a family of distributions is ordered according to the ρ-concave order, then it is rotation-ordered and ordered in the sense of Blackwell.[12]

Assume that $\max_{p\in[0,\frac{1}{2}]} \{\rho(p;\alpha)\}$ and $\max_{p\in[0,\frac{1}{2}]} \{\rho(p;\alpha)\}$ exist for all $\alpha \in (0, \infty)$. Denote $\overline{\rho}(\alpha) = \max_{p\in[0,\frac{1}{2}]} \{\rho(p;\alpha)\}$ and $\underline{\rho}(\alpha) = \min_{p\in[0,\frac{1}{2}]} \{\rho(p;\alpha)\}$ for notational convenience.

Lemma 4 *Suppose $0 < \underline{\rho} \leq \overline{\rho} < \infty$. Then $\frac{1}{2}(2p)^{\frac{1}{\underline{\rho}}} \leq G(p) \leq \frac{1}{2}(2p)^{\frac{1}{\overline{\rho}}}$ for $p \in [0, \frac{1}{2}]$.*

By Lemma 4, $G(p)$ can be bounded by two constant cumulative density functions with constant ρ-concavity. A completely informative information structure corresponds to the case where $\lim_{\alpha\to\infty} \underline{\rho}(\alpha) = \infty$ and a completely uninformative information structure is equivalent to $\lim_{\alpha\to\infty} \overline{\rho}(\alpha) = 0$.[13] The following assumptions are imposed on the family of distribution $\{G(\cdot;\alpha)\}$ indexed by $\alpha \in (0, \infty)$.

Assumption 5 *(a) Log concavity: $\rho(p;\alpha) \in (0, \infty)$ for $(p, \alpha) \in [0, \frac{1}{2}] \times (0, \infty)$.*

(b) ρ-concave order: If $\alpha_1 > \alpha_2$, $\rho(p;\alpha_1) > \rho(p;\alpha_2)$ for $p \in [0, \frac{1}{2}]$.

(c) Regularity 1: $\forall\alpha$, $\rho(p;\alpha)$ is weakly decreasing in p for $p \in [0, \frac{1}{2}]$.[14]

(d) Regularity 2: There exists α such that $\rho(p;\alpha) = 1$ for all $p \in [0, \frac{1}{2}]$.

(e) Normalization: $\lim_{\alpha\to\infty} \underline{\rho}(\alpha) = \infty$ and $\lim_{\alpha\to 0} \overline{\rho}(\alpha) = 0$.

By Assumption 5(a), we focus on $G(p;\alpha)$ which is log-concave in $p \in [0, \frac{1}{2}]$. Together with Assumption 5(c), Assumption 5(d) guarantees that the concavity/convexity of $G(\cdot)$ will not change for given α. Assumption 5(e) restates Assumption 4 in the language of the ρ-concavity.

[12] See Appendix C for more details.

[13] See Appendix C for detailed proof.

[14] As will be clear later, this assumption generates a well-behaved profit function for $p \in [0, \frac{1}{2}]$.

4.2.3 The optimal replacement policy

Denote $\hat{p}$ as the cutoff of the posterior belief and $\tilde{\pi}(\hat{p})$ as the board's profit as a function of $\hat{p}$. Then

$$\tilde{\pi}(\hat{p}) = \frac{1}{4} \underbrace{\int_{\hat{p}}^{1} tg(t)dt}_{\text{incentive effect}} \left\{ \underbrace{\frac{1}{2}G(\hat{p}) + \int_{\hat{p}}^{1} tg(t)dt}_{\text{selection effect}} + \underbrace{(\hat{p} - \frac{1}{2}) \int_{0}^{\hat{p}} g(t)dt}_{\text{commitment effect}} \right\}.$$

The profit function can be further simplified by combining the selection effect and the commitment effect,

$$\tilde{\pi}(\hat{p}) = \frac{1}{4} \underbrace{\int_{\hat{p}}^{1} tg(t)dt}_{\text{incentive effect}} \left\{ \underbrace{\int_{\hat{p}}^{1} tg(t)dt \;+\; \hat{p}G(\hat{p})}_{\text{selection+commitment effect}} \right\}.$$

The expression of the total selection and commitment effect is intuitive. In equilibrium, the board's expected profit of replacement is equal to the expected profit created by the marginal incumbent manager with expected ability $\hat{p}$. Hence the board is replacing the incumbent manager of ability $p \leq \hat{p}$ with $\hat{p}$ taking commitment into consideration. It can be verified that the total of the selection effect and the commitment effect is increasing in $\hat{p}$ and thus is maximized at $\hat{p} = 1$.

The first order derivative with respect to $\hat{p}$ yields,

$$\tilde{\pi}'(\hat{p}) = \frac{1}{4} \left[-\hat{p}g(\hat{p}) \left(1 - \int_{\hat{p}}^{1} G(t)dt \right) + G(\hat{p}) \int_{\hat{p}}^{1} tg(t)dt \right].$$

$$\Rightarrow \; \tilde{\pi}'(\hat{p}) \lesseqqgtr 0 \;\Leftrightarrow\; \frac{\hat{p}g(\hat{p})}{G(\hat{p})} \gtreqqless \frac{\int_{\hat{p}}^{1} tg(t)dt}{1 - \int_{\hat{p}}^{1} G(t)dt}.$$

From the first order condition, $\hat{p}g(\hat{p})$ is the marginal incentive effect and $G(\hat{p})$ is the marginal selection plus commitment effect. Whether profit is increasing or decreasing in $\hat{p}$ largely depends on the ratio of these two marginal effects, which is also the elasticity of $G(\cdot)$ at point $\hat{p}$. Since $\tilde{\pi}(1) = 0$, the incentive effect dominates the selection plus commitment effect when $\hat{p}$ is close to 1. To relate ρ-concavity to the profit function, notice that $\frac{\hat{p}g(\hat{p})}{G(\hat{p})} = \left(\frac{\int_0^{\hat{p}} \rho(t)dt}{\hat{p}}\right)^{-1}$, which is the inverse of the average ρ-concavity of $G(\cdot)$ from 0 to $\hat{p}$. This ratio is weakly increasing if $\rho(p;\alpha)$ is weakly decreasing in

p for $p \in [0, \frac{1}{2}]$ by Assumption 5(c). This assumption guarantees that the marginal incentive effect changes faster than the marginal selection plus commitment effect and yields a well-behaved profit function for $\hat{p} \in [0, \frac{1}{2}]$. Assumption 5(b) ($\rho$-concave order) guarantees that the marginal selection plus commitment effect changes faster than the marginal incentive effect for given $\hat{p} \in [0, \frac{1}{2}]$ as α increases. Consequently, the selection plus commitment effect takes over as board's monitoring technology improves and anti-entrenchment is more likely to emerge.

Proposition 2 *Suppose the family of distribution* $\{G(\cdot;\alpha)\}$, *indexed by* $\alpha \in (0, \infty)$, *satisfies Assumption* 5. *Then there exists* α_1 *and* α_2 *such that*

1. $\hat{s}^*(\alpha) = 0$ *for* $\alpha \in (0, \alpha_1]$;
2. $\hat{s}^*(\alpha) \in (0, \frac{1}{2})$ *for* $\alpha \in (\alpha_1, \alpha_2)$;
3. $\hat{s}^*(\alpha) \in (\frac{1}{2}, 1)$ *for* $\alpha \in (\alpha_2, \infty)$,

where α_1 *satisfies* $\rho(p;\alpha_1) = 1 \ \forall\, p \in [0, \frac{1}{2}]$ *and* $\alpha_2 > \alpha_1$.

Proposition 2 characterizes the optimal replacement policy for all α. When α is small, the board provides full job security and never fires the incumbent manager. When α is moderate, the board replaces the incumbent manager less frequently than the socially optimal level and entrenchment is optimal. When α is large, the board uses an aggressive replacement policy and anti-entrenchment emerges. There exists a clear cutoff between entrenchment and anti-entrenchment: once anti-entrenchment is optimal for informativeness level α', the optimal replacement policy is never entrenchment under a more informative information structure $\alpha > \alpha'$.

4.2.4 A tractable example

Example 1 *Suppose* $G(p)$ *has the following functional form,*

$$G(p) = \begin{cases} \frac{1}{2}(2p)^{\frac{1}{\alpha}} & \text{for} \quad \hat{p} \in [0, \frac{1}{2}] \\ 1 - \frac{1}{2}[2(1-p)]^{\frac{1}{\alpha}} & \text{for} \quad \hat{p} \in (\frac{1}{2}, 1] \end{cases} .$$

Then the optimal cutoff is

1. *for* $\alpha \leq 1$, $\hat{p}^*(\alpha) = 0$;

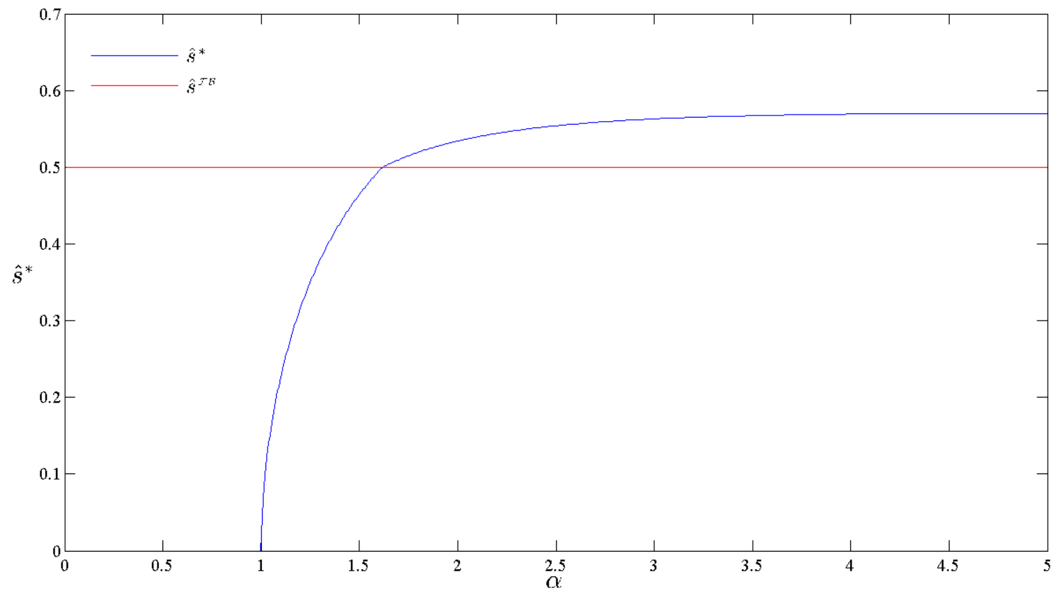

Figure 2: The optimal replacement policy

2. *for* $1 < \alpha < \frac{\sqrt{5}+1}{2}$*,* $\hat{p}^* \in (0, \frac{1}{2})$*;*

3. *for* $\alpha > \frac{\sqrt{5}+1}{2}$*,* $\hat{p}^* \in (\frac{1}{2}, 1)$*.*

Given $G(\cdot)$, the two corresponding conditional density functions are

$$f_1(s) = \begin{cases} (2s)^\alpha & \text{for} \quad s \in [0, \frac{1}{2}] \\ 2 - \left[2(1-s)\right]^\alpha & \text{for} \quad s \in (\frac{1}{2}, 1] \end{cases}$$

and

$$f_0(s) = \begin{cases} 2 - (2s)^\alpha & \text{for} \quad s \in [0, \frac{1}{2}] \\ \left[2(1-s)\right]^\alpha & \text{for} \quad s \in (\frac{1}{2}, 1] \end{cases} .$$

Figure 2 shows the optimal cutoff for different informativeness levels of the monitoring technology. Turnover is increasing for $\alpha \in \left[1, \frac{\sqrt{5}+1}{2}\right]$ when the manager is entrenched. The relationship between turnover and the informativeness of the board's monitoring technology is an inverted-U shape for $\alpha > \frac{\sqrt{5}+1}{2}$. As α approaches infinity, the optimal cutoff converges to $\frac{1}{2}$.

Figure 3 shows that severance pay in the optimal contract is decreasing in the informativeness of the board's monitoring technology. When the information structure

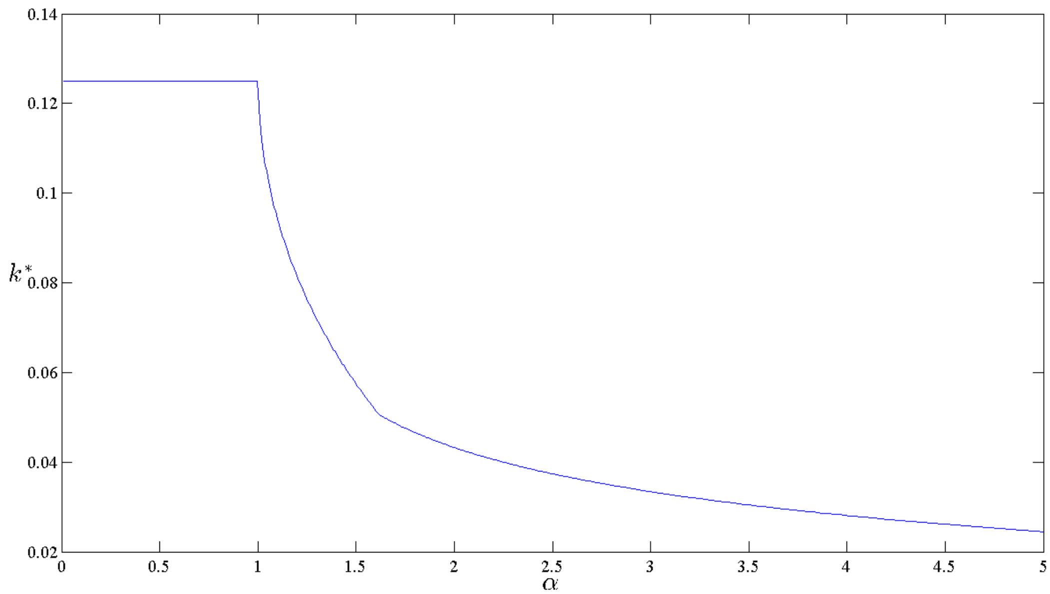

Figure 3: Severance pay in the optimal contract

becomes more informative, it is easier for the board to obtain net commitment gain. Thus the board is less willing to commit to not replacing the incumbent manager and the size of severance pay offered in the optimal contract decreases as a result. This generates a testable implication of the model: the size of the severance package is decreasing in the informativeness of the board's monitoring technology.

4.3 Discussion

An optimal replacement policy that differs from the first best stems from two important assumptions: the signal is non-contractible and severance pay is constant with respect to outcome, i.e., the board cannot provide performance-based severance pay. Without either of these assumptions, neither entrenchment nor anti-entrenchment emerges: the optimal replacement policy is always $\hat{s}^* = \frac{1}{2}$.

First consider what happens if the board's signal is contractible, while maintaining the assumption that severance pay is constant. A contract is fully characterized by $\{w(s), r(s), k(s)\}$, where $s \in [0, 1]$. $\{w(s), k(s)\}$ is the promised wage and severance pay after signal s. $r(s) \in [0, 1]$ specifies the retaining probability of the incumbent manager at signal s. In particular, $r(s) = 1$ indicates that the incumbent manager is retained while $r(s) = 0$ indicates that the incumbent is fired.[15]

Proposition 3 *Suppose that the signal is contractible and severance pay is constant with respect to outcome. Then* $k^*(s) = 0$. *Moreover,* $r^*(s) = 1$ *for* $s \in [\frac{1}{2}, 1]$ *and* $r^*(s) = 0$ *for* $s \in [0, \frac{1}{2})$.

Allowing the board to contract on signals gives the board commitment power on its retention decision at no cost. Severance pay is a costly commitment device, and is no longer used in the optimal contract.

The board can design a contract to induce any effort level without deviating from the socially optimal replacement cutoff. The manager is risk-neutral and only cares about the expected wage. Thus, the board can incentivize the incumbent manager by increasing the expected wage payment, which is determined by both the wage function $w(s)$ and the replacement policy $r(s)$. For given effort q and replacement policy $r(s)$, the board can adjust the wage function $w(s)$ to induce q without changing $r(s)$. That is, the board can optimize effort and selection separately if the signal is contractible, and the replacement cutoff is equal to $\frac{1}{2}$ in the optimal contract.

Next consider what happens when the board can condition the severance pay on the outcome, while maintaining the assumption that the signal is non-contractible. A contract is in the form of a triple (w_1, w_2, k). w_1 is the wage rate when the incumbent manager stays as in the baseline model. The tuple (w_2, k) constitutes a severance package. w_2 is the payment to the incumbent manager if he is forced out and $y = 1$. k is the constant severance pay as in the baseline model.

Proposition 4 *Suppose that the signal is non-contractible and the board can provide performance-based severance pay. Then* $k^* = 0$, $w_1^* = w_2^*$ *and* $\hat{s}^* = \frac{1}{2}$.

Constant severance pay is less effective to the board than performance-based severance pay when the incumbent manager is no longer in office because a lump-sum payment rewards failure. Thus, only performance-based severance pay is employed in the optimal contract.

[15] Due to the board's risk neutrality, randomization is not optimal except for the case where the board is indifferent between retaining and firing the incumbent manager, in which we assume the incumbent manager is retained with probability 1.

Again the board has no incentive to deviate from the first best cutoff. Due to manager's risk neutrality, the effort choice of the incumbent manager is only determined by the expected wage. For a given effort level q that the board wants to motivate, the expected wage is fixed, which is also the total cost to hire the incumbent manager. Since q is fixed, it remains to maximize the expected ability of the manager who stays in office in period 2. Hence the replacement cutoff stays at the first best in the optimal contract.[16]

5 Extensions

In this section, we show that the main result of the optimal replacement policy is robust to several different specifications.

5.1 More effort vs. better selection

It is interesting to study whether the main result on entrenchment (anti-entrenchment) remains optimal when effort becomes more important than ability. Anti-entrenchment is less likely to emerge when selection becomes less important. We model this by decreasing the variance of the manager's ability or increasing the importance of effort relative to ability in the success probability.

Assume the ability space is $\theta \in \left\{\frac{1}{2} - \delta, \frac{1}{2} + \delta\right\}$, where $\delta \in \left(0, \frac{1}{2}\right]$ and the success probability is equal to $q^{1+\tau}\theta$, where $\tau \in (-1, 1)$. δ is a measure of the variance of manager's ability ex ante while τ is a measure of the relative importance of effort compared to selection. The baseline model corresponds to $(\delta, \tau) = (\frac{1}{2}, 0)$. It is intuitive that entrenchment remains optimal as α approaches 0. Thus we focus on the case where α approaches infinity.

Proposition 5 (Comparative static)

1. *if* $\delta > \frac{1}{2}^{\frac{1-\tau}{2}} - \frac{1}{2}$, *there exists* $\bar{\alpha}_A$ *such that anti-entrenchment is optimal for* $\alpha > \bar{\alpha}_A$;
2. *if* $\delta < \frac{1}{2}^{\frac{1-\tau}{2}} - \frac{1}{2}$, *there exists* $\bar{\alpha}_E$ *such that entrenchment is optimal for* $\alpha > \bar{\alpha}_E$.

[16] In practice, a severance package usually comes in the form of a combination of a lump-sum payment and a stock option. This can be rationalized by the assumption that the manager is more risk-averse than the board. If that is the case, the optimal contract will involve some degree of lump-sum payment in response to risk sharing.

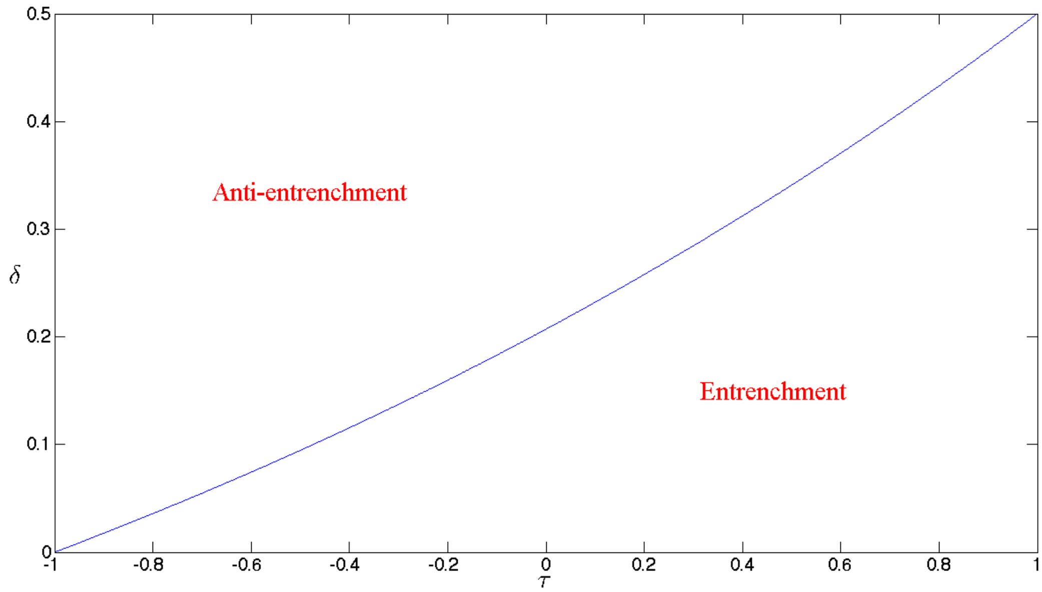

Figure 4: The optimal replacement policy at the limiting distribution

Fixing τ, selection becomes more important as the variance of the manager's ability increases. The intuition from the baseline model applies when δ is large, and anti-entrenchment emerges in the optimal contract. When δ is small, the marginal productivity of a low ability manager is close to that of a high ability manager. Since motivating the low ability manager is also important, the optimal contract leads to entrenchment as $\alpha \to \infty$.

5.2 Costly execution

Suppose the outcome depends on the incumbent manager's effort q in period 1, effort e of the manager in period 2 as well as the ability of the manager in office in period 2. Effort q can be interpreted as the project quality of the project selected by the incumbent manager and e can be interpreted as the effort required to execute the project. Now the board needs to also offer a contract to the replacement manager. Moreover, the optimal contract with the incumbent manager must balance a two-dimensional moral hazard problem as well as the selection problem.

Assume the success probability is equal to $\tilde{\theta}\big[(1-\lambda)q + \lambda\tilde{e}\big]$, where $\tilde{\theta} \in \{0,1\}$ is the manager's ability in period 2, $q \in [0,1]$ is the effort of the incumbent manager in period 1 and $\tilde{e} \in [0,1]$ is the effort of the manager in period 2. Period 1 effort q and period 2 effort e are substitutes and $\lambda \in [0,1]$ measures the relative importance of period 1 effort. When $\lambda = 0$ the model simplifies to the baseline model. After the board makes a retention decision, the manager in office at the beginning of period 2 exerts effort e to execute the project. The cost function to the incumbent manager is assumed to be separable and quadratic, i.e., $C_i(q,e) = \frac{1}{2}q^2 + \frac{1}{2}e^2$. The cost function to the replacement manager is assumed to be $C_r(q,e) = C_i(0,e)$.

Lemma 5 (First best outcome with costly execution) *The socially optimal cutoff is equal to* $\frac{1}{2}$.

The proof is similar to Lemma 1 and is omitted. With costly execution, the first best replacement cutoff remains at $\frac{1}{2}$. In fact, this result is very general. The optimal cutoff is always $\frac{1}{2}$ as long as the marginal impact of manager's ability is positive.

The board provides two contracts, contract (w,k) to the incumbent manager and a wage w_r to the replacement manager. Denote variables with subscript r as the variables related to the replacement manager after retention.

We first calculate the optimal contract with the replacement manager w_r after the incumbent's departure for a given belief of period 1 effort q. Given q and w_r, the replacement manager chooses e_r to maximize:

$$\frac{1}{2}\big[(1-\lambda)q + \lambda e_r\big]w_r - \frac{1}{2}e_r^2 \Rightarrow e_r(w_r) = \frac{1}{2}\lambda w_r.$$

Note that the replacement manager's effort on execution is independent of q. This is because q and e are assumed to be substitutes. The board chooses w_r to maximize:

$$\frac{1}{2}\big[(1-\lambda)q + \lambda e_r\big](1-w_r) \Rightarrow w_r^* = \max\Big\{\frac{1}{2} - \frac{1-\lambda}{\lambda^2}q, 0\Big\}.$$

Since q and e are substitutes by assumption, the optimal wage to the replacement manager is weakly decreasing in the belief about q. When first period effort q is large or λ is small, the board provides a contract with $w_r = 0$ to the replacement manager.

Let $\underline{\pi}(q)$ be the board's expected profit under optimal contract after replacement. $\underline{\pi}(q)$ can be calculated as follows,

$$\underline{\pi}(q) = \begin{cases} \frac{1}{4}\left(\frac{1}{2}\lambda + \frac{1-\lambda}{\lambda}q\right)^2 & \text{for} \quad q \leq \frac{1}{2}\frac{\lambda^2}{1-\lambda} \\ \frac{1}{2}(1-\lambda)q & \text{for} \quad q > \frac{1}{2}\frac{\lambda^2}{1-\lambda} \end{cases}.$$

Next we calculate the equilibrium for a given contract (w, k) with the incumbent manager. For a fixed contract (w, k) and belief about cutoff $\hat{s}$, the incumbent manager chooses (q, e) to maximize:

$$\frac{1}{2}[1 - F_1(\hat{s})]\big[(1-\lambda)q + \lambda e\big]w + \frac{1}{2}[F_1(\hat{s}) + F_0(\hat{s})]k - \frac{1}{2}q^2 - \frac{1}{4}[(1 - F_1(\hat{s})) + (1 - F_0(\hat{s}))]e^2.$$

$$\Rightarrow q = (1-\lambda)\frac{1 - F_1(\hat{s})}{2}w \quad \text{and} \quad e = \lambda\frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}w.$$

Note that q is decreasing in $\hat{s}$ while e is increasing in $\hat{s}$. A higher equilibrium replacement cutoff leads to a lower first period effort q and higher second period effort e by the incumbent manager. The first period effort q is decreasing in $\hat{s}$ because a higher cutoff implies lower retaining probability in period 2 and dis-incentivizes the incumbent manager as in the baseline model. Conditional on the fact that the incumbent manager is retained, a higher cutoff yields a higher estimate of the incumbent's ability and thus the incumbent is willing to exert more effort in period 2 (θ and e are assumed to be compliments). As a result, e is increasing in equilibrium cutoff $\hat{s}$. Similarly to Casamatta and Guembel (2010), the incumbent is easier to motivate, but for different reasons. The incumbent manager is easier to motivate in Casamatta and Guembel (2010) due to his reputational concern while in our model it is due to the incumbent's learning of his ability. For a given wage rate w, the replacement manager chooses $e_r = \frac{1}{2}\lambda w$ while the incumbent manager chooses $e = \lambda\frac{1-F_1(\hat{s})}{[1-F_1(\hat{s})]+[1-F_0(\hat{s})]}w > \frac{1}{2}\lambda w$. The incumbent manager learns from his retention that his ability is above average. Since ability and effort are assumed to be compliments, a higher estimate of ability implies a higher marginal return on effort. Thus, the incumbent manager exerts more effort in period 2 than the potential replacement manager given the same wage.

For a fixed contract (w, k) and belief about effort (q, e), the board chooses cutoff $\hat{s}$ to maximize:

$$\frac{1}{2}[1 - F_1(\hat{s})]\big[(1-\lambda)q + \lambda e\big](1 - w) + \frac{1}{2}[F_1(\hat{s}) + F_0(\hat{s})]\left[\underline{\pi}(q) - k\right].$$

$$\Rightarrow \frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})}\big[(1-\lambda)q+\lambda e\big](1-w)=\underline{\pi}(q)-k.$$

Not every $\hat{s}$ can be implemented. For instance, $\hat{s}$ very close to 1 cannot be induced by a contract. This is due to the limited liability assumption of severance pay. Extremely high cutoff can only be induced if severance pay is allowed to be negative. This is different from the baseline model. Using an aggressive replacement policy results in a small q, making the board's outside option very unattractive. On the other hand, an aggressive replacement policy improves learning and makes e very high, increasing the value of keeping the current manager. Thus, unless the board is compensated by negative severance pay, a very aggressive replacement policy cannot be induced by a contract subject to the limited liability constraint.

Similarly to the baseline model, the expected profit can be written as a function of cutoff $\hat{s}$ alone, assuming away the limited liability constraint of k,[17]

$$\pi(\hat{s})=\frac{1}{8}\Bigg\{(1-\lambda)^2\underbrace{\frac{1-F_1(\hat{s})}{2}}_{\text{incentive effect}}+\lambda^2\underbrace{\frac{1-F_1(\hat{s})}{[1-F_1(\hat{s})]+[1-F_0(\hat{s})]}}_{\text{learning effect}}\Bigg\}$$
$$\cdot\Bigg\{\underbrace{\Big[[1-F_1(\hat{s})]+\frac{1}{2}\big[F_1(\hat{s})+F_0(\hat{s})\big]\Big]}_{\text{selection effect}}+\underbrace{\Big[F_1(\hat{s})+F_0(\hat{s})\Big]\Big(\frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})}-\frac{1}{2}\Big)}_{\text{commitment effect}}\Bigg\}.$$

Proposition 6 (Optimal replacement policy with costly execution) *1. If $\lambda\in[0,\sqrt{2}-1)$, there exists $\overline{\alpha}$ such that $\hat{s}^*(\alpha)>\frac{1}{2}$ for $\alpha>\overline{\alpha}$.*

2. For $\lambda\in[0,1]$, there exists $\underline{\alpha}$ such that $\hat{s}^*(\alpha)<\frac{1}{2}$ for $\alpha<\underline{\alpha}$.

A learning effect enters into the board's profit function along with the three aforementioned effects. When the information structure is sufficiently noisy ($\alpha\to 0$), the incumbent's learning is very slow for all $s\in(0,1)$. The learning effect plays a minor role in determining the optimal replacement policy since $\frac{1-F_1(\hat{s})}{[1-F_1(\hat{s})]+[1-F_0(\hat{s})]}\lambda^2$ can be considered as a constant. Thus, entrenchment is expected to be optimal when α is sufficiently small independent of the size of λ.

When the information structure is sufficiently informative ($\alpha\to\infty$), the incum-

[17]The intuition can be clearly illustrated assuming that every $\hat{s}$ can be induced. The limited liability constraint of severance pay is taken into consideration in the proof of Proposition 6.

bent's learning becomes very fast. When execution becomes sufficiently important, entrenchment can be optimal to the board. When period 1 effort q is sufficiently important relative to period 2 effort e (i.e., $\lambda < \sqrt{2} - 1$), the incentive effect is more important than the learning effect in board's contractual problem. Thus, the main insight in the baseline model follows through and entrenchment is expected to emerge in the optimal contract.

5.3 Signal of outcome instead of ability

In the baseline model, it is assumed that the board observes a signal of the incumbent manager's ability rather than the outcome under the incumbent's management. Since the signal is not related to the incumbent manager's effort, the incumbent cannot increase his probability of retention by exerting more effort. When the board receives a signal related to effort, the incumbent manager is able to increase his probability of being retained by exerting more effort.

Suppose for outcome $y \in \{0, 1\}$, signal s is drawn from a distribution with density $h_y(\cdot)$ and cdf $H_y(\cdot)$. Similarly, we assume $\{h_1(\cdot), h_0(\cdot)\}$ satisfies Assumptions 1 - 4. The signal provides information about the expected outcome and the incumbent manager's ability.

The social planner chooses $(\hat{s}, q)$ to maximize:

$$\max_{\{\hat{s},q\}} \frac{1}{2}q\left[1 - H_1(\hat{s})\right] + \frac{1}{2}q\left[\frac{1}{2}qH_1(\hat{s}) + \left(1 - \frac{1}{2}q\right)H_0(\hat{s})\right] - \frac{1}{2}q^2.$$

Lemma 6 (First best outcome) *$\hat{s}^{\mathcal{FB}} = \frac{1}{2}$ and $q^{\mathcal{FB}} = \frac{1+H_0(\frac{1}{2})-H_1(\frac{1}{2})}{2+H_0(\frac{1}{2})-H_1(\frac{1}{2})}$ in the first best outcome.*

The proof is similar to Lemma 1 and is omitted. Given effort level q, the Bayesian update of the incumbent manager's ability is

$$\varphi(\hat{s}, q) = \frac{\frac{1}{2}qh_1(\hat{s}) + \frac{1}{2}(1 - q)h_0(\hat{s})}{\frac{1}{2}qh_1(\hat{s}) + \left(1 - \frac{1}{2}q\right)h_0(\hat{s})}.$$

It can be verified that $\varphi(\frac{1}{2}, q) = \frac{1}{2}$ independent of q and α. Thus it is socially optimal to replace the incumbent manager if and only if the posterior belief of his ability falls below the prior.

Given a contract (w, k) and belief about replacement cutoff $\hat{s}$, the manager chooses q to maximize:

$$\frac{1}{2}q[1 - H_1(\hat{s})]w + \left\{\frac{1}{2}qH_1(\hat{s}) + \left(1 - \frac{1}{2}q\right) H_0(\hat{s})\right\} k - \frac{1}{2}q^2.$$

The manager's best response is:

$$q = \max\left\{\frac{1}{2}[1 - H_1(\hat{s})]w - \frac{1}{2}[H_0(\hat{s}) - H_1(\hat{s})]k, 0\right\}.$$

Unlike the baseline model, here the incumbent manager is directly dis-incentivized by severance pay. An increase in severance pay increases the opportunity cost of exerting effort and leads directly to a decrease in effort. If the severance pay is high enough, the incumbent manager exerts no effort at all and is willing to be fired. Under this extension, severance pay is a double-edged sword. By the direct effect (better outside option if the incumbent manager is replaced), effort decreases. By the indirect effect (better job security with lower equilibrium replacement cutoff), effort increases. The design of the optimal contract should take this non-trivial incentive of k on q into consideration.

Fixed (w, k) and q, the board chooses $\hat{s}$ to maximize:

$$\frac{1}{2}q\big[1 - H_1(\hat{s})\big](1 - w) + \left\{\frac{1}{2}qH_1(\hat{s}) + \left(1 - \frac{1}{2}q\right) H_0(\hat{s})\right\} \left(\frac{1}{2}q - k\right).$$

The board's indifference condition is:

$$\frac{\frac{1}{2}qh_1(\hat{s})}{\frac{1}{2}qh_1(\hat{s}) + \left(1 - \frac{1}{2}q\right)h_0(\hat{s})}(1 - w) = \frac{1}{2}q - k.$$

$\hat{s}$ is the solution to $\zeta(\hat{s}, q) = \max\left\{\min\left\{\frac{\frac{1}{2}q - k}{1 - w}, 1\right\}, 0\right\}$, where $\zeta(\hat{s}, q)$ is the estimate of the outcome under the incumbent's management at $\hat{s}$ given q,

$$\zeta(\hat{s}, q) \equiv \frac{\frac{1}{2}qh_1(\hat{s})}{\frac{1}{2}qh_1(\hat{s}) + \left(1 - \frac{1}{2}q\right) h_0(\hat{s})}.$$

It is difficult to write the expected profit as a function of $\hat{s}$ alone because q is now affected by (w, k) directly and by equilibrium cutoff $\hat{s}$ indirectly. However, we can still discuss the optimal replacement policy under extreme information structure.

Multiple equilibria may exist for some contract (w, k) because incentive on effort q is not monotone in k as in the baseline model. For the same reason, equilibria may not be Pareto-ranked. We further assume that the equilibrium most favorable to the board is selected when multiple equilibria exist.

Proposition 7 (Optimal replacement policy) *There exist $\overline{\alpha}$ and $\underline{\alpha}$ such that,*

1. *$\hat{s}^*(\alpha) > \frac{1}{2}$ for $\alpha > \overline{\alpha}$;*
2. *$\hat{s}^*(\alpha) < \frac{1}{2}$ for $\alpha < \underline{\alpha}$.*

When the information structure is sufficiently noisy ($\alpha \to 0$), $H_0(s)$ is very close to $H_1(s)$ for $s \in [0, 1]$. The direct negative effect of severance pay on effort is small and the model is back to the baseline in the limit. Knowing that the board has noisy monitoring technology, the incumbent manager has little incentive to manipulate the realization of the signal. Entrenchment is expected to be optimal when α is sufficiently small.

When the information structure is sufficiently informative ($\alpha \to \infty$), the magnitude of the direct negative effect of k (i.e., $\frac{1}{2}[H_0(\hat{s}) - H_1(\hat{s})]$) is very large. Under this scenario, the board can simply avoid the disadvantage of k by setting $k = 0$. Moreover, this does not contradict the possibility of obtaining a net commitment gain. In fact, we can construct a contract with high wage and zero severance pay that yields anti-entrenchment and dominates all possible contracts that yield entrenchment. Thus, anti-entrenchment emerges in the optimal contract as the board's information structure becomes sufficiently informative.

6 Conclusion

This paper explores how the problem of motivating the incumbent manager to exert effort and keeping the flexibility to choose a high ability manager interacts with the equilibrium replacement policy. We focus on the situation where the board observes a non-contractible signal after the incumbent manager exerts effort and solve for the optimal contract. We show that the information technology that the board uses to assess the incumbent manager's ability is an important determinant of the optimal contract and of managerial turnover. Unlike the existing literature on managerial turnover, which aims to rationalize entrenchment, we show that anti-entrenchment can also be optimal for shareholders in some situations. This result is robust to allowing costly execution and the possibility that the board observes a signal of the outcome rather than incumbent manager's ability. The model highlights the board's monitoring technology as an important determinant of managerial turnover.

There are several interesting questions that can be pursued using the stylized model introduced in this paper. For future research, it would be interesting to endogenize the informativeness of the board's monitoring technology. In practice, informativeness is often the choice of the board. Some boards actively monitor their CEOs while some tend to be passive monitors. Endogenizing the board's monitoring technology could help us better understand the differences of monitoring intensity that occur across industries.

Another intriguing research avenue would be to incorporate voluntary departure into the model by allowing the possibility that manager possesses private information about the firm's profit. As Inderst and Mueller (2010) point out, managers sometimes have private information about a firm's performance. In such scenarios, the optimal contract needs to provide incentives for the incumbent manager to step down voluntarily. It would be interesting to build a unified model with both forced departure and voluntary departure and study the interaction between them.

Appendix A: Proofs of the propositions

Proof of Lemma 1. The first best outcome is the solution to the following maximization problem:

$$\max_{\{\hat{s},q\}} \frac{1}{2}\left[1 - F_1(\hat{s})\right]q + \frac{1}{4}\left[F_1(\hat{s}) + F_0(\hat{s})\right]q - \frac{1}{2}q^2.$$

The first order condition with respect to $\hat{s}$ yields:

$$f_1(\hat{s}) = f_0(\hat{s}) \Rightarrow \hat{s}^{\mathcal{FB}} = \frac{1}{2}.$$

The first order condition with respect to q yields:

$$q^{\mathcal{FB}} = \frac{1}{2} + \frac{1}{4}\big[F_0(\hat{s}^{\mathcal{FB}}) - F_1(\hat{s}^{\mathcal{FB}})\big].$$

■

Proof of Proposition 1. It is useful to first prove the two lemmas.

Lemma A1 (Uniform convergence of $F_1(\cdot)$ as $\alpha \to \infty$) *For any given $\epsilon > 0$, there exists N such that for $\alpha > N$, $F_1(s;\alpha) < \epsilon$ for $s \in [0, \frac{1}{2}]$ and $F_1(s;\alpha) < (2s-1) + \epsilon$ for $s \in [\frac{1}{2}, 1]$.*

Proof. By the definition of the completely informative information structure, given $\epsilon' = \frac{1}{2}\epsilon$ and $\Delta \in (0, \frac{1}{2})$, there exists N such that $f_1(\Delta;\alpha) < \epsilon'$ for $\alpha > N$. Thus,

$$F_1\Big(\frac{1}{2};\alpha\Big) = \int_0^{\frac{1}{2}} f_1(t;\alpha)dt = \int_0^{\Delta} f_1(t;\alpha)dt + \int_{\Delta}^{\frac{1}{2}} f_1(t;\alpha)dt \le \Delta\epsilon' + \Big(\frac{1}{2} - \Delta\Big).$$

Let $\Delta = \frac{1}{2} - \epsilon'$. $F_1\Big(\frac{1}{2};\alpha\Big)$ can be bounded above by

$$F_1(s;\alpha) \le F_1\Big(\frac{1}{2};\alpha\Big) \le \epsilon'\Big(\frac{1}{2} - \epsilon'\Big) + \epsilon' < 2\epsilon' = \epsilon \text{ for } s \in [0, \frac{1}{2}].$$

Similarly, for all $s \in [\frac{1}{2}, 1]$,

$$F_1(s;\alpha) = 2s - F_0(s;\alpha) = (2s-1) + F_1(1-s;\alpha) < (2s-1) + \epsilon \text{ for } \alpha > N.$$

■

Lemma A2 (Uniform convergence of $F_1(\cdot)$ as $\alpha \to 0$) *For any given $\epsilon > 0$, there exists N such that for $\alpha < N$, $F_1(s;\alpha) > s - \epsilon$ for all $s \in [0,1]$.*

Proof. By the definition of the completely uninformative information structure, for any given ϵ and $\Delta \in (0, \frac{1}{2})$, there exists N such that $f_1(\delta;\alpha) > 1 - \epsilon$. Thus,

$$\begin{aligned} s - F_1(s;\alpha) = \int_0^s \big[1 - f_1(t;\alpha)\big]dt &\le \int_0^{\frac{1}{2}} \big[1 - f_1(t;\alpha)\big]dt \\ &= \int_0^{\Delta} \big[1 - f_1(t;\alpha)\big]dt + \int_{\Delta}^{\frac{1}{2}} \big[1 - f_1(t;\alpha)\big]dt \\ &\le \Delta + \epsilon(\frac{1}{2} - \Delta) \text{ for } s \in [0, \frac{1}{2}]. \end{aligned}$$

Let $\Delta = \frac{1}{2}\epsilon$. $s - F_1(s;\alpha)$ can be bounded above by

$$s - F_1(s;\alpha) \leq \Delta + \epsilon(\frac{1}{2} - \Delta) < \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon \text{ for } s \in [0, \frac{1}{2}].$$

Similarly, for all $s \in [\frac{1}{2}, 1]$,

$$s - F_1(s;\alpha) = s - \big[2s - F_0(s;\alpha)\big] = (1-s) - F_1(1-s;\alpha) < \epsilon.$$

■

Recall that the expected profit function is

$$\pi(\hat{s}) = \frac{1}{8}\big[1 - F_1(\hat{s})\big]\left\{\frac{1}{2}\big[1 - F_1(\hat{s})\big] + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\right\}.$$

By Assumption 3, $f_0(s) = f_1(1-s)$ and $F_0(s) = 1 - F_1(1-s)$. The expected profit function can be written as

$$\pi(\hat{s}) = \frac{1}{16}\big[1 - F_1(\hat{s})\big]\Big[\big[1 - F_1(\hat{s})\big] + \hat{s}f_1(\hat{s})\Big].$$

1. Anti-entrenchment:

$$\pi(\hat{s};\alpha) < \frac{1}{16}(1+\hat{s}) < \frac{3}{32} \text{ for all } \hat{s} \in [0, \frac{1}{2}].$$

   By Lemma A1, for any ϵ, there exists N such that for $\alpha > N$, $1 - F_1(\hat{s};\alpha) > 2 - 2\hat{s} - \epsilon$ for $\hat{s} \in (\frac{1}{2}, 1)$. Moreover, given $\hat{s} \in (\frac{1}{2}, 1)$ and ϵ', there exists N' such that $\frac{1}{2}f_1(\hat{s}) = \frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} > 1 - \epsilon'$ for $\alpha > N'$.
   Let $\overline{\alpha} = \max\{N, N'\}$. Then for $\alpha > \overline{\alpha}$,

$$\pi(\hat{s};\alpha) > \frac{1}{16}\big(2 - 2\hat{s} - \epsilon\big)\left[(2 - 2\hat{s} - \epsilon) + 2\hat{s}(1-\epsilon')\right].$$

   Let $\hat{s} = \frac{1}{2}(1+\epsilon)$ and $\epsilon' = \frac{\epsilon}{1+\epsilon}$. Then,

$$\begin{aligned}\pi\Big(\frac{1}{2}(1+\epsilon);\alpha\Big) &> \frac{1}{16}(1-2\epsilon)\left[(1-2\epsilon) + (1+\epsilon)(1-\epsilon')\right] \\ &= \frac{1}{8}(1-2\epsilon)(1-\epsilon).\end{aligned}$$

To prove the proposition, it suffices to find ϵ such that

$$\frac{1}{8}(1-2\epsilon)(1-\epsilon) \geq \frac{3}{32}.$$

This inequality holds when $\epsilon \leq \frac{3-\sqrt{7}}{4}$.

2. Entrenchment:

   It suffices to prove that there exists $\underline{\alpha}$ such that for $\alpha < \underline{\alpha}$ and, $\pi(\hat{s}) < \pi(0) = \frac{1}{16}$ for all $\hat{s} \in [\frac{1}{2}, 1]$. Since $f_1(s) < 2$ for $s \in [0,1)$ by normalization, it directly follows that $1 - F_1(s) < 2(1-s)$. Thus, $1 - F_1(1-\Delta) < 2\Delta$ for $\Delta \in (0, \frac{1}{2})$.

$$\Rightarrow \pi(\hat{s}) < \frac{1}{4}\Delta(\Delta+1) \text{ for } \hat{s} \in [1-\Delta, 1].$$

   For Δ to be sufficiently small, $\frac{1}{4}\Delta(\Delta+1) < \frac{1}{16}$. In particular, let $\Delta = \frac{\sqrt{2}-1}{2}$. Then $\hat{s} \in [1-\Delta, 1]$ cannot be optimal replacement policy.

   It remains to prove that there extists $\underline{\alpha}$ such that for $\alpha < \underline{\alpha}$, $\pi(\hat{s}) < \pi(0)$ for all $\hat{s} \in [\frac{1}{2}, 1-\Delta]$. By the definition of the completely uninformative information structure, for any ϵ', there exists N' such that $\frac{1}{2}f_1(\hat{s}) = \frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} < \frac{1}{2} + \epsilon'$ for $\hat{s} \in [\frac{1}{2}, 1-\Delta]$ and $\alpha < N'$.

   By Lemma A2, for any ϵ, there exists N such that for $\alpha < N$, $F_1(\hat{s};\alpha) > \hat{s} - \epsilon$ for $\hat{s} \in [0,1]$. Thus,

$$\pi(\hat{s}) < \frac{1}{16}(1-\hat{s}+\epsilon)\big[(1-\hat{s}+\epsilon) + (1+2\epsilon')\big].$$

   Let $\epsilon' = \epsilon = \frac{\sqrt{3}}{3} - \frac{1}{2}$ and $\underline{\alpha} = \min\{N, N'\}$. Then for $\alpha < \underline{\alpha}$,

$$\pi(\hat{s}) < \frac{1}{16}(1-\hat{s}+\epsilon)(2-\hat{s}+3\epsilon) \leq \frac{3}{16}\big(\frac{1}{2}+\epsilon\big)^2 = \frac{1}{16} \text{ for all } \hat{s} \in [\frac{1}{2}, 1].$$

■

Proof of Lemma 3. For existence, it suffices to construct an example. Suppose $\{f_1(\cdot), f_0(\cdot)\}$ induces $g(\cdot)$. By the definition of the information structure,

$$\frac{1}{2}F_1(s) + \frac{1}{2}F_0(s) = s \text{ for all } s \in [0,1] \Leftrightarrow \frac{1}{2}f_1(s) + \frac{1}{2}f_0(s) = 1 \text{ for all } s \in [0,1].$$

Meanwhile, we have

$$g(p) = \Big[\frac{1}{2}f_1(\varphi^{-1}(p)) + \frac{1}{2}f_0(\varphi^{-1}(p)))\Big]\frac{d\varphi^{-1}(p)}{dp}.$$

Thus, $g(p)dp = d\varphi^{-1}(p) \Rightarrow \varphi(G(p)) = p \Rightarrow \tilde{f}_1(x) = 2G^{-1}(x)$ and $\tilde{f}_0(x) = 2[1 - G^{-1}(x)]$. This finishes the proof of existence.

For uniqueness, suppose two information structures $\{f_1(s), f_0(s)\}$ and $\{f_1^\dagger(s), f_1^\dagger(s)\}$ induce the same $g(p)$. By the definition of the information structure,

$$\frac{1}{2}f_1^\dagger(s) + \frac{1}{2}f_0^\dagger(s) = 1 = \frac{1}{2}f_1(s) + \frac{1}{2}f_0(s).$$

By the definition of p,

$$\frac{1}{2}f_1(s) + \frac{1}{2}f_0(s) = \frac{f_1(s)}{2p} = \frac{f_0(s)}{2(1-p)}.$$

By the derivation of $g(p)$,

$$g(p) = \left[\frac{1}{2}f_1(\varphi^{-1}(p)) + \frac{1}{2}f_0(\varphi^{-1}(p))\right]\frac{d\varphi^{-1}(p)}{dp}.$$

$$\Rightarrow pg(p) = \frac{1}{2}f_1(\varphi^{-1}(p))\frac{d\varphi^{-1}(p)}{dp}.$$

$$\Rightarrow \int_0^p tg(t)dt = \frac{1}{2}F_1(\varphi^{-1}(p)) = \frac{1}{2}F_1^\dagger(\varphi^{\dagger-1}(p)).$$

Similarly,

$$(1-p)g(p) = \frac{1}{2}f_0(\varphi^{-1}(p))\frac{d\varphi^{-1}(p)}{dp}.$$

$$\Rightarrow \int_0^p (1-t)g(t)dt = \frac{1}{2}F_0(\varphi^{-1}(p)) = \frac{1}{2}F_0^\dagger(\varphi^{\dagger-1}(p)).$$

Thus,

$$\frac{1}{2}F_1^\dagger(\varphi^{\dagger-1}(p)) + \frac{1}{2}F_0^\dagger(\varphi^{\dagger-1}(p)) = \frac{1}{2}F_1(\varphi^{-1}(p)) + \frac{1}{2}F_0(\varphi^{-1}(p)).$$

$$\Rightarrow \varphi^{\dagger-1}(p) = \varphi^{-1}(p) \Rightarrow f_1^\dagger(s) = f_1(s).$$

Since $\frac{1}{2}f_1^\dagger(s) + \frac{1}{2}f_0^\dagger(s) = \frac{1}{2}f_1(s) + \frac{1}{2}f_0(s)$, it follows directly that $f_0^\dagger(s) = f_0(s)$. This finishes the proof of uniqueness. ■

Proof of Lemma 4. By definition, $\underline{\rho} \leq 1 - \frac{G(p)g'(p)}{g^2(p)} \leq \overline{\rho}$. Integrating both sides from 0 to p yields

$$\underline{\rho}p \leq \frac{G(p)}{g(p)} - \frac{G(0)}{g(0)} \leq \overline{\rho}p.$$

$$\Rightarrow \frac{1}{\overline{\rho}}\frac{1}{p} \leq \frac{g(p)}{G(p)} \leq \frac{1}{\underline{\rho}}\frac{1}{p} \Leftrightarrow \frac{1}{\overline{\rho}} \leq \frac{pg(p)}{G(p)} \leq \frac{1}{\underline{\rho}}.$$

Integrating both sides from p to $\frac{1}{2}$ yields

$$\frac{1}{2}(2p)^{\frac{1}{\underline{\rho}}} \leq G(p) \leq \frac{1}{2}(2p)^{\frac{1}{\overline{\rho}}}.$$

■

Proof of Proposition 2.

Lemma A3 *If $G(p) \leq p$ for all $p \in [0, \frac{1}{2}]$, entrenchment is optimal to the board. Moreover, if $G(p)$ is convex in p for $p \in [0, \frac{1}{2}]$, $\hat{p}^* = 0$.*

Proof. We finish the proof in two steps:

1. $\tilde{\pi}(1-\hat{p}) < \tilde{\pi}(0)$ for $\hat{p} \in [0, \frac{1}{2}]$.

   It is equivalent to prove that

$$\int_{1-\hat{p}}^{1} tg(t)dt\left(1 - \int_{1-\hat{p}}^{1} G(t)dt\right) < \int_{0}^{1} tg(t)dt\left(1 - \int_{0}^{1} G(t)dt\right).$$

   Because $G(1-\hat{p}) = 1 - G(\hat{p})$, $\int_0^1 G(t)dt = \frac{1}{2}$. Thus, the right-hand side can be further simplified as

$$\int_{0}^{1} tg(t)dt\left(1 - \int_{0}^{1} G(t)dt\right) = \left(1 - \int_{0}^{1} G(t)dt\right)^2 = \frac{1}{4}.$$

For the left-hand side,

$$
\begin{aligned}
&\int_{1-\hat{p}}^{1} tg(t)dt\left(1-\int_{1-\hat{p}}^{1} G(t)dt\right)\\
&=\left(1-\int_{1-\hat{p}}^{1} G(t)dt-(1-\hat{p})G(1-\hat{p})\right)\left(1-\int_{1-\hat{p}}^{1} G(t)dt\right)\\
&<\left(1-\int_{0}^{\hat{p}}(1-G(t))dt-\frac{1}{2}(1-\hat{p})[1-G(\hat{p})]\right)^2\\
&=\left(\frac{1-\hat{p}}{2}(1+G(\hat{p}))+\int_0^{\hat{p}} G(t)dt\right)^2\\
&\leq\left(\frac{1-\hat{p}}{2}(1+\hat{p})+\int_0^{\hat{p}} tdt\right)^2=\frac{1}{4}.
\end{aligned}
$$

2. $\tilde{\pi}(\hat{p})$ is strictly decreasing in $\hat{p}$ for $\hat{p}\in[0,\frac{1}{2}]$ if $G(p)$ is convex in p for $p\in[0,\frac{1}{2}]$.

   First, notice that

$$
\int_{\hat{p}}^{1} tg(t)dt=\int_{\hat{p}}^{1} tdG(t)=1-\int_{\hat{p}}^{1} G(t)dt-\hat{p}G(\hat{p})<1-\int_{\hat{p}}^{1} G(t)dt \text{ for } \hat{p}\in(0,\frac{1}{2}].
$$

   Second, when $g(\hat{p})$ is increasing in $\hat{p}$, we have

$$
G(\hat{p})=\int_0^{\hat{p}} g(t)dt\leq\int_0^{\hat{p}} g(t)dt=\hat{p}g(\hat{p}).
$$

   Thus, $\tilde{\pi}'(p)<0$ for $p\in(0,\frac{1}{2}]$.

■

It directly follows that $\hat{p}^*=0$ for $\alpha\leq\alpha_1$ by Lemma A3. For $\alpha>\alpha_1$, it is useful to first prove the following two lemmas.

Lemma A4 *If $\rho(p;\alpha)$ is weakly decreasing in p, $\frac{G(p)}{pg(\hat{p})}$ is weakly decreasing in p for $p\in[0,\frac{1}{2}]$.*

Proof. By the definition of ρ-concavity,

$$
\rho(t)=1-\frac{G(t)g'(t)}{g^2(t)}.
$$

Integrating both sides from 0 to p yields,

$$
\int_0^p \rho(t)dt=\frac{G(p)}{g(p)}\Rightarrow\frac{G(p)}{pg(p)}=\frac{\int_0^p \rho(t)dt}{p}.
$$

$$\Rightarrow \left(\frac{\int_0^p \rho(t)dt}{p}\right)' = \frac{\rho(p)p - \int_0^p \rho(t)dt}{p^2} = \frac{\int_0^p \big[\rho(p) - \rho(t)\big]dt}{p^2} \leq 0.$$

■

Lemma A5 *For $\alpha_1 > \alpha_2$, $G(p;\alpha_1) > G(p;\alpha_2)$ for $p \in (0, \frac{1}{2})$.*

Proof. By Lemma A4,

$$\int_0^p \rho(t)dt = \frac{G(p)}{g(p)} \Rightarrow \ln\big(\frac{1}{2}\big) - \ln G(p;\alpha) = \int_p^{\frac{1}{2}} \frac{1}{\int_0^\omega \rho(t;\alpha)dt} d\omega.$$

It can be verified that $\int_p^{\frac{1}{2}} \frac{1}{\int_0^\omega \rho(t;\alpha)dt} d\omega$ is decreasing in α by the definition of ρ-concave order. Thus, $G(p;\alpha)$ is increasing in α. ■

Rearranging the first order condition with respect to $\hat{p}$ yields

$$\tilde{\pi}'(\hat{p}) \gtreqless 0 \Leftrightarrow \frac{G(\hat{p})}{\hat{p}g(\hat{p})} \gtreqless \frac{1 - \int_{\hat{p}}^1 G(t)dt}{\int_{\hat{p}}^1 tg(t)dt} = \frac{\frac{1}{2} + \int_0^{\hat{p}} G(t)dt}{\frac{1}{2} + \int_0^{\hat{p}} G(t)dt - pG(p)}.$$

By Lemma A4, the left-hand side is decreasing in $\hat{p}$. It is can be verified that the right-hand side is increasing in $\hat{p}$. Thus, the board's profit function for $\hat{p} \in [0, \frac{1}{2}]$ is well-behaved.

Notice that $\lim_{p\to 0} \frac{G(\hat{p})}{\hat{p}g(\hat{p})} = \rho(0) > 1$ for $\alpha > \alpha_1$, and $\lim_{p\to 0} \frac{1-\int_{\hat{p}}^1 G(t)dt}{\int_{\hat{p}}^1 tg(t)dt} = 1$. It suffices to compare the end points of the two curves.

If $2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt > \frac{\frac{1}{2}+\int_0^{\frac{1}{2}} G(t)dt}{\frac{1}{4}+\int_0^{\frac{1}{2}} G(t)dt}$, $\tilde{\pi}(\hat{p})$ is increasing in $\hat{p} \in [0, \frac{1}{2}]$ and the optimal cutoff $\hat{p}^*$ lies between $\frac{1}{2}$ and 1.

If $2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt < \frac{\frac{1}{2}+\int_0^{\frac{1}{2}} G(t)dt}{\frac{1}{4}+\int_0^{\frac{1}{2}} G(t)dt}$, $\tilde{\pi}(\hat{p})$ is first increasing and then decreasing in $\hat{p} \in [0, \frac{1}{2}]$. The maximal can be pinned down by the first order condition for $\hat{p} \in [0, \frac{1}{2}]$. We further argue that this local maximal is indeed the global maximal for $\hat{p} \in [0, 1]$. To see this, notice that second order derivative of the profit function with respect to $\hat{p}$ is

$$\tilde{\pi}''(\hat{p}) = \frac{1}{4}\Big[- \hat{p}g'(\hat{p})\Big(1 - \int_{\hat{p}}^1 G(t)dt\Big) - 3\hat{p}g(\hat{p})G(\hat{p})\Big].$$

Because $G(p)$ is concave for $p \in [0, \frac{1}{2}]$ for $\alpha > \alpha_1$, $G(p)$ is convex for $p \in [\frac{1}{2}, 1]$. This directly implies $g'(p) > 0$ for $p \in [\frac{1}{2}, 1]$. Thus $\pi''(\hat{p}) < 0$ for $p \in [\frac{1}{2}, 1]$. Because

$\tilde{\pi}'(\frac{1}{2}) < 0$, profit is decreasing in $\hat{p}$ for $\hat{p} \in [\frac{1}{2}, 1]$.

By the definition of ρ-concavity, $2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt$ is increasing in α. By Lemma A5, $\int_0^{\frac{1}{2}} G(t;\alpha)dt$ is increasing in $\alpha \Rightarrow \frac{\frac{1}{2}+\int_0^{\frac{1}{2}} G(t)dt}{\frac{1}{4}+\int_0^{\frac{1}{2}} G(t)dt}$ is decreasing in α. By Assumptions 5(d) and 5(e),

$$\lim_{\alpha\to\alpha_1} 2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt = 1 \text{ and } \lim_{\alpha\to\infty} 2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt = \infty.$$

Thus, there exists $\alpha_2 > \alpha_1$ such that for $\alpha > \alpha_2$, anti-entrenchment is optimal; for $\alpha < \alpha_2$, entrenchment is optimal. ■

Proof of Example 1. Given the functional form of $g(\cdot)$, it can be verified that the board's profit function is

$$\tilde{\pi}(\hat{p};\alpha) = \begin{cases} \frac{1}{4}\left[\frac{1}{2} + \frac{1}{4}\frac{\alpha}{\alpha+1}(2\hat{p})^{\frac{\alpha+1}{\alpha}}\right]\left[\frac{1}{2} - \frac{1}{4}\frac{1}{1+\alpha}(2\hat{p})^{\frac{\alpha+1}{\alpha}}\right] & \text{for} \quad \hat{p} \in [0, \frac{1}{2}] \\ \frac{1}{16}[2(1-\hat{p})]^{\frac{\alpha+1}{\alpha}}\left[\frac{\alpha}{\alpha+1} + \frac{\hat{p}}{1-\hat{p}}\right]\left[\hat{p} + \frac{1}{4}\frac{\alpha}{\alpha+1}[2(1-\hat{p})]^{\frac{\alpha+1}{\alpha}}\right] & \text{for} \quad \hat{p} \in (\frac{1}{2}, 1] \end{cases}.$$

It is obvious that $\alpha_1 = 1$. By the proof of Proposition 2, α_2 is the informativeness such that the first derivative of profit at $\hat{p} = \frac{1}{2}$ is equal to 0. $\Rightarrow \alpha_2 = \frac{\sqrt{5}+1}{2}$. ■

Proof of Proposition 3. Designing a contract based on the signal is equivalent to designing a contract based on the posterior belief about the incumbent's ability $p \in [0, 1]$. By abuse of notation, denote $\{w(p), r(p), k(p)\}$ as the contract based on the board's posterior belief. It suffices to prove that $r^*(p) = 1$ for $p \in [\frac{1}{2}, 1]$ and $r^*(p) = 0$ for $p \in [0, \frac{1}{2})$ in the optimal contract.

Given contract $\{w(p), r(p), k(p)\}$, the incumbent manager chooses q to maximize:

$$\int_0^1 \{r(p)qpw(p) + [1 - r(p)]k(p)\}g(p)dp - C(q).$$

The first order condition with respect to q yields

$$C'(q) = \int_0^1 r(p)qw(p)g(p)dp.$$

Note that $k(p)$ cannot provide incentive on the effort level. Because the incumbent manager is protected by limited liability, $k^*(p) = 0$ in the optimal contract.

The board chooses $\{w(p), r(p)\}$ to maximize

$$\begin{aligned}&\int_0^1 \Big\{\big[r(p)qp(1-w(p))\big] + \frac{1}{2}q(1-r(p))\Big\}g(p)dp\\ &= q\bigg(\int_0^1 \Big[r(p)p + \frac{1}{2}(1-r(p))\Big]g(p)dp - \int_0^1 r(p)pw(p)g(p)dp\bigg).\end{aligned}$$

Given effort level q, $\int_0^1 r(p)pw(p)g(p)dp = C'(q)$ is a constant by the incumbent manager's first order condition. It is equivalent to maximize:

$$\int_0^1 \Big[r(p)p + \frac{1}{2}(1-r(p))\Big]g(p)dp.$$

The integral is maximized by setting $r(p) = 1$ for $p \in [\frac{1}{2}, 1]$ and $r(p) = 0$ for $p \in [0, \frac{1}{2})$.

■

Proof of Proposition 4. We first proved that $k^* = 0$ in the optimal contract. Given (w_1, w_2, k) and belief of cutoff $\hat{s}$, the incumbent manager chooses q to maximize

$$\frac{1}{2}[1-F_1(\hat{s})]qw_1 + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\Big(\frac{1}{2}qw_2 + k\Big) - C(q).$$

$$\Rightarrow q = \frac{1}{2}[1-F_1(\hat{s})]w_1 + \frac{1}{4}\big[F_1(\hat{s}) + F_0(\hat{s})\big]w_2.$$

Given (w_1, w_2, k) and belief of project quality q, the board's indifference condition yields:

$$\frac{1}{2}q(1-w_2) - k = \frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}q(1-w_1).$$

The board chooses (w_1, w_2, k) to maximize expected profit:

$$\begin{aligned}&\frac{1}{2}\big[1-F_1(\hat{s})\big]q(1-w_1) + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\Big[\frac{1}{2}q(1-w_2) - k\Big]\\ &= q(1-w_1)\left\{\frac{1}{2}\big[1-F_1(\hat{s})\big] + \frac{1}{2}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\right\}.\end{aligned}$$

Given $(q, \hat{s})$ the board would like to induce, it is obvious that profit is decreasing in w_1. By the two equilibrium conditions, it is easy to verify that $w_1(k)$ is increasing in k and $w_2(k)$ is decreasing in k. Thus, $k^* = 0$.

The board's profit maximization problem can be written as

$$\max_{\{w_1,w_2,q,\hat{s}\}} \pi(w_1,w_2,q,\hat{s}) := \frac{1}{2}\big[1-F_1(\hat{s})\big]q(1-w_1)+\frac{1}{4}\big[F_1(\hat{s})+F_0(\hat{s})\big]q(1-w_2)$$

s.t.

$$q-\Big(\frac{1}{2}[1-F_1(\hat{s})]w_1+\frac{1}{4}\big[F_1(\hat{s})+F_0(\hat{s})\big]w_2\Big)=0$$

and

$$\frac{1}{2}(1-w_2)-\varphi(\hat{s})(1-w_1)=0.$$

Let $\mathcal{L}$ be the Lagrangian and denote λ_1 and λ_2 as Lagrangian multipliers on the two constraints respectively.

$$\frac{\partial\mathcal{L}(w_1,w_2,q,\hat{s},\lambda_1,\lambda_2)}{\partial w_1}=0 \Rightarrow -\frac{1}{2}\big(q+\lambda_1\big)\big[1-F_1(\hat{s})\big]+\varphi(\hat{s})\lambda_2=0.$$

$$\frac{\partial\mathcal{L}(w_1,w_2,q,\hat{s},\lambda_1,\lambda_2)}{\partial w_2}=0 \Rightarrow -\frac{1}{2}(q+\lambda_1)\big[F_1(\hat{s})+F_0(\hat{s})\big]-\lambda_2=0.$$

It can be verified that $\hat{s}=0$ is never optimal. Thus, $\varphi(\hat{x})>0$. Then $\lambda_2=0$ and $\lambda_1=-q$ must be true. The first order condition of the Lagrangian with respect to $\hat{s}$ yields

$$\frac{\partial\mathcal{L}(w_1,w_2,q,\hat{s},\lambda_1,\lambda_2)}{\partial\hat{s}}=0.$$

$$\begin{aligned}
&\Rightarrow -q(1-w_1)f_1(\hat{s})+\frac{1}{2}\big[f_1(\hat{s})+f_0(\hat{s})\big]q(1-w_2)\\
&\qquad\qquad +\lambda_1\left(f_1(\hat{s})w_1-\frac{1}{2}\big[f_1(\hat{s})+f_0(\hat{s})\big]w_2\right)=0\\
&\Rightarrow \varphi(\hat{s})=\frac{1}{2}\Rightarrow \hat{s}^*=\frac{1}{2}.
\end{aligned}$$

Because $\frac{1}{2}(1-w_2)-\varphi(\hat{s})(1-w_1)=0$, it follows directly that $w_1^*=w_2^*$. ■

Proof of Proposition 5. Given contract (w,k) and $\hat{s}$, the incumbent manager's best response is

$$q=\left\{(1+\tau)\Big[\frac{1}{2}\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big]+\frac{1}{2}\big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]w\right\}^{\frac{1}{1-\tau}}.$$

Similarly, the board's indifference condition is

$$\frac{1}{2}q^{1+\tau} - k = \left[\frac{1}{2} + \frac{f_1(\hat{s}) - f_0(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\delta\right] q^{1+\tau}(1-w).$$

Plugging the two equilibrium conditions into the board's profit function yields

$$\pi(\hat{s}) = M\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]^{\frac{1+\tau}{1-\tau}}$$
$$\cdot\Big\{\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big] + \big[F_1(\hat{s})+F_0(\hat{s})\big]\Big[\frac{1}{2} + \frac{f_1(\hat{s}) - f_0(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\delta\Big]\Big\},$$

where $M = \frac{1}{4}^{\frac{2}{1-\tau}}(1-\tau)(1+\tau)^{\frac{2+2\tau}{1-\tau}}$.

Next, we calculate the board's expected profit for a given cutoff $\hat{s}$ as $\alpha \to \infty$:

$$\lim_{\alpha\to\infty}\pi(\hat{s};\alpha) = \begin{cases} M\Big[1-(1-2\delta)\hat{s}\Big]^{\frac{1+\tau}{1-\tau}} & \text{for} \quad \hat{s}\in[0,\frac{1}{2}) \\ M\Big[\delta+\frac{1}{2}\Big]^{\frac{1+\tau}{1-\tau}}(1+\delta) & \text{for} \quad \hat{s}=\frac{1}{2} \\ M\Big[(1+2\delta)(1-\hat{s})\Big]^{\frac{1+\tau}{1-\tau}}(1+2\delta) & \text{for} \quad \hat{s}\in(\frac{1}{2},1] \end{cases}.$$

1. Entrenchment as $\alpha \to \infty$:

   Notice that $\frac{f_1(\hat{s})-f_0(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} \leq 1$. Then $\pi(\hat{s};\alpha)$ can be bounded above by

   $$\pi(\hat{s};\alpha) \leq M\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]^{\frac{1+\tau}{1-\tau}}$$
   $$\cdot\Big\{\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big] + \big[F_1(\hat{s})+F_0(\hat{s})\big]\Big[\frac{1}{2}+\delta\Big]\Big\}.$$

   Denote the right-hand side as $\pi_E(\hat{s};\alpha)$. By Lemma A1, $F_1(\hat{s};\alpha)$ converges uniformly to $\max\{0, 2\hat{s}-1\}$ as $\alpha \to \infty$. Thus, $\pi_E(\hat{s};\alpha)$ converges uniformly to $M\Big[(1+2\delta)(1-\hat{s})\Big]^{\frac{1+\tau}{1-\tau}}(1+2\delta)$ for $\hat{s}\in[\frac{1}{2},1]$ as $\alpha\to\infty$. Because $\pi(0;\alpha) = M$, entrenchment is optimal for sufficiently large α if

   $$M > \max_{\hat{s}\in[\frac{1}{2},1]}\Big\{M\big[(1+2\delta)(1-\hat{s})\big]^{\frac{1+\tau}{1-\tau}}\big(1+2\delta\big)\Big\} \Rightarrow \delta < \frac{1}{2}^{\frac{1-\tau}{2}} - \frac{1}{2}.$$

2. Anti-entrenchment as $\alpha \to \infty$:

Notice that $\frac{f_1(\hat{s})-f_0(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} \leq 0$ for $\hat{s} \in [0, \frac{1}{2}]$. Then $\pi(\hat{s};\alpha)$ can be bounded above by

$$\begin{aligned}\pi(\hat{s};\alpha) \leq & M\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big]+\big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]^{\frac{1+\tau}{1-\tau}} \\ & \cdot\Big\{\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big]+\big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]+\frac{1}{2}\big[F_1(\hat{s})+F_0(\hat{s})\big]\Big\}.\end{aligned}$$

Denote the right-hand side as $\pi_A(\hat{s};\alpha)$. By Lemma A1, $F_1(\hat{s};\alpha)$ converges uniformly to $\max\big\{0, 2\hat{s}-1\big\}$. Thus, $\pi_A(\hat{s};\alpha)$ converges uniformly to $\xi(\hat{s};\delta,\tau) = M\Big[1-(1-2\delta)\hat{s}\Big]^{\frac{1+\tau}{1-\tau}}\big(1+2\delta\hat{s}\big)$ for $\hat{s} \in [0, \frac{1}{2}]$ as $\alpha \to \infty$.

Given $(\delta,\tau) \in (0, \frac{1}{2}) \times (-1, 1)$, it can be verified that there exists $\nu(\delta,\tau) < \frac{1}{2}$ such that $\xi(\hat{s};\delta,\tau) < M\big[\delta+\frac{1}{2}\big]^{\frac{1+\tau}{1-\tau}}(1+2\delta)$ for $\hat{s} \in [\nu(\delta,\tau), \frac{1}{2}]$. Thus, $\hat{s} \in [\nu(\delta,\tau), \frac{1}{2}]$ can never be optimal for sufficiently large α.

Because $f_1(\hat{s})$ is strictly increasing in $\hat{s}$ and $\lim_{\alpha\to\infty} f_1(\hat{s};\alpha) = 0$ for all $\hat{s} \in [0, \frac{1}{2})$, $f_1(\hat{s};\alpha)$ converges uniformly to 0 for $\hat{s} \in [0, \nu(\delta,\tau)]$ as $\alpha \to \infty$. Then $\pi(\hat{s};\alpha)$ converges uniformly to $M\Big[1-(1-2\delta)\hat{s}\Big]^{\frac{1+\tau}{1-\tau}}$ for $\hat{s} \in [0, \nu(\delta,\tau)]$ as $\alpha \to \infty$. Thus, entrenchment is optimal for sufficiently large α if

$$\max_{\hat{s}\in[0,\nu(\delta,\tau)]} M\Big[1-(1-2\delta)\hat{s}\Big]^{\frac{1+\tau}{1-\tau}} < M\Big[\delta+\frac{1}{2}\Big]^{\frac{1+\tau}{1-\tau}}(1+2\delta) \Rightarrow \delta > \frac{1}{2}^{\frac{1-\tau}{2}} - \frac{1}{2}.$$

■

Proof of Proposition 6. Given $\hat{s}$, a contract can always be constructed to induce $\hat{s}$. However, it is not necessarily $w = \frac{1}{2}$. $k \geq 0$ does not hold for all $\hat{s} \in [0,1]$ when $w = \frac{1}{2}$. To see this, notice that the severance pay k is

$$k(\hat{s}, w) = \underline{\pi}\big(q(\hat{s}, w)\big) - \frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})}\big[(1-\lambda)q(\hat{s}, w)+\lambda e(\hat{s}, w)\big](1-w).$$

Given $\hat{s}$, letting w be sufficiently close to 1 generates a positive severance pay k. Define $\hat{\mathcal{S}} = \Big\{\hat{s}\,\big|k(\hat{s}, \frac{1}{2}) \geq 0 \,\&\, \hat{s} \in [0,1]\Big\}$, which is the the set of cutoffs that can be induced by contracts that satisfy $w = \frac{1}{2}$ and $k \geq 0$. If $\hat{s} \in \hat{\mathcal{S}}$, the board's expected profit can be written as

$$\overline{\pi}(\hat{s}) = \frac{1}{16}\Big[1-F_1(\hat{s})\Big]\Big[(1-\lambda)^2+\lambda^2\frac{1}{1-\hat{s}}\Big]\Big[[1-F_1(\hat{s})]+\hat{s}f_1(\hat{s})\Big].$$

If $\hat{s} \notin \hat{\mathcal{S}}$, $w = \frac{1}{2}$ cannot be sustained. Define $\mathcal{W}(\hat{s}) = \Big\{ w \,\Big|\, k(\hat{s}, w) \geq 0 \,\&\, w \in [0,1] \Big\}$, which is the set of wages that can induce $\hat{s}$ without violating the limited liability constraint of k.

1. Entrenchment:

   It suffices to prove that $\pi(\hat{s}; \alpha) < \pi(0; \alpha)$ for all $\hat{s} \in [\frac{1}{2}, 1]$ for sufficiently small α. $\pi(0; \alpha)$ is independent of α and can be calculated as

$$\pi(0; \alpha) = \frac{1}{16}\big[(1-\lambda)^2 + \lambda^2\big].$$

Lemma A6 *There exist $\Delta \in (0, \frac{1}{2})$ and N such that for $\alpha < N$, $\pi(\hat{s}; \alpha) < \pi(0; \alpha)$ for all $\hat{s} \in [1-\Delta, 1]$.*

Proof. By Lemma A2, for any $\epsilon > 0$ there exists N such that for $\alpha < N$, $1 - F_1(1-\Delta) < \Delta + \epsilon$ for all $\Delta \in [0,1]$. Note that $\frac{1}{4}\Big(\frac{1}{2}\lambda + \frac{1-\lambda}{\lambda}q\Big)^2 \geq \frac{1}{2}(1-\lambda)q$ for all q.

The expected profit of replacement can be bounded above by

$$\begin{aligned}
\underline{\pi}(q) - k &\leq \frac{1}{4}\Big(\frac{1}{2}\lambda + \frac{1-\lambda}{\lambda}q\Big)^2 = \frac{1}{16}\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]w\Big)^2 \\
&\leq \frac{1}{16}\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]\Big)^2.
\end{aligned}$$

Thus, the board's expected profit can be bounded for $\hat{s} \in [1-\Delta, 1]$:

$$\begin{aligned}
\pi(\hat{s}) \leq{}& \frac{1}{2}\big[1 - F_1(\hat{s})\big]\left\{\frac{1 - F_1(\hat{s})}{2}(1-\lambda)^2 + \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}\lambda^2\right\} w(1-w) \\
&+ \frac{1}{32}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]\Big)^2 \\
\leq{}& \frac{1}{8}\big[1 - F_1(\hat{s})\big]\left\{\frac{1 - F_1(\hat{s})}{2}(1-\lambda)^2 + \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}\lambda^2\right\} \\
&+ \frac{1}{16}\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]\Big)^2 \\
<{}& \frac{1}{8}(\Delta + \epsilon)\Big[\frac{\Delta + \epsilon}{2}(1-\lambda)^2 + \lambda^2\Big] + \frac{1}{16}\Big[\lambda + \frac{(1-\lambda)^2}{\lambda}(\Delta + \epsilon)\Big]^2.
\end{aligned}$$

Note that the last expression is increasing in $\Delta + \epsilon$. It suffices to prove that

$$\frac{1}{16}\lambda^2 < \frac{1}{16}\big[(1-\lambda)^2 + \lambda^2\big].$$

Consequently, we can always find sufficiently small Δ and ϵ such that

$$\frac{1}{8}(\Delta+\epsilon)\Big[\frac{\Delta+\epsilon}{2}(1-\lambda)^2+\lambda^2\Big]+\frac{1}{16}\Big[\lambda+\frac{(1-\lambda)^2}{\lambda}(\Delta+\epsilon)\Big]^2<\frac{1}{16}\big[(1-\lambda)^2+\lambda^2\big].$$

■

Next, notice that $\overline{\pi}(\hat{s})$ is the maximum expected profit without the limited liability constraint of k. Thus $\pi(\hat{s})\leq\overline{\pi}(\hat{s})$ for all $\hat{s}\in[0,1]$.

Given ϵ, the expected profit for $\alpha<N$ for all $\hat{s}\in[\frac{1}{2},1-\Delta]$ is bounded above by

$$\begin{aligned}\pi(\hat{s}) &\leq \frac{1}{16}\Big[1-F_1(\hat{s})\Big]\Big[(1-\lambda)^2+\lambda^2\frac{1}{1-\hat{s}}\Big]\Big[[1-F_1(\hat{s})]+\hat{x}f_1(\hat{s})\Big]\\ &\leq \frac{1}{16}(1-\hat{s}+\epsilon)\Big[(1-\lambda)^2+\lambda^2\frac{1}{1-\hat{s}}\Big]\big[(1-\hat{s}+\epsilon)+\hat{s}(1+\epsilon)\big]\\ &\leq \frac{1}{16}\Big(1+\frac{\epsilon}{\delta}\Big)\Big[\frac{1}{2}(1-\lambda)^2+\lambda^2\Big](1+2\epsilon).\end{aligned}$$

It remains to prove that $\frac{1}{16}\Big[\frac{1}{2}(1-\lambda)^2+\lambda^2\Big]<\frac{1}{16}\big[(1-\lambda)^2+\lambda^2\big]$, which is obvious.

2. Anti-entrenchment:

   For $\hat{s}\in[0,\frac{1}{2}]$, the expected profit can be bounded above by

$$\begin{aligned}\pi(\hat{s};\alpha) &\leq \frac{1}{16}\Big[1-F_1(\hat{s};\alpha)\Big]\Big[(1-\lambda)^2+\lambda^2\frac{1}{1-\hat{s}}\Big]\Big[[1-F_1(\hat{s};\alpha)]+\hat{x}f_1(\hat{s};\alpha)\Big]\\ &< \frac{3}{32}\Big[(1-\lambda)^2+2\lambda^2\Big].\end{aligned}$$

   It remains to find $\hat{s}>\frac{1}{2}$ that yields a profit no less than $\frac{3}{32}[(1-\lambda)^2+2\lambda^2]$. For notational convenience, let $\psi=\big(\frac{\lambda}{1-\lambda}\big)^2$. $\lambda<\sqrt{2}-1$ directly implies that $\psi<\frac{1}{2}$. By the limited liability constraint of k, we have

$$\frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})}\big[(1-\lambda)q+\lambda e\big](1-w)\leq\underline{\pi}(q).$$

   Notice that $\frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})}<1$, it suffices to satisfy

$$\frac{1-F_1(\hat{s})}{4}(1-\lambda)^2w\geq\left\{\frac{1-F_1(\hat{s})}{2}(1-\lambda)^2+\frac{1-F_1(\hat{s})}{[1-F_1(\hat{s})]+[1-F_0(\hat{s})]}\lambda^2\right\}w(1-w),$$

   and

$$\frac{1-F_1(\hat{s};\alpha)}{2}w\geq\frac{1}{2}\psi.$$

The second inequality comes from the construction that the board will not induce effort from the replacement manager. Let $\hat{s} = \frac{1}{2} + \kappa(\lambda)$ and $w = \frac{1}{2} + \iota(\lambda)$. Then it suffices to find (κ, ι) that yields a higher expected profit given λ. Note that the first inequality is independent of α. By Lemma A1, $\frac{1-F_1(\hat{s};\alpha)}{2}$ can be arbitrarily close to $1 - \hat{s}$ when α is sufficiently large. Thus, these two conditions can be further simplified as

$$\frac{1}{2}\Big(\frac{1}{2} - \kappa\Big) \geq \Big[\Big(\frac{1}{2} - \kappa\Big) + \psi\Big]\Big(\frac{1}{2} - \iota\Big),$$

and

$$2\Big(\frac{1}{2} + \iota\Big)\Big(\frac{1}{2} - \kappa\Big) \geq \psi.$$

$$\Rightarrow \iota \geq \max\Big\{\frac{\frac{1}{2}\psi}{\frac{1}{2} - \kappa + \psi}, \frac{\psi}{1 - 2\kappa} - \frac{1}{2}\Big\}.$$

Let $\iota = \frac{\frac{1}{2}\psi}{\frac{1}{2}-\kappa+\psi}$. It can be verified that $\psi < \frac{1}{4}$ if $\kappa < \frac{1}{2} - \psi$. The board's expected profit from the contract with the incumbent manager (w, k) that induces $\hat{s} = \frac{1}{2}+\kappa$ with wage $w = \frac{1}{2} + \iota$ as $\alpha \to \infty$ is

$$\lim_{\alpha\to\infty} \pi(\hat{s}; \alpha) = \Big[(1 - \lambda)^2\big(\frac{1}{2} - \kappa\big) + \lambda^2\Big]\Big(\frac{1}{4} - \frac{\frac{1}{2}\psi}{\frac{1}{2} - \kappa + \psi}^{2}\Big).$$

Note that

$$\begin{aligned}
\lim_{\kappa\to 0}\lim_{\alpha\to\infty} \pi(\hat{s}; \alpha) &= \lim_{\kappa\to 0}\Big\{\Big[(1 - \lambda)^2\big(\frac{1}{2} - \kappa\big) + \lambda^2\Big]\Big[\frac{1}{4} - \Big(\frac{\frac{1}{2}\psi}{\frac{1}{2} - \kappa + \psi}\Big)^2\Big]\Big\} \\
&= \frac{1}{2}\big[(1 - \lambda)^2 + 2\lambda^2\big]\Big[\frac{1}{4} - \Big(\frac{\frac{1}{2}\psi}{\frac{1}{2} + \psi}\Big)^2\Big] \\
&> \frac{3}{32}\big[(1 - \lambda)^2 + 2\lambda^2\big].
\end{aligned}$$

Thus, we can find sufficiently small κ such that $\lim_{\alpha\to\infty} \pi(\hat{s}; \alpha) > \frac{3}{32}\big[(1 - \lambda)^2 + 2\lambda^2\big]$. That is, anti-entrenchment is optimal to the board when α is sufficiently large and $\lambda < \sqrt{2} - 1$.

■

Proof of Proposition 7.

1. Entrenchment:

   It can be verified that $\pi(0;\alpha) = \frac{1}{16}$. Similarly, $\pi(1;\alpha) = 0$. Thus, $\hat{s} = 1$ is never optimal. It suffices to prove that there exists N such that for $\alpha < N$, $\pi(\hat{s}) < \pi(0)$ for all $\hat{s} \in [\frac{1}{2}, 1]$.

   Lemma A7 *There exist $\Delta \in (0, \frac{1}{2})$ and N such that for $\alpha < N$, $\pi(\hat{s};\alpha) < \pi(0;\alpha)$ for all $\hat{s} \in [1-\Delta, 1]$.*

   Proof. Since $q = \max\left\{\frac{1}{2}[1 - H_1(\hat{s})]w - \frac{1}{2}[H_0(\hat{s}) - H_1(\hat{s})]k, 0\right\}$, the effort level of the incumbent manager can be bounded above by

$$q \leq \frac{1}{2}[1 - H_1(\hat{s})]w.$$

   Thus, the expected profit can be bounded above by

$$\begin{aligned}
\pi(\hat{s}, q) &\leq \frac{1}{2}q[1 - H_1(\hat{s})](1 - w) + \left\{\frac{1}{2}qH_1(\hat{s}) + \left(1 - \frac{1}{2}q\right) H_0(\hat{s})\right\} \left(\frac{1}{2}q - k\right) \\
&\leq \frac{1}{2}q[1 - H_1(\hat{s})](1 - w) + \frac{1}{2}q \left\{\frac{1}{2}qH_1(\hat{s}) + \left(1 - \frac{1}{2}q\right) H_0(\hat{s})\right\} \\
&\leq \frac{1}{2}q\big[[1 - H_1(\hat{s})] + 1\big] \leq q < 1 - H_1(\hat{s}).
\end{aligned}$$

   Let $\Delta = \frac{1}{32}$. By Lemma A2, for $\epsilon' = \frac{1}{32}$, there exists N such that for $\alpha < N$, $H_1(\hat{s}) \geq \hat{s} - \epsilon'$ for all $\hat{s} \in [0, 1]$. Since $\hat{s} \geq 1 - \Delta$, we have

$$\pi(\hat{s}, q) < 1 - H_1(\hat{s}) \leq 1 - \hat{s} + \epsilon' \leq \Delta + \epsilon' = \frac{1}{16} = \pi(0;\alpha).$$

   ■

   Lemma A8 *Given any $\Delta \in (0, \frac{1}{2})$ and $q \in [0, 1]$, for any $\epsilon > 0$, there exists N' such that for $\alpha < N'$, $\frac{\frac{1}{2}qh_1(\hat{s})}{\frac{1}{2}qh_1(\hat{s}) + \left(1 - \frac{1}{2}q\right)h_0(\hat{s})} \leq \frac{1}{2}q + \epsilon$ for $\hat{s} \in [\frac{1}{2}, 1 - \Delta]$.*

   Proof. For any $\epsilon > 0$, let $\epsilon' = \frac{\epsilon}{1+\epsilon}$. By the definition of the completely uninformative information structure, there exists N' such that for $\alpha < N'$,

$h_1(1-\delta;\alpha) < 1+\epsilon'$.

$$\begin{aligned}\frac{\frac{1}{2}qh_1(\hat{s})}{\frac{1}{2}qh_1(\hat{s})+\left(1-\frac{1}{2}q\right)h_0(\hat{s})}-\frac{1}{2}q &= \frac{1}{2}q\Big(1-\frac{1}{2}q\Big)\frac{h_1(\hat{s})-h_0(\hat{s})}{\frac{1}{2}qh_1(\hat{s})+\left(1-\frac{1}{2}q\right)h_0(\hat{s})} \\ &\leq \frac{1}{2}\frac{h_1(\hat{s})-h_0(\hat{s})}{h_0(\hat{s})} = \frac{h_1(\hat{s})-1}{2-h_1(\hat{s})} \\ &\leq \frac{h_1(1-\delta;\alpha)-1}{2-h_1(1-\delta;\alpha)} \leq \frac{\epsilon'}{1-\epsilon'} = \epsilon.\end{aligned}$$

■

By Lemma A8, for all $\hat{s}\in[\frac{1}{2},1-\Delta]$, $\pi(\hat{s},q)$ can be bounded above by

$$\begin{aligned}\pi(\hat{s},q) &\leq \frac{1}{2}q[1-H_1(\hat{s})](1-w)+\left\{\frac{1}{2}qH_1(\hat{s})+\left(1-\frac{1}{2}q\right)H_0(\hat{s})\right\}\left(\frac{1}{2}q+\epsilon\right)(1-w) \\ &\leq \frac{1}{2}q(1-w)\Big[\big[1-H_1(\hat{s})\big]+\frac{1}{2}qH_1(\hat{s})+\Big(1-\frac{1}{2}q\Big)H_0(\hat{s})\Big]+\epsilon \\ &\leq \frac{1}{4}[1-H_1(\hat{s})][2-H_1(\hat{s})]w(1-w)+\epsilon \\ &\leq \frac{1}{16}(1-\hat{s}+\epsilon)(2-\hat{s}+\epsilon)+\epsilon = \frac{1}{16}\Big(\frac{1}{2}+\epsilon\Big)\Big(\frac{3}{2}+\epsilon\Big)+\epsilon.\end{aligned}$$

The last expression is strictly less than $\frac{1}{16}$ for sufficiently small ϵ.

2. Anti-entrenchment:

For $\hat{s}\in[0,\frac{1}{2}]$, $\zeta(\hat{s},q)\leq\frac{1}{2}q$. Thus,

$$\begin{aligned}\pi(\hat{s},q) &= \frac{1}{2}q[1-H_1(\hat{s})](1-w)+\left\{\frac{1}{2}qH_1(\hat{s})+\left(1-\frac{1}{2}q\right)H_0(\hat{s})\right\}\left(\frac{1}{2}q-k\right) \\ &\leq \frac{1}{2}q[1-H_1(\hat{s})](1-w)+\zeta(\hat{s},q)(1-w)H_0(\hat{s}) \\ &\leq \frac{1}{2}q(1-w)\big[1+H_0(\hat{s})-H_1(\hat{s})\big] \\ &\leq \frac{1}{4}w(1-w)\big[1-H_1(\hat{s})\big]\big[1+H_0(\hat{s})-H_1(\hat{s})\big] \leq \frac{1}{8}.\end{aligned}$$

Next, we consider a fixed contract $(w',k')=(\frac{4}{5},0)$. It can be verified that this contract will not yield an equilibrium with a replacement cutoff $\hat{s}$ below $\frac{1}{2}$. To see this, notice that the effort level under this contract is

$$q=\frac{2}{5}[1-H_1(\hat{s})].$$

The expected profit of replacement is,

$$\frac{1}{2}q - k = \frac{1}{5}[1 - H_1(\hat{s})].$$

The expected profit created by the manager on the margin is

$$\frac{\frac{1}{2}qh_1(\hat{s})}{\frac{1}{2}qh_1(\hat{s}) + \left(1 - \frac{1}{2}q\right)h_0(\hat{s})}(1-w) \le \frac{1}{2}q(1-w) = \frac{1}{25}[1 - H_1(\hat{s})], \text{ for } \hat{s} \in [0, \frac{1}{2}].$$

The indifference condition of the board never holds for $\hat{s} \in [0, \frac{1}{2}]$. Thus, the only possible equilibrium replacement policy under this contract is $\hat{s} > \frac{1}{2}$. It remains to prove that the profit of the contract is above $\frac{1}{8}$ for sufficiently large α.

Lemma A9 *For any $\Delta \in (0, \frac{1}{2})$, there exists N such that for $\alpha > N$, $\hat{s}(\alpha) < \frac{1}{2} + \Delta$ with contract $(w', k') = (\frac{4}{5}, 0)$.*

Proof. It suffices to prove that for any $\Delta \in (0, \frac{1}{2})$, there exists N such that for $\alpha > N$, the board's indifference condition never holds for all $\hat{s} \in [\frac{1}{2} + \Delta, 1]$ with contract $(w', k') = (\frac{4}{5}, 0)$.

The board's indifference condition can be simplified as

$$\frac{h_1(\hat{s};\alpha)}{h_0(\hat{s};\alpha)} = 1 + \frac{4}{H_1(\hat{s};\alpha)}.$$

Since $H_1(\hat{s}) \ge 2\hat{s} - 1$,

$$1 + \frac{4}{H_1(\hat{s};\alpha)} \le 1 + \frac{4}{H_1(\frac{1}{2} + \Delta;\alpha)} \le 1 + \frac{2}{\Delta}.$$

$\frac{h_1(\hat{s};\alpha)}{h_0(\hat{s};\alpha)}$ approaches infinity as $\alpha \to \infty$ while $1 + \frac{4}{H_1(\hat{s};\alpha)}$ is bounded, which is a contradiction. ■

For notational convenience, define $\Lambda(\hat{s};\alpha) = 1 - H_1(\hat{s})$. The board's expected profit can be written as

$$\begin{aligned}
&\pi\big(\hat{s}(w',k';\alpha),q(w',k';\alpha)\big)\\
&=\frac{1}{5}\Lambda^2(\hat{s};\alpha)\Big(\frac{7}{5}-\frac{2}{5}\Lambda(\hat{s};\alpha)\Big)+\frac{1}{5}\Lambda(\hat{s};\alpha)\Big(1-\frac{1}{5}\Lambda(\hat{s};\alpha)\Big)(2\hat{s}-1)\\
&\geq\frac{1}{5}\Lambda^2(\hat{s};\alpha)\Big(\frac{7}{5}-\frac{2}{5}\Lambda(\hat{s};\alpha)\Big)\geq\frac{1}{5}\Lambda^2(\hat{s};\alpha).
\end{aligned}$$

By Lemma A1, given any $\epsilon>0$, there exists N such that for $\alpha>N$, $\Lambda(\hat{s};\alpha)>2(1-\hat{s})-\epsilon$ for all $\hat{s}\in[\frac{1}{2},1]$. Thus,

$$\pi\big(\hat{s}(w',k';\alpha),q(w',k';\alpha)\big)\geq\frac{1}{5}\big[2(1-\hat{s})-\epsilon\big]^2\geq\frac{1}{5}(1-2\Delta-\epsilon)^2.$$

Let $\Delta=\epsilon=\frac{1}{24}$. Then,

$$\pi\big(\hat{s}(w',k';\alpha),q(w',k';\alpha)\big)\geq\frac{1}{5}(1-2\Delta-\epsilon)^2=\frac{49}{320}\geq\frac{1}{8}.$$

■

Appendix B: Normalization of information structure

In this section we first show that normalizing the signal space $\mathcal{S}$ to $[0,1]$ and assuming $\frac{1}{2}F_1(s)+\frac{1}{2}F_0(s)=s$ are without loss of generality. Next we show that the three assumptions imposed on $\{f_1(\cdot),f_0(\cdot)\}$ can be derived from similar assumptions on information structures without such normalization.

Suppose instead the board receives a noisy signal $x\in\mathcal{X}$ about the incumbent manager's ability θ_i. x is drawn from distribution with cdf $\tilde{F}_{\theta_i}(\cdot)$ and pdf $\tilde{f}_{\theta_i}(\cdot)$ for $\theta_i\in\{0,1\}$ with support $\mathcal{X}=[\underline{x},\overline{x}]$, where $-\infty\leq\underline{x}<\overline{x}\leq\infty$. Together with the signal space $\mathcal{X}$, the two conditional distributions $\{\tilde{f}_1(\cdot),\tilde{f}_0(\cdot)\}$ define an information structure.

Given an information structure $\{\tilde{f}_1(\cdot),\tilde{f}_0(\cdot),\mathcal{X}\}$, define a new signal x by applying the probability integral transformation to $x=\frac{1}{2}\tilde{F}_1(x)+\frac{1}{2}\tilde{F}_0(x)$. Then the unconditional distribution of s is uniform on $[0,1]$. Let $F_\theta(s)$ and $f_\theta(s)$ be the corresponding conditional cdf and pdf for $\theta_i\in\{0,1\}$ respectively. It can be verified that $\frac{1}{2}F_1(s)+\frac{1}{2}F_0(s)=s$ for all $s\in[0,1]$.

Assumption 6 *The monotone likelihood ratio property (MLRP): $\frac{\tilde{f}_1(x)}{\tilde{f}_0(x)}$ is strictly increasing in $x \in [\underline{x}, \overline{x}]$.*

Assumption 6 directly implies Assumption 1. For binary states, the MLRP assumption is without loss of generality since it can always be satisfied by relabeling signals according to the likelihood ratio.

Lemma B1 *Suppose two information structures $\{\tilde{f}_1(\cdot), \tilde{f}_0(\cdot), \mathcal{X}\}$ and $\{\tilde{f}_1^{\dagger}(\cdot), \tilde{f}_0^{\dagger}(\cdot), \mathcal{X}^{\dagger}\}$ generate the same distribution of posterior beliefs with prior $\Pr(\theta_i = 1) = \frac{1}{2}$. Then they yield the same distribution of posterior beliefs with all prior $\Pr(\theta_i = 1) \in (0, 1)$.*

The proof of Lemma B1 is similar to Lemma 3 and thus is omitted. Since entrenchment (anti-entrenchment) is defined by comparing the expected ability of the incumbent manager with that of the replacement manager, only the posterior belief about the incumbent manager matters. By Lemma B1, we can restrict attention to the information structures that satisfy $\frac{1}{2}F_1(s) + \frac{1}{2}F_0(s) = s$ for $s \in [0, 1]$ without loss of generality.

Assumption 7 *Perfectly informative at extreme signals: $\lim_{x \to \underline{x}} \frac{\tilde{f}_1(x)}{\tilde{f}_0(x)} = 0$ and $\lim_{x \to \overline{x}} \frac{\tilde{f}_1(x)}{\tilde{f}_0(x)} = +\infty$.*

Assumption 8 *There exists $\check{x} \in (\underline{x}, \overline{x})$ such that $\tilde{f}_0(x) = \tilde{f}_1(2\check{x} - x)$.*

Assumptions 7 and 8 directly imply Assumptions 2 and 3, respectively. We close this section by introducing two indexed families of information structures that satisfy Assumptions 6 – 8. The corresponding normalized signals after probability integral transformation also satisfy Assumption 4.

Example 2 (Normal Distribution) *Suppose $x = \theta_i + \epsilon$ for $\theta_i \in \{0, 1\}$, where $\epsilon \sim \mathcal{N}(0, \alpha^{-1})$. Then $x|\theta \sim \mathcal{N}(\theta, \alpha^{-1})$.*

Example 3 (Beta Distribution) *Suppose $\tilde{f}_1(x; \alpha) = (1 + \alpha)x^{\alpha}$ and $\tilde{f}_0(x; \alpha) = (1 + \alpha)(1 - x)^{\alpha}$ for $x \in [0, 1]$. Then $\tilde{F}_1(x; \alpha) = x^{1+\alpha}$ and $\tilde{F}_0(s; \alpha) = 1 - (1 - x)^{1+\alpha}$. This example is borrowed from Taylor and Yildirim (2011).*

For both examples, $\alpha \in (0, \infty)$ is interpreted as the informativeness of the information structure.

Appendix C: Properties of the ρ-concave order

By Lemma A5, the ρ-concave order implies the rotation order first introduced by Johnson and Myatt (2006) with $\Pr(\theta_i = \frac{1}{2})$. It can be verified that for a different prior, the rotation order does not remain. Intuitively, if the information structure becomes more informative, more densities concentrate on $p = 0$ and $p = 1$, and the distribution becomes more disperse.

Lemma C1 (Bayesian update) *Suppose $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the ρ-concave order. Then $\varphi(s|G_1) \geq \varphi(s|G_2)$ for $s \in (\frac{1}{2}, 1]$ and $\varphi(s|G_1) \leq \varphi(s|G_2)$ for $s \in (0, \frac{1}{2}]$.*

Proof. Since $G_1(0) = G_2(0) = 0$ and $G_1(\frac{1}{2}) = G_2(\frac{1}{2}) = \frac{1}{2}$ and $G_1(p) \geq G_2(p)$ for $p \in [0, \frac{1}{2}]$ by Lemma A5, $G_1^{-1}(s) \leq G_2^{-1}(s)$ for $s \in [0, \frac{1}{2}]$. Thus $\varphi(s|G_1) = \frac{G_1^{-1}(s)}{G_1^{-1}(s)+[1-G_1^{-1}(s)]} \leq \frac{G_2^{-1}(s)}{G_2^{-1}(s)+[1-G_2^{-1}(s)]} = \varphi(s, |G_2)$. The proof for $s \in (\frac{1}{2}, 1]$ is similar. ■

Lemma C1 shows the implication of the ρ-concave order on the Bayesian update of the incumbent manager's ability. The posterior belief $\varphi(x; \alpha)$ rotates counter-clockwise via $(\frac{1}{2}, \frac{1}{2})$ the as information structure becomes more informative. In other words, a fixed signal x has more information value to the board as the information structure becomes more informative.

Lemma C2 (Comparison with Blackwell's sufficiency) *If $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the ρ-concave order, $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the sense of Blackwell.*

Proof.

Lemma C3 *$F_1(s|G_1) \leq F_1(s|G_2)$ and $F_0(s|G_1) \geq F_0(s|G_2)$ for $s \in [0, 1]$.*

Proof. From the proof of Lemma C1, $G_1^{-1}(s) \leq G_2^{-1}(s)$ for $s \in [0, \frac{1}{2}]$.

1. For $s \in [0, \frac{1}{2}]$,

$$F_1(s|G_1) = \int_0^s f_1(t|G_1)dt = \int_0^s 2G_1^{-1}(t)dt \leq \int_0^s 2G_2^{-1}(t)dt = F_1(s|G_2).$$

2. For $s \in (\frac{1}{2}, 1]$,

$$\begin{aligned}
F_1(s|G_1) = \int_0^s f_1(t|G_1)dt &= \int_0^{1-s} f_1(t|G_1)dt + \int_{1-s}^{s} f_1(t|G_1)dt \\
&= \int_0^{1-s} f_1(t|G_1)dt + \frac{1}{2}(2s-1) \\
&\leq \int_0^{1-s} f_1(t|G_2)dt + \int_{1-s}^{s} f_1(t|G_2)dt = F_1(s|G_2).
\end{aligned}$$

Thus, $F_1(s|G_1) \leq F_1(s|G_2)$ for $s \in [0,1]$. Similarly, $F_0(s|G_1) \geq F_0(s|G_2)$. ■

Note that for binary states, Blackwell's order is equivalent to Lehmann's order. Thus, it suffices to prove that for $\omega \in (0,1)$,

$$F_1(F_0^{-1}(\omega|G_1)|G_1) \leq F_1(F_0^{-1}(\omega|G_2)|G_2).$$

Suppose we have the contrary, then there exists ω' such that,

$$F_1(F_0^{-1}(\omega'|G_1)|G_1) > F_1(F_0^{-1}(\omega'|G_2)|G_2).$$

By Lemma C3, it follows directly that $F_0^{-1}(\omega'|G_1) > F_0^{-1}(\omega'|G_2)$. However, $F_0^{-1}(\omega'|G_1) > F_0^{-1}(\omega'|G_2)$ cannot be true. To see this, let $s_1 = F_0^{-1}(\omega'|G_1)$ and $s_2 = F_0^{-1}(\omega'|G_2)$. Then $s_1 > s_2$ and $F_0(s_1|G_1) = F_0(s_2|G_1) = \omega'$. Again by Lemma C3, we have $F_0(s_1|G_1) > F_0(s_2|G_1) \geq F_0(s_2|G_2)$, which is a contradiction. ■

References

Almazan, Andres and Javier Suarez (2003), "Entrenchment and severance pay in optimal governance structures." *Journal of Finance*, 58, 519–547.

Anderson, Simon P. and Régis Renault (2003), "Efficiency and surplus bounds in Cournot competition." *Journal of Economic Theory*, 113, 253–264.

Casamatta, Catherine and Alexander Guembel (2010), "Managerial legacies, entrenchment, and strategic inertia." *Journal of Finance*, 65, 2403–2436.

Hermalin, Benjamin E. and Michael S. Weisbach (1998), "Endogenously chosen boards of directors and their monitoring of the CEO." *American Economic Review*, 88, 96–118.

Huang, Peggy (2011), "Marital prenups? a look at CEO severance agreements." *Available at SSRN 1928264*.

Inderst, Roman and Holger M. Mueller (2010), "CEO replacement under private information." *Review of Financial Studies*, 23, 2935–2969.

Johnson, Justin P. and David P. Myatt (2006), "On the simple economics of advertising, marketing, and product design." *American Economic Review*, 96, 756–784.

Kaplan, Steven N. and Bernadette A. Minton (2012), "How has CEO turnover changed?" *International Review of Finance*, 12, 57–87.

Laux, Volker (2008), "Board independence and CEO turnover." *Journal of Accounting Research*, 46, 137–171.

Laux, Volker (2014), "Corporate governance, board oversight, and CEO turnover." *Foundations and Trends in Accounting*, 8, 1–73.

Manso, Gustavo (2011), "Motivating innovation." *Journal of Finance*, 66, 1823–1860.

Mares, Vlad and Jeroen M. Swinkels (2014), "On the analysis of asymmetric first price auctions." *Journal of Economic Theory*, 152, 1–40.

Rusticus, Tjomme O. (2006), *Executive severance agreements*. Ph.D. thesis, University of Pennsylvania.

Shleifer, Andrei and Robert W. Vishny (1989), "Management entrenchment: The case of manager-specific investments." *Journal of Financial Economics*, 25, 123–139.

Taylor, Curtis R. and Huseyin Yildirim (2011), "Subjective performance and the value of blind evaluation." *Review of Economic Studies*, 78, 762–794.

Taylor, Lucian A. (2010), "Why are CEOs rarely fired? Evidence from structural estimation." *Journal of Finance*, 65, 2051–2087.

Weyl, E. Glen and Michal Fabinger (2013), "Pass-through as an economic tool: Principles of incidence under imperfect competition." *Journal of Political Economy*, 121, 528–583.