Managerial Turnover and Entrenchment

ONLINE APPENDIX

(Not Intended for Publication)

Zenan Wu [*] Xi Weng [†]

In this online appendix we include the materials omitted from the main text of the paper. The appendices are ordered according to where they are first referenced in the main text. In Appendix A we introduce the ρ-concave information order and characterize the equilibrium replacement policy for all α accordingly; in Appendix B we show that the main result derived in Propositions 1 and 2 are robust to the three extensions discussed in Section 6.2.

A ρ-concave Order and Optimal Replacement

In this section, we introduce the ρ-concave order, which is defined on the distribution of the board's posterior belief about the incumbent manager's ability, and characterize the equilibrium replacement policy for intermediate α.

A.1 Distribution of Posterior Beliefs

We first derive the distribution of the board's posterior belief about the incumbent's ability given an information structure $\left\{f_1(\cdot), f_0(\cdot)\right\}$. After observing signal s, the board's Bayesian update about θ_1, denoted by p, is as follows:

$$p := \varphi(s) \equiv \frac{f_1(s)}{f_1(s)+f_0(s)}.$$

Assumption 1 implies that $\varphi(s)$ is strictly increasing in s. Assumption 2 implies that the support of p is $[0,1]$. Let $g(p)$ and $G(\cdot)$ be the PDF and CDF of the posterior belief. Because $\mathbb{E}(\mathbb{E}(\theta|s)) = \frac{1}{2}$, the only constraint for $g(\cdot)$ is that $\int_0^1 pg(p)dp = \frac{1}{2}$. Given an information structure $\left\{f_1(\cdot), f_0(\cdot)\right\}$, $g(\cdot)$ can be derived as

$$g(p) = \frac{1}{2}\left[f_1(\varphi^{-1}(p)) + f_0(\varphi^{-1}(p))\right]\frac{d\varphi^{-1}(p)}{dp}. \tag{A1}$$

[*] School of Economics, Peking University, Beijing, China. Email: zenan@pku.edu.cn

[†] Guanghua School of Management, Peking University, Beijing, China. Email: wengxi125@gsm.pku.edu.cn

Lemma A1 *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 $\{f_1(\cdot), f_0(\cdot)\}$ that induces $g(\cdot)$.*

Proof. For existence, it suffices to construct an example. Consider the following information structure:

$$f_1(s) = 2G^{-1}(s), \text{ and } f_0(s) = 2[1 - G^{-1}(s)].$$

It is straightforward to see that

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

Therefore, we must have that

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

Moreover, we have that

$$p := \varphi(s) \equiv \frac{f_1(s)}{f_1(s) + f_0(s)} = G^{-1}(s).$$

The above equation, together with (A1), implies that the distribution of the posterior belief p is $g(\cdot)$ with the constructed information structure. This completes the proof of existence.

For uniqueness, suppose that there exist two information structures $\{f_1(s), f_0(s)\}$ and $\left\{f_1^\dagger(s), f_1^\dagger(s)\right\}$ that induce $g(\cdot)$. For notational convenience, we use the dagger symbol to refer to the variables when the information structure is $\left\{f_1^\dagger(s), f_1^\dagger(s)\right\}$. Therefore, we must have that

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

It follows from the definition of p that

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

The above condition, together with (A1), implies instantly that

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

which in turn implies that

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

Therefore, we must have that

$$\int_0^p tg(t)dt = \int_0^p \frac{1}{2}f_1(\varphi^{-1}(t))\frac{d\varphi^{-1}(t)}{dt}dt = \frac{1}{2}F_1(\varphi^{-1}(p)) = \frac{1}{2}F_1^\dagger(\varphi^{\dagger-1}(p)).$$

Similarly, we have that

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

which in turn implies that

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

Therefore, we must have that

$$\begin{aligned}\varphi^{\dagger-1}(p) &\equiv \frac{1}{2} F_1^\dagger(\varphi^{\dagger-1}(p)) + \frac{1}{2} F_0^\dagger(\varphi^{\dagger-1}(p)) \\ &= \int_0^p (1-t)g(t)dt + \int_0^p tg(t)dt \\ &\equiv \frac{1}{2} F_1(\varphi^{-1}(p)) + \frac{1}{2} F_0(\varphi^{-1}(p)) \equiv \varphi^{-1}(p).\end{aligned}$$

The above equation implies $\varphi^{\dagger-1}(p) = \varphi^{-1}(p)$ and hence $f_1^\dagger(s) = f_1(s)$ for all $s \in [0,1]$; together with the fact that $\frac{1}{2} f_1^\dagger(s) + \frac{1}{2} f_0^\dagger(s) = \frac{1}{2} f_1(s) + \frac{1}{2} f_0(s)$, we must have that $f_0^\dagger(s) = f_0(s)$ for all $s \in [0,1]$. This completes the proof of uniqueness. ■

Lemma A1 states that there exists a one-to-one mapping between $g(\cdot)$ and information structure $\{f_1(\cdot), f_0(\cdot)\}$. Therefore, working on the information structure $\{f_1(\cdot), f_0(\cdot)\}$ is equivalent to working on the distribution of the posterior belief $g(\cdot)$. Consequently, we can define information order on $g(\cdot)$. Moreover, Assumption 3 implies that $g(p) = g(1-p)$ and $G(p) = 1 - G(1-p)$ for all $p \in [0,1]$. This implies that it suffices to order different information structures based on $G(p)$ for $p \in [0, \frac{1}{2}]$.

A.2 ρ-concave Order

Next, we use ρ-concavity to define the informativeness of the information structure. Given $G(\cdot)$, we define local ρ-concavity at p as follows:

$$\rho(p) := 1 - \frac{G(p)g'(p)}{g^2(p)}. \tag{A2}$$

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. It can be verified from (A2) that log-concavity of $G(\cdot)$ is equivalent to $\rho(p) \geq 0$, and concavity of $G(\cdot)$ is equivalent to $\rho(p) \geq 1$. Next, we focus on $G(\cdot)$ such that $\rho(p) \in (0, \infty)$. This assumption is a necessary condition to guarantee the initial condition $G(0) = 0$ is satisfied.[1]

Definition 2 (ρ-concave order) *Consider two CDFs $G_1(p)$ and $G_2(p)$ with support $[0,1]$ that satisfy $G_1(p) + G_1(1-p) = G_2(p) + G_2(1-p) = 1$ for all $p \in [0,1]$. $G_1(p)$ 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 $p \in [0, \frac{1}{2}]$.*

[1] It is useful to point out that a completely uninformative information structure is well-defined with this constraint.

The above definition states that $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the ρ-concave order if $G_1(\cdot)$ is everywhere more concave than $G_2(\cdot)$ as measured by local ρ-concavity. Suppose that $\max_{p\in[0,\frac{1}{2}]}\{\rho(p)\}$ and $\min_{p\in[0,\frac{1}{2}]}\{\rho(p)\}$ exist, and denote the corresponding maximum and minimum as $\overline{\rho}$ and $\underline{\rho}$ respectively. The following lemma characterizes the shape of $G(\cdot)$.

Lemma A2 *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 all* $p\in[0,\frac{1}{2}]$.

Proof. From the definition of $\underline{\rho}$ and $\overline{\rho}$, we have that

$$\underline{\rho}\leq 1-\frac{G(p)g'(p)}{g^2(p)}\leq\overline{\rho},\ \text{for all } p\in\left[0,\frac{1}{2}\right].$$

Integrating all sides of the above inequality from 0 to p yields

$$\underline{\rho}p\leq\frac{G(p)}{g(p)}-\frac{G(0)}{g(0)}\leq\overline{\rho}p\Leftrightarrow\frac{1}{\overline{\rho}}\frac{1}{p}\leq\frac{g(p)}{G(p)}\leq\frac{1}{\underline{\rho}}\frac{1}{p}\Leftrightarrow\frac{1}{\overline{\rho}}\leq\frac{pg(p)}{G(p)}\leq\frac{1}{\underline{\rho}}.$$

Integrating all sides of the above inequality from p to $\frac{1}{2}$ yields

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

This completes the proof. ■

Lemma A2 implies that $G(\cdot)$ can be bounded by two CDFs with constant ρ-concavity. It is straightforward that a completely informative information structure corresponds to $\lim_{\alpha\to\infty}\underline{\rho}(\alpha)=\infty$ and a completely uninformative information structure is equivalent to $\lim_{\alpha\to\infty}\overline{\rho}(\alpha)=0$. The following assumptions are imposed on the family of distribution $\{G(\cdot;\alpha)\}$ indexed by $\alpha\in(0,\infty)$.

Assumption A1 *(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* $p\in[0,\frac{1}{2}]$.[2]

(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* $\lim_{\alpha\to 0}\overline{\rho}(\alpha)=0$.

It follows from Assumption A1(a) that $G(p;\alpha)$ is log-concave in $p\in[0,\frac{1}{2}]$ for all $\alpha\in(0,\infty)$. Assumption A1(c) and Assumption A1(d) guarantee that $G(\cdot)$ is either concave or convex for a given α. Last, Assumption A1(e) simply restates Assumption 4 in the language of ρ-concavity.

[2]As we will show later, this assumption guarantees that the profit function for $p\in[0,\frac{1}{2}]$ is well-behaved in the sense that it is either monotonic or single-peaked.

A.3 Equilibrium Replacement Policy under Optimal Contract

Denote $\hat{p}$ as the cutoff of the posterior belief and $\tilde{\pi}(\hat{p})$ as the board's profit in terms of $\hat{p}$. Then $\tilde{\pi}(\hat{p})$ can be derived as

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

It can be verified that the sum of the selection effect and the commitment cost effect is increasing in $\hat{p}$ and thus is maximized at $\hat{p} = 1$. Taking the first-order derivative of $\tilde{\pi}(\hat{p})$ with respect to $\hat{p}$ yields

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

$$\Rightarrow \ \tilde{\pi}'(\hat{p}) \lesseqqgtr 0 \ \Leftrightarrow \ \frac{\hat{p}g(\hat{p})}{G(\hat{p})} \gtreqqless \frac{\int_{\hat{p}}^{1} tg(t)dt}{1-\int_{\hat{p}}^{1} G(t)dt}. \tag{A3}$$

From the above first-order condition, $\hat{p}g(\hat{p})$ is the marginal incentive effect and $G(\hat{p})$ is the marginal selection plus commitment cost 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 cost 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 $p \in [0, \frac{1}{2}]$ [see Assumption A1(c)]. This assumption guarantees that the marginal incentive effect changes faster than the marginal selection plus commitment cost effect, and hence yields a well-behaved profit function for $\hat{p} \in [0, \frac{1}{2}]$. Assumption A1(b) guarantees that the marginal selection plus commitment cost 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 cost effect takes over as the board's monitoring technology improves and anti-entrenchment is more likely to emerge. The next proposition formalizes the above discussions.

Proposition A1 *Suppose the family of distribution* $\{G(\cdot;\alpha)\}$*, indexed by* $\alpha \in (0,\infty)$*, satisfies Assumption A1. Then there exist two cutoffs* α_1 *and* α_2 *with* $\alpha_2 > \alpha_1 > 0$ *such that: (i)* $\hat{s}^*(\alpha) = 0$ *for* $\alpha \in (0, \alpha_1]$*; (ii)* $\hat{s}^*(\alpha) \in (0, \frac{1}{2})$ *for* $\alpha \in (\alpha_1, \alpha_2)$*; (iii)* $\hat{s}^*(\alpha) \in (\frac{1}{2}, 1)$ *for* $\alpha \in (\alpha_2, \infty)$*. Moreover,* α_1 *satisfies* $\rho(p;\alpha_1) = 1 \ \forall\, p \in [0, \frac{1}{2}]$*.*

Proof. It is useful to prove several intermediate results.

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

Proof. We complete the proof in the following two steps:

Step I: Inducing anti-entrenchment is never optimal. It suffices to show that $\tilde{\pi}(1-\hat{p}) < \tilde{\pi}(0)$ for all $\hat{p} \in [0, \frac{1}{2}]$, which is equivalent to

$$\int_{1-\hat{p}}^{1} tg(t)dt\Big(1 - \int_{1-\hat{p}}^{1} G(t)dt\Big) < \int_{0}^{1} tg(t)dt\Big(1 - \int_{0}^{1} G(t)dt\Big). \tag{A4}$$

Because $G(1-\hat{p}) = 1 - G(\hat{p})$, $\int_0^1 G(t)dt = \frac{1}{2}$. The right-hand side of (A4) can be simplified as

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

while the left-hand side of (A4) can be bounded above by

$$\begin{aligned}
\int_{1-\hat{p}}^{1} tg(t)dt\Big(1 - \int_{1-\hat{p}}^{1} G(t)dt\Big) &= \Big(1 - \int_{1-\hat{p}}^{1} G(t)dt - (1-\hat{p})G(1-\hat{p})\Big)\Big(1 - \int_{1-\hat{p}}^{1} G(t)dt\Big) \\
&< \Big(1 - \int_{0}^{\hat{p}} \big[1 - G(t)\big]\, dt - \frac{1}{2}(1-\hat{p})[1 - G(\hat{p})]\Big)^2 \\
&= \Big(\frac{1-\hat{p}}{2}\big[1 + G(\hat{p})\big] + \int_{0}^{\hat{p}} G(t)dt\Big)^2 \\
&\le \Big(\frac{1-\hat{p}}{2}(1+\hat{p}) + \int_{0}^{\hat{p}} tdt\Big)^2 = \frac{1}{4},
\end{aligned}$$

where the last inequality follows from the fact that $G(t) \le t$ due to the convexity of $G(t)$ for $t \in [0, \frac{1}{2}]$. Therefore, condition (A4) holds.

Step II: $\tilde{\pi}(\hat{p})$ is strictly decreasing in $\hat{p}$ for $\hat{p} \in [0, \frac{1}{2}]$ if $G(\cdot)$ is convex in p for $p \in [0, \frac{1}{2}]$. First, notice that

$$\begin{aligned}
\int_{\hat{p}}^{1} tg(t)dt = \int_{\hat{p}}^{1} tdG(t) &= 1 - \int_{\hat{p}}^{1} G(t)dt - \hat{p}G(\hat{p}) \\
&< 1 - \int_{\hat{p}}^{1} G(t)dt, \text{ for all } \hat{p} \in (0, \frac{1}{2}].
\end{aligned} \tag{A5}$$

Second, note that $g(\hat{p})$ is increasing in $\hat{p}$ from the convexity of $G(\cdot)$. Therefore, we must have that

$$G(\hat{p}) = \int_{0}^{\hat{p}} g(t)dt \le \int_{0}^{\hat{p}} g(\hat{p})dt = \hat{p}g(\hat{p}). \tag{A6}$$

Inequalities (A5) and (A6) together imply that

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

Therefore, we have that $\tilde{\pi}'(p) < 0$ for all $p \in (0, \frac{1}{2}]$ from (A3). This completes the proof. ■

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

Proof. From the definition of ρ-concavity, we have that

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

Integrating both sides of the above equality from 0 to p yields

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

Therefore, we have that

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

This completes the proof. ■

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

Proof. It follows from the proof of Lemma A4 that

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

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 from the definition of the ρ-concave order. Therefore, $G(p;\alpha)$ is increasing in α. This completes the proof. ■

We can now prove Proposition A1. Lemma A3 implies instantly that $\hat{p}^* = 0$ for $\alpha \leq \alpha_1$. For $\alpha > \alpha_1$, recall that $\tilde{\pi}'(\hat{p}) \gtreqless 0$ is equivalent to

$$\frac{G(\hat{p})}{\hat{p}g(\hat{p})} \gtreqless \frac{1 - \int_{\hat{p}}^1 G(t)dt}{\int_{\hat{p}}^1 tg(t)dt} = \frac{\frac{1}{2} + \int_0^{\hat{p}} G(t)dt}{\frac{1}{2} + \int_0^{\hat{p}} G(t)dt - \hat{p}G(\hat{p})}. \tag{A7}$$

Lemma A4 implies that the left-hand side of (A7) is decreasing in $\hat{p}$. In addition, it can be verified that the right-hand side of (A7) is increasing in $\hat{p}$. Therefore, $\tilde{\pi}'(\hat{p})$ is either monotonic or single-peaked for $\hat{p} \in [0, \frac{1}{2}]$. Notice that $\lim_{p\to 0} \frac{G(\hat{p})}{\hat{p}g(\hat{p})} = \rho(0) > 1$ for $\alpha > \alpha_1$, and $\lim_{p\to 0} \frac{1-\int_{\hat{p}}^1 G(t)dt}{\int_{\hat{p}}^1 tg(t)dt} = 1$. Therefore, it suffices to consider the following two cases:

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

Case II: $2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt < \frac{\frac{1}{2}+\int_0^{\frac{1}{2}} G(t)dt}{\frac{1}{4}+\int_0^{\frac{1}{2}} G(t)dt}$. Then $\tilde{\pi}(\hat{p})$ is first increasing and then decreasing in $\hat{p} \in [0, \frac{1}{2}]$. The maximum can be pinned down by the first-order condition for $\hat{p} \in [0, \frac{1}{2}]$. Next we show that this local maximum is indeed the global maximum for $\hat{p} \in [0, 1]$. To see this, notice that second-order derivative of the profit function with respect to $\hat{p}$ can be derived as follows:

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

Because $G(p)$ is concave for $p \in [0, \frac{1}{2}]$ if $\alpha > \alpha_1$, we have that $G(p)$ is convex for $p \in [\frac{1}{2}, 1]$, which in turn implies that $g'(p) > 0$ for $p \in [\frac{1}{2}, 1]$. Therefore, $\pi''(\hat{p}) < 0$ for all $p \in [\frac{1}{2}, 1]$. Together with the fact that $\tilde{\pi}'(\frac{1}{2}) < 0$, we have that $\tilde{\pi}(\hat{p})$ is decreasing in $\hat{p}$ for $\hat{p} \in [\frac{1}{2}, 1]$. This indicates that the local maximum for $p \in [0, \frac{1}{2}]$ is also the global maximum for $p \in [0, 1]$.

Last, we prove the existence of the cutoff α_2. On the one hand, it follows from the definition of ρ-concavity that $2\int_0^{\frac{1}{2}} \rho(t;\alpha)dt$ is increasing in α. On the other hand, Lemma A5 states that $\int_0^{\frac{1}{2}} G(t;\alpha)dt$ is increasing in α, implying that $\frac{\frac{1}{2}+\int_0^{\frac{1}{2}} G(t)dt}{\frac{1}{4}+\int_0^{\frac{1}{2}} G(t)dt}$ is decreasing in α. Moreover, it follows from Assumptions A1(d) and A1(e) that

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

Therefore, there exists a threshold α_2 above which anti-entrenchment is optimal and below which entrenchment is optimal. This completes the proof. ■

A.4 Properties of ρ-concave Order

In this section, we discuss some properties of the ρ-concave order. First, it follows from Lemma A5 and $G(\frac{1}{2}) = \frac{1}{2}$ that the ρ-concave order implies the rotation order introduced by Johnson and Myatt (2006). Next, we show that the ρ-concave order is a stronger condition than Blackwell's order: if a family of distributions is ordered according to the ρ-concave order, then it is also ordered in the sense of Blackwell.

Proposition A2 (Comparison with Blackwell's Sufficiency) *If $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the ρ-concave order, then $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the sense of Blackwell.*

Proof. It is useful to first prove the following two intermediary results.

Lemma A6 *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. First, notice that $G_1(0) = G_2(0) = 0$ and $G_1(\frac{1}{2}) = G_2(\frac{1}{2}) = \frac{1}{2}$. Together with the fact that $G_1(p) \geq G_2(p)$ for all $p \in [0, \frac{1}{2}]$ from Lemma A5, we must have that $G_1^{-1}(s) \leq G_2^{-1}(s)$ for all

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

$$\varphi(s|G_1) = \frac{G_1^{-1}(s)}{G_1^{-1}(s) + [1 - G_1^{-1}(s)]} \leq \frac{G_2^{-1}(s)}{G_2^{-1}(s) + [1 - G_2^{-1}(s)]} = \varphi(s, |G_2), \text{ for } s \in (0, \frac{1}{2}].$$

The proof for $s \in (\frac{1}{2}, 1]$ is similar. This completes the proof. ■

Lemma A7 *Suppose $G_1(\cdot)$ is more informative than $G_2(\cdot)$ in the ρ-concave order. Then $F_1(s|G_1) \leq F_1(s|G_2)$ and $F_0(s|G_1) \geq F_0(s|G_2)$ for all $s \in [0, 1]$.*

Proof. It follows from the proof of Lemma A6 that $G_1^{-1}(s) \leq G_2^{-1}(s)$ for all $s \in [0, \frac{1}{2}]$.

Case I: $s \in [0, \frac{1}{2}]$**.** Then we have that

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

Case II: $s \in (\frac{1}{2}, 1]$**.** Similarly, we have that

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

Thus, we have that $F_1(s|G_1) \leq F_1(s|G_2)$ for all $s \in [0, 1]$, which in turn implies that

$$F_0(s|G_1) \equiv 2s - F_1(s|G_1) \geq 2s - F_1(s|G_2) \equiv F_0(s|G_2).$$

This completes the proof. ■

Now we can prove Proposition A2. Note that for binary states, Blackwell's order is equivalent to Lehmann's order. Therefore, it suffices to show that

$$F_1(F_0^{-1}(\omega|G_1)|G_1) \leq F_1(F_0^{-1}(\omega|G_2)|G_2), \text{ for all } \omega \in (0, 1).$$

Suppose to the contrary that there exists $\omega' \in (0, 1)$ such that

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

It follows from Lemma A7 that $F_0^{-1}(\omega'|G_1) > F_0^{-1}(\omega'|G_2)$. However, $F_0^{-1}(\omega'|G_1) > F_0^{-1}(\omega'|G_2)$ does not hold. To see this, let $s_1 := F_0^{-1}(\omega'|G_1)$ and $s_2 := F_0^{-1}(\omega'|G_2)$. It follows immediately that $s_1 > s_2$ and $F_0(s_1|G_1) = F_0(s_2|G_1) = \omega'$. Moreover, from Lemma A7 we must have that $F_0(s_1|G_1) > F_0(s_2|G_1) \geq F_0(s_2|G_2)$, which is a contradiction. This completes the proof. ■

B Extensions and Robustness

B.1 Variance of Incumbent Manager's Ability.

Recall from Section 6.2 that 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 $e_1^{1+\tau}\theta$, where $\tau \in (-1, 1)$. Then we have the following proposition as α becomes sufficiently large:

Proposition A3 *Suppose that $\theta \in \left\{\frac{1}{2}-\delta, \frac{1}{2}+\delta\right\}$ and the success probability is equal to $e_1^{1+\tau}\theta$. Then:*

i. if $\delta > \frac{1}{2}^{\frac{1-\tau}{2}} - \frac{1}{2}$, there exists $\bar{\alpha}_A$ such that anti-entrenchment emerges under the optimal contract if $\alpha > \bar{\alpha}_A$;

ii. if $\delta < \frac{1}{2}^{\frac{1-\tau}{2}} - \frac{1}{2}$, there exists $\bar{\alpha}_E$ such that entrenchment emerges under the optimal contract if $\alpha > \bar{\alpha}_E$.

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

$$e_1 = \left\{(1+\tau)\Big[\frac{1}{2}\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \frac{1}{2}\big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]w\right\}^{\frac{1}{1-\tau}}.$$

Similarly, the board's indifference condition is

$$\frac{1}{2}e_1^{1+\tau} - k = \left[\frac{1}{2} + \frac{f_1(\hat{s}) - f_0(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\delta\right] e_1^{1+\tau}(1-w).$$

Let $\pi_a(\hat{s})$ denote the board's expected profit by inducing the equilibrium cutoff $\hat{s}$. We have that

$$\begin{aligned}\pi_a(\hat{s}) = \mathcal{M} \cdot &\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]^{\frac{1+\tau}{1-\tau}} \\ &\times \left\{\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big] + \big[F_1(\hat{s}) + F_0(\hat{s})\big]\Big[\frac{1}{2} + \frac{f_1(\hat{s}) - f_0(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\delta\Big]\right\},\end{aligned}$$

where $\mathcal{M} := \frac{1}{4}^{\frac{2}{1-\tau}}(1-\tau)(1+\tau)^{\frac{2+2\tau}{1-\tau}}$. Fixing a replacement cutoff $\hat{s}$, the board's expected profit as $\alpha \to \infty$ can be derived as follows:

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

Entrenchment as $\alpha \to \infty$. Notice that for all $\hat{s} \in [0,1]$, $\pi(\hat{s};\alpha)$ can be bounded above by

$$\begin{aligned}\pi_a(\hat{s};\alpha) \leq & \mathcal{M} \cdot \Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]^{\frac{1+\tau}{1-\tau}} \\ & \times \left\{\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big] + \big[F_1(\hat{s})+F_0(\hat{s})\big]\Big[\frac{1}{2}+\delta\Big]\right\} =: \pi_E(\hat{s};\alpha),\end{aligned}$$

where the inequality follows from $\frac{f_1(\hat{s})-f_0(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} \leq 1$. Lemma 3 states that $F_1(\hat{s};\alpha)$ converges uniformly to $\max\big\{0, 2\hat{s}-1\big\}$ as $\alpha \to \infty$. Thus, $\pi_E(\hat{s};\alpha)$ converges uniformly to

$$\mathcal{M} \cdot \Big[(1+2\delta)(1-\hat{s})\Big]^{\frac{1+\tau}{1-\tau}}(1+2\delta), \text{ for all } \hat{s} \in \Big[\frac{1}{2}, 1\Big] \text{ as } \alpha \to \infty.$$

Note that $\pi_a(0;\alpha) = \mathcal{M}$. Therefore, inducing entrenchment is optimal for sufficiently large α if

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

Anti-entrenchment as $\alpha \to \infty$. Similar to the proof of entrenchment, $\pi(\hat{s};\alpha)$ can be bounded above by

$$\begin{aligned}\pi_a(\hat{s};\alpha) \leq & \mathcal{M} \cdot \Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big]^{\frac{1+\tau}{1-\tau}} \\ & \times \left\{\Big[\big(\frac{1}{2}+\delta\big)\big[1-F_1(\hat{s})\big] + \big(\frac{1}{2}-\delta\big)\big[1-F_0(\hat{s})\big]\Big] + \frac{1}{2}\big[F_1(\hat{s})+F_0(\hat{s})\big]\right\} =: \pi_A(\hat{s};\alpha),\end{aligned}$$

where the inequality follows from the fact that $\frac{f_1(\hat{s})-f_0(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} \leq 0$ for $\hat{s} \in [0, \frac{1}{2}]$. Again, Lemma 3 implies that $F_1(\hat{s};\alpha)$ converges uniformly to $\max\big\{0, 2\hat{s}-1\big\}$. Thus, $\pi_A(\hat{s};\alpha)$ converges uniformly to

$$\xi(\hat{s};\delta,\tau) := \mathcal{M} \cdot \Big[1-(1-2\delta)\hat{s}\Big]^{\frac{1+\tau}{1-\tau}} \cdot (1+2\delta\hat{s}), \text{ for } \hat{s} \in \Big[0, \frac{1}{2}\Big] \text{ as } \alpha \to \infty.$$

Fixing $(\delta,\tau) \in (0,\frac{1}{2}] \times (-1,1)$, it can be verified that there exists $\nu(\delta,\tau) < \frac{1}{2}$ such that

$$\xi(\hat{s};\delta,\tau) < \mathcal{M} \cdot \Big[\delta + \frac{1}{2}\Big]^{\frac{1+\tau}{1-\tau}} \cdot (1+2\delta) = \lim_{\alpha\downarrow\frac{1}{2}} \lim_{\alpha\to\infty} \pi_a(\hat{s};\alpha), \text{ for } \hat{s} \in \Big[\nu(\delta,\tau), \frac{1}{2}\Big].$$

Therefore, inducing $\hat{s} \in [\nu(\delta,\tau), \frac{1}{2}]$ is not optimal as $\alpha \to \infty$. Moreover, 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 $\alpha \to \infty$. Therefore, $\pi_a(\hat{s};\alpha)$ converges uniformly to $\mathcal{M}\Big[1-(1-2\delta)\hat{s}\Big]^{\frac{1+\tau}{1-\tau}}$ for $\hat{s} \in [0,\nu(\delta,\tau)]$ as $\alpha \to \infty$; and inducing entrenchment is optimal for sufficiently large α if

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

This completes the proof. ■

B.2 Signal of Outcome Instead of Ability

In this section, we extend the baseline model to allow the signal to be related to the incumbent manager's effort. Therefore, fixing a replacement cutoff, the incumbent is able to increase the probability of his retention by exerting more effort. To formalize the idea, suppose $\lambda = 0$ for simplicity, and the signal s is drawn from a distribution with density $h_y(\cdot)$ and CDF $H_y(\cdot)$ for $y \in \{0, 1\}$. Similar to the baseline model in Section 4, we assume the pair of density functions $\{h_1(\cdot), h_0(\cdot)\}$ satisfies Assumptions 1 - 4. The social planner chooses $(\hat{s}, e_1)$ to maximize

$$\max_{\{\hat{s}, e_1\}} \frac{1}{2}e_1\left[1 - H_1(\hat{s})\right] + \frac{1}{2}e_1\left[\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) H_0(\hat{s})\right] - \frac{1}{2}e_1^2.$$

It can be verified that the socially optimal replacement cutoff is $\frac{1}{2}$ and the period-1 effort is

$$\frac{1 + H_0(\frac{1}{2}) - H_1(\frac{1}{2})}{2 + H_0(\frac{1}{2}) - H_1(\frac{1}{2})}.$$

The intuition is as follows. Given period-1 effort e_1 and signal s, the board's Bayesian update of the incumbent manager's ability, denoted by $\varphi_h(s, e_1)$, is

$$\varphi_h(s, e_1) = \frac{\frac{1}{2}e_1 h_1(s) + \frac{1}{2}(1 - e_1)h_0(s)}{\frac{1}{2}e_1 h_1(s) + \left(1 - \frac{1}{2}e_1\right) h_0(s)}.$$

Note that $\varphi_h(\frac{1}{2}, e_1) = \frac{1}{2}$ independent of period-1 effort e_1 and of the informativeness of the information structure α. Therefore, it is socially optimal to replace the incumbent manager if and only if the posterior belief of his ability falls below the prior.

Next, we consider the incumbent's effort choice fixing a contract (w, k) and belief about the replacement cutoff $\hat{s}$. The manager chooses e_1 to maximize

$$\frac{1}{2}e_1[1 - H_1(\hat{s})]w + \left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) H_0(\hat{s})\right\} k - \frac{1}{2}e_1^2.$$

The first-order condition for the above maximization problem with respect to e_1 yields

$$e_1 = \max\left\{\frac{1}{2}\left[1 - H_1(\hat{s})\right] w - \frac{1}{2}\left[H_0(\hat{s}) - H_1(\hat{s})\right] k, 0\right\}.$$

Unlike (2) in the baseline model, it follows from the above condition that the incumbent manager is directly dis-incentivized by severance pay. An increase in severance pay increases the opportunity cost of exerting effort and leads to a decrease in effort. If the severance pay is high enough, the incumbent manager is willing to be fired and exerts no effort. In this extension, severance pay is a double-edged sword. By the direct effect (better outside option if the incumbent is replaced), severance pay decreases effort. By the indirect effect (better job security with a lower equilibrium replacement cutoff), it increases effort. The design of the optimal contract should take into consideration this non-trivial effect of severance pay on period-1 effort.

Fix (w, k) and belief about e_1. The board chooses the replacement cutoff $\hat{s}$ to maximize

$$\frac{1}{2}e_1\big[1 - H_1(\hat{s})\big](1-w) + \left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) H_0(\hat{s})\right\}\left(\frac{1}{2}e_1 - k\right).$$

The board's indifference condition is

$$\frac{\frac{1}{2}e_1 h_1(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \big(1 - \frac{1}{2}e_1\big)h_0(\hat{s})}(1-w) = \frac{1}{2}e_1 - k.$$

Carrying out the algebra, $\hat{s}$ is the solution to

$$\zeta(\hat{s}, e_1) = \max\left\{\min\left\{\frac{\frac{1}{2}e_1 - k}{1-w}, 1\right\}, 0\right\},$$

where $\zeta(\hat{s}, e_1)$ is the estimate of the outcome under the incumbent's management at $\hat{s}$ given e_1:

$$\zeta(\hat{s}, e_1) := \frac{\frac{1}{2}e_1 h_1(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) h_0(\hat{s})}.$$

It is difficult to write the expected profit as a function of $\hat{s}$ alone as in Section 4 because e_1 is now affected by the contract (w, k) directly as well as the equilibrium cutoff $\hat{s}$ indirectly. However, we can still discuss the optimal replacement policy under extreme cases where α goes to 0 and to ∞. Note that multiple equilibria may exist for some contract (w, k) because incentive on effort e_1 is not monotonic in k as in the baseline model in Section 4. To proceed, we further assume that the equilibrium most favorable to the board is selected when multiple equilibria exist.

Proposition A4 *Suppose $\lambda = 0$ and the board receives a signal of outcome under the management of the incumbent (i.e., $\theta_1 e_1$) rather than the his ability (i.e., θ_1). Then $\hat{s}^*(\alpha) > \frac{1}{2}$ if α is sufficiently large, and $\hat{s}^*(\alpha) < \frac{1}{2}$ if α is sufficiently small.*

Proof. Fixing w and the equilibrium cutoff $\hat{s}$, the board's expected profit, denoted by $\pi_h(\hat{s}, w)$, is as follows:

$$\pi_h(\hat{s}, w) := \frac{1}{2}e_1\left[1 - H_1(\hat{s})\right](1-w) + \left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) H_0(\hat{s})\right\}\left(\frac{1}{2}e_1 - k\right).$$

s.t.

$$e_1 = \max\left\{\frac{1}{2}\left[1 - H_1(\hat{s})\right] w - \frac{1}{2}\left[H_0(\hat{s}) - H_1(\hat{s})\right] k, 0\right\},$$

and

$$\zeta(\hat{s}, e_1)(1-w) = \frac{1}{2}e_1 - k.$$

Entrenchment. It can be verified that the maximum profit to induce $\hat{s} = 0$ is $\frac{1}{16}$. The profit can be obtained by offering a contract with wage $w = \frac{1}{2}$ and a sufficiently large k. Similarly, inducing $\hat{s} = 1$ leads to $e_1 = 0$ and generates zero profit to the board, and cannot be optimal. Therefore, it suffices to prove that $\pi_h(\hat{s}, w) < \frac{1}{16}$ for all $\hat{s} \in [\frac{1}{2}, 1]$ as α becomes sufficiently small. It is useful to prove several intermediate results.

Lemma A8 *There exists N_h such that for $\alpha < N_h$, $\pi_h(\hat{s}, w) < \frac{1}{16}$ for all $\hat{s} \in [\frac{31}{32}, 1]$.*

Proof. Note that the period-1 effort can be bounded above by

$$e_1 = \max\left\{\frac{1}{2}\left[1 - H_1(\hat{s})\right] w - \frac{1}{2}\left[H_0(\hat{s}) - H_1(\hat{s})\right] k, 0\right\} \leq \frac{1}{2}\left[1 - H_1(\hat{s})\right] w, \tag{A8}$$

where the last inequality follows from that fact that $w \geq 0$. Thus, the expected profit can be bounded above by

$$\begin{aligned}
\pi_h(\hat{s}, w) &= \frac{1}{2}e_1[1 - H_1(\hat{s})](1-w) + \left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) H_0(\hat{s})\right\}\left(\frac{1}{2}e_1 - k\right) \\
&\leq \frac{1}{2}e_1[1 - H_1(\hat{s})](1-w) + \frac{1}{2}e_1\left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) H_0(\hat{s})\right\} \\
&\leq \frac{1}{2}e_1\left[\left[1 - H_1(\hat{s})\right] + 1\right] \leq e_1 \leq \frac{1}{2}\left[1 - H_1(\hat{s})\right] w < 1 - H_1(\hat{s}),
\end{aligned}$$

where the first inequality follows from $k \geq 0$, the second inequality follows from $H_1(\hat{s}) \leq H_0(\hat{s}) \leq 1$ and $w \geq 0$, the third inequality follows from $H_1(\hat{s}) \geq 0$, and the fourth inequality follows from (A8).

It follows from Lemma 4 that for $\epsilon'_h = \frac{1}{32}$, there exists N_h such that for $\alpha < N_h$, $H_1(\hat{s}) \geq \hat{s} - \epsilon'_h$ for all $\hat{s} \in [0, 1]$. Therefore, for $\alpha < N_h$, we have that

$$\pi_h(\hat{s}, w) < 1 - H_1(\hat{s}) \leq 1 - \hat{s} + \epsilon'_h \leq \frac{1}{32} + \epsilon'_h = \frac{1}{16}, \text{ for all } \hat{s} \in \left[\frac{31}{32}, 1\right].$$

This completes the proof. ■

Lemma A9 *Fix $e_1 \in [0, 1]$. For any $\epsilon_h > 0$, there exists N'_h such that for $\alpha < N'_h$,*

$$\frac{\frac{1}{2}e_1 h_1(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right) h_0(\hat{s})} \leq \frac{1}{2}e_1 + \epsilon_h, \text{ for all } \hat{s} \in \left[\frac{1}{2}, \frac{31}{32}\right].$$

Proof. Let $\epsilon' := \frac{\epsilon_h}{1+\epsilon_h}$. It follows from the definition of the completely uninformative information structure that there exists N'_h such that $h_1(\frac{31}{32};\alpha) < 1+\epsilon'$ for $\alpha < N'_h$. Therefore, we have that

$$\frac{\frac{1}{2}e_1 h_1(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \left(1-\frac{1}{2}e_1\right) h_0(\hat{s})} - \frac{1}{2}e_1 = \frac{1}{2}e_1\left(1-\frac{1}{2}e_1\right)\frac{h_1(\hat{s}) - h_0(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \left(1-\frac{1}{2}e_1\right) h_0(\hat{s})} \leq \frac{1}{2}\frac{h_1(\hat{s}) - h_0(\hat{s})}{h_0(\hat{s})},$$

where the inequality follows from the fact that $e_1(1-\frac{1}{2}e_1) \leq 1$ and $h_1(\hat{s}) \geq 1 \geq h_0(\hat{s})$ for $\hat{s} \geq \frac{1}{2}$. Next, it follows from $h_1(\hat{s}) + h_0(\hat{s}) = 2$ for all $\hat{s} \in [0,1]$ that

$$\frac{1}{2}\frac{h_1(\hat{s}) - h_0(\hat{s})}{h_0(\hat{s})} = \frac{h_1(\hat{s}) - 1}{2 - h_1(\hat{s})} \leq \frac{h_1(\frac{31}{32};\alpha) - 1}{2 - h_1(\frac{31}{32};\alpha)} \leq \frac{\epsilon'}{1-\epsilon'} \equiv \epsilon_h,$$

where the first inequality follows from the fact that $h_1(\hat{s})$ in strictly increasing in $\hat{s}$ and $\hat{s} \leq \frac{31}{32}$. This completes the proof. ■

Now we can prove the entrenchment result. Lemma A8 states that inducing $\hat{s} \in [\frac{32}{33}, 1]$ is not optimal as $\alpha \to 0$. Moreover, Lemma 4 implies that for all $\epsilon > 0$, we have that $H_1(\hat{s}) \geq \hat{s} - \epsilon$ for any $\hat{s} \in [0,1]$ as α becomes sufficiently small. Therefore, as $\alpha \to 0$, $\pi_h(\hat{s}, w)$ can be bounded above by

$$\begin{aligned}
\pi_h(\hat{s}, w) &\leq \frac{1}{2}e_1[1 - H_1(\hat{s})](1-w) + \left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1-\frac{1}{2}e_1\right) H_0(\hat{s})\right\}\left(\frac{1}{2}e_1 + \epsilon_h\right)(1-w) \\
&\leq \frac{1}{2}e_1(1-w)\left[\left[1 - H_1(\hat{s})\right] + \frac{1}{2}e_1 H_1(\hat{s}) + \left(1-\frac{1}{2}e_1\right) H_0(\hat{s})\right] + \epsilon_h \\
&\leq \frac{1}{4}\left[1 - H_1(\hat{s})\right]\left[2 - H_1(\hat{s})\right] w(1-w) + \epsilon_h \\
&\leq \frac{1}{16}\left[1 - H_1(\hat{s})\right]\left[2 - H_1(\hat{s})\right] + \epsilon_h \\
&\leq \frac{1}{16}(1 - \hat{s} + \epsilon)(2 - \hat{s} + \epsilon) + \epsilon_h \\
&\leq \frac{1}{16}\left(\frac{1}{2} + \epsilon\right)\left(\frac{3}{2} + \epsilon\right) + \epsilon_h, \text{ for all } \hat{s} \in \left[\frac{1}{2}, \frac{31}{32}\right].
\end{aligned}$$

The first inequality follows from Lemma A9 and the board's indifference condition; the second inequality follows from $w \geq 0$ and $H_1(\hat{s}) \leq H_0(\hat{s}) \leq 1$; the third inequality follows from (A8) and again $H_1(\hat{s}) \leq H_0(\hat{s}) \leq 1$; the fourth inequality follows from $w(1-w) \leq \frac{1}{4}$; the fifth inequality follows from Lemma 4; and the last inequality follows from $\hat{s} \geq \frac{1}{2}$. Note that the last expression can be arbitrarily close to $\frac{1}{16} \times \frac{1}{2} \times \frac{3}{2} = \frac{3}{64}$ as $\epsilon \to 0$ and $\epsilon_h \to 0$, which is strictly less than $\frac{1}{16}$. This completes the proof of entrenchment.

Anti-entrenchment. First, for all $\hat{s} \in [0, \frac{1}{2}]$, we have that

$$\zeta(\hat{s}, e_1) = \frac{\frac{1}{2}e_1 h_1(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \left(1-\frac{1}{2}e_1\right) h_0(\hat{s})} \leq \frac{\frac{1}{2}e_1 h_1(\hat{s})}{\frac{1}{2}e_1 h_1(\hat{s}) + \left(1-\frac{1}{2}e_1\right) h_1(\hat{s})} = \frac{1}{2}e_1. \tag{A9}$$

Thus, $\pi_h(\hat{s}, w)$ can be bounded above by

$$\begin{aligned}\pi_h(\hat{s}, w) &\equiv \frac{1}{2}e_1\left[1 - H_1(\hat{s})\right](1-w) + \left\{\frac{1}{2}e_1 H_1(\hat{s}) + \left(1 - \frac{1}{2}e_1\right)H_0(\hat{s})\right\}\left(\frac{1}{2}e_1 - k\right) \\ &\leq \frac{1}{2}e_1[1 - H_1(\hat{s})](1-w) + \zeta(\hat{s}, e_1)(1-w)H_0(\hat{s}) \\ &\leq \frac{1}{2}e_1(1-w)\left[1 + H_0(\hat{s}) - H_1(\hat{s})\right] \\ &\leq \frac{1}{4}w(1-w)\left[1 - H_1(\hat{s})\right]\left[1 + H_0(\hat{s}) - H_1(\hat{s})\right] \leq \frac{1}{8}, \text{ for all } \hat{s} \in \left[0, \frac{1}{2}\right] \text{ and } w \in [0,1],\end{aligned}$$

where the first inequality follows from $H_1(\hat{s}) \leq H_0(\hat{s})$, the second inequality follows from (A9), and the third inequality follows from (A8).

Next, we consider a fixed contract $(w_h, k_h) := (\frac{4}{5}, 0)$. First, we show that this contract induces anti-entrenchment for all α. To see this, notice that the effort level under this contract is

$$e_1 = \frac{2}{5}\left[1 - H_1(\hat{s})\right], \tag{A10}$$

and hence the expected profit of replacement is

$$\frac{1}{2}e_1 - k_h = \frac{1}{5}[1 - H_1(\hat{s})]. \tag{A11}$$

On the other hand, the board's expected profit from retaining the incumbent manager after observing signal $s \in [0, \frac{1}{2}]$ is

$$\zeta(s, e_1)(1 - w_h) \leq \frac{1}{2}e_1(1 - w_h) = \frac{1}{25}[1 - H_1(\hat{s})] < \frac{1}{5}[1 - H_1(\hat{s})] = \frac{1}{2}e_1 - k_h, \text{ for all } s \in \left[0, \frac{1}{2}\right],$$

where the first inequality follows from (A9), the first equality follows from $w_h = \frac{4}{5}$ and (A10), and the last equality follows from (A11). Therefore, the equilibrium replacement policy under this contract must be that $\hat{s} > \frac{1}{2}$.

Second, we show that the profit under the contract (w_h, k_h) is strictly greater than $\frac{1}{8}$ as $\alpha \to \infty$.

Lemma A10 *The equilibrium cutoff under the optimal contract* $\hat{s}(\alpha) < \frac{13}{24}$ *with contract* $(w_h, k_h) = (\frac{4}{5}, 0)$ *as* $\alpha \to \infty$.

Proof. It suffices to prove that the board's indifference condition does not hold for all $\hat{s} \in [\frac{13}{24}, 1]$ under contract $(w_h, k_h) = (\frac{4}{5}, 0)$ as $\alpha \to \infty$. First, note that the board's indifference condition can be simplified as

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

Moreover, we have that

$$1 + \frac{4}{H_1(\hat{s};\alpha)} \leq 1 + \frac{4}{H_1(\frac{13}{24};\alpha)} \leq 1 + \frac{4}{\frac{1}{24}} = 97, \text{ for all } \hat{s} \in \left[\frac{13}{24}, 1\right].$$

where the first inequality follows from the postulated $\hat{s} \geq \frac{13}{24}$ and the second inequality follows from

$$H_1\left(\frac{13}{24};\alpha\right) = \int_0^{\frac{1}{2}} h_1(s;\alpha)ds + \int_{\frac{1}{2}}^{\frac{13}{24}} h_1(s;\alpha)ds \geq 0 + \int_{\frac{1}{2}}^{\frac{13}{24}} ds = \frac{1}{24}.$$

Therefore, $1+\frac{4}{H_1(\hat{s};\alpha)}$ is bounded above by 97 while $\frac{h_1(\hat{s};\alpha)}{h_0(\hat{s};\alpha)}$ approaches infinity from Assumption 4 as $\alpha \to \infty$, a contradiction. This completes the proof. ■

Now we can prove the anti-entrenchment result. The board's expected profit can be bounded above by

$$\begin{aligned}\pi_h\left(\hat{s}(w_h,k_h;\alpha),w_h\right) &= \frac{1}{5}\left[1-H_1(\hat{s};\alpha)\right]^2\left(\frac{7}{5}-\frac{2}{5}\left[1-H_1(\hat{s};\alpha)\right]\right)\\ &\quad+\frac{1}{5}\left[1-H_1(\hat{s};\alpha)\right]\left(1-\frac{1}{5}\left[1-H_1(\hat{s};\alpha)\right]\right)(2\hat{s}-1)\\ &\geq \frac{1}{5}\left[1-H_1(\hat{s};\alpha)\right]^2\left(\frac{7}{5}-\frac{2}{5}\left[1-H_1(\hat{s};\alpha)\right]\right) \geq \frac{1}{5}\left[1-H_1(\hat{s};\alpha)\right]^2,\end{aligned}$$

where the first inequality follows from $\hat{s} \geq \frac{1}{2}$ and the second inequality follows from $H_1(\hat{s};\alpha) \geq 0$. It follows from Lemma 3 that given any $\epsilon > 0$,

$$1 - H_1(\hat{s};\alpha) > 2(1-\hat{s}) - \epsilon, \text{ for all } \hat{s} \in \left[\frac{1}{2},1\right] \text{ as } \alpha \to \infty. \tag{A12}$$

Therefore, we have that

$$\pi_h\left(\hat{s}(w_h,k_h;\alpha),w_h\right) \geq \frac{1}{5}\left[2\left[1-\hat{s}(w_h,k_h;\alpha)\right]-\epsilon\right]^2 \geq \frac{1}{5}\left(\frac{11}{12}-\epsilon\right)^2, \text{ as } \alpha \to \infty,$$

where the first inequality follows from (A12), and the second inequality follows from Lemma A10. Let $\epsilon = \frac{1}{24}$. Then the above inequality reduces to

$$\pi_h\left(\hat{s}(w_h,k_h;\alpha),w_h\right) \geq \frac{1}{5}\left(\frac{11}{12}-\epsilon\right)^2 = \frac{49}{320} > \frac{1}{8}.$$

This completes the proof. ■

B.3 Signal is private information to the board

In this section, we assume that the signal is *private information* to the board and show that the results derived in Proposition 2 are robust. Fixing contract (w,k), the incumbent manager's best response to cutoff $\hat{s}$ is the effort profile (e_1,e_2) that maximizes

$$\int_{\hat{s}}^1 \left\{\frac{f_1(s)}{f_1(s)+f_0(s)}\left[(1-\lambda)e_1+\lambda e_2\right]w - \frac{1}{2}e_2^2\right\} d\frac{F_1(s)+F_0(s)}{2} + \frac{F_1(\hat{s})+F_0(\hat{s})}{2}k - \frac{1}{2}e_1^2.$$

The first-order conditions with respect to e_1 and e_2 yield

$$e_1(\hat{s}; w, k) = \frac{1 - F_1(\hat{s})}{2}(1-\lambda)w, \tag{A13}$$

and

$$e_2(\hat{s}; w, k) = \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}\lambda w. \tag{A14}$$

Similarly, for a fixed contract (w, k) and belief about effort (e_1, e_2), the board chooses cutoff $\hat{s}$ to maximize

$$\int_{\hat{s}}^{1} \left\{ \frac{f_1(s)}{f_1(s) + f_0(s)} \left[(1-\lambda)e_1 + \lambda e_2\right](1-w) \right\} d\frac{F_1(s) + F_0(s)}{2} + \frac{1}{2}[F_1(\hat{s}) + F_0(\hat{s})]\left[\underline{\pi}(e_1) - k\right].$$

$$\Rightarrow \hat{s}(e_1, e_2; w, k) \text{ solves } \frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\left[(1-\lambda)e_1 + \lambda e_2\right](1-w) = \underline{\pi}(e_1) - k.^{3} \tag{A15}$$

Given contract (w, k), the optimal cutoff and effort $\big(\hat{s}(w,k), e_1(w,k), e_2(w,k)\big)$ are pinned down by equations (A13), (A14), and (A15).[4] Alternatively, we can calculate the corresponding severance pay k and effort decision (e_1, e_2) given the wage rate w and the equilibrium cutoff $\hat{s}$ that the board would like to induce as follows:

$$e_1(\hat{s}, w) = \frac{1 - F_1(\hat{s})}{2}(1-\lambda), \tag{A16}$$

$$e_2(\hat{s}, w) = \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}\lambda w, \tag{A17}$$

and

$$k(\hat{s}, w) = \underline{\pi}(e_1(\hat{s}, w)) - \frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\left[(1-\lambda)e_1(\hat{s}, w) + \lambda e_2(\hat{s}, w)\right](1-w). \tag{A18}$$

As it also does in the analysis in Section 5, the board chooses $(w, \hat{s})$ to maximize expected profit as follows:

$$\begin{aligned}\pi_p(\hat{s}, w) :=& w(1-w) \times \left\{ (1-\lambda)^2 \frac{1 - F_1(\hat{s})}{2} + \lambda^2 \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]} \right\} \\ & \times \left\{ \frac{1}{2}\left[1 - F_1(\hat{s})\right] + \frac{1}{2}\left[F_1(\hat{s}) + F_0(\hat{s})\right]\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})} \right\},\end{aligned} \tag{A19}$$

subject to the non-negativity constraint for k, which is $k(\hat{s}, w) \geq 0$. It is straightforward to see

[3] It is without loss of generality to assume that indifference condition (A15) always holds. We consider three possibilities: (i) $\hat{s} \in (0,1)$; (ii) $\hat{s} = 1$; (iii) $\hat{s} = 0$. First, it is clear that condition (A15) holds if the board would like to induce $\hat{s} \in (0,1)$. Second, it is never optimal to induce $\hat{s} = 1$. To see this, notice that if $\hat{s} = 1$, then the incumbent manager exerts no effort in the first period, and it can be verified that firm's maximum profit is $\frac{1}{16}\lambda^2$. On the other hand, if firm chooses $(w, k) = (\frac{1}{2}, \infty)$, the incumbent manager would expect $\hat{s} = 0$. This contract generates an expected profit of $\frac{1}{16}[\lambda^2 + (1-\lambda)^2]$, which is greater than $\frac{1}{16}\lambda^2$. Third, we consider the case where the board would like to induce $\hat{s} = 0$ and condition (A15) does not hold. Then it must be the case that the left-hand side of condition (A15) is strictly greater than the right-hand side for all $\hat{s} \in [0,1]$, that is, $\underline{\pi}(e_1) - k < 0$. In that case, the board can set $k = \underline{\pi}(e_1)$ without changing the expected profit, and condition (A15) holds at $\hat{s} = 0$.

[4] Again, it is assumed that the equilibrium most favorable to the board is selected when multiple equilibria exist.

that $w^* = \frac{1}{2}$ under the optimal contract if the non-negativity constraint for k does not bind. In this case, the above expected profit function (A19) can be rewritten in terms of $\hat{s}$ alone as follows:

$$\pi_p\left(\hat{s}, \frac{1}{2}\right) = \frac{1}{4} \times \left\{(1-\lambda)^2 \underbrace{\frac{1-F_1(\hat{s})}{2}}_{\text{incentive effect}} + \lambda^2 \underbrace{\frac{1-F_1(\hat{s})}{[1-F_1(\hat{s})]+[1-F_0(\hat{s})]}}_{\text{learning effect}}\right\}$$
$$\times \left\{\underbrace{\left[\frac{1}{2}[1-F_1(\hat{s})] + \frac{1}{4}\left[F_1(\hat{s})+F_0(\hat{s})\right]\right]}_{\text{selection effect}} + \underbrace{\frac{1}{2}\left[F_1(\hat{s})+F_0(\hat{s})\right]\left(\frac{f_1(\hat{s})}{f_1(\hat{s})+f_0(\hat{s})} - \frac{1}{2}\right)}_{\text{commitment cost effect}}\right\}.$$

Proposition A5 *Suppose* $\{f_1(\cdot;\alpha), f_0(\cdot;\alpha)\}$ *satisfies Assumptions 1 - 4 and the signal is private information to the board. Then,*

i. anti-entrenchment emerges in the optimal contract if α *is sufficiently large for* $\lambda \in (0, \sqrt{2}-1]$*;*

ii. entrenchment emerges in the optimal contract if α *is sufficiently small for* $\lambda \in (0, 1]$*.*

Proof. Define $\hat{\mathcal{S}}_p(\alpha)$ as the set of cutoffs that can be induced by contracts that satisfy $w = \frac{1}{2}$ and $k \geq 0$, that is,

$$\hat{\mathcal{S}}_p(\alpha) := \left\{\hat{s}\,\middle|\,k(\hat{s}, \frac{1}{2}) \geq 0 \;\&\; \hat{s} \in [0,1]\right\}.$$

If $\hat{s} \in \hat{\mathcal{S}}_p(\alpha)$, the board's expected profit can be written as

$$\bar{\pi}_p(\hat{s}) := \pi_p\left(\hat{s}, \frac{1}{2}\right) \equiv \frac{1}{16}\left[1-F_1(\hat{s})\right]\left[(1-\lambda)^2 + \lambda^2 \frac{1}{1-\hat{s}}\right]\left[[1-F_1(\hat{s})] + \hat{s} f_1(\hat{s})\right].$$

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

Entrenchment. Note that $\frac{1}{2} \in \mathcal{W}_p(0;\alpha)$, and hence it suffices to prove that $\pi_p(\hat{s}, w;\alpha) < \bar{\pi}_p(0;\alpha)$ for all $\hat{s} \in [\frac{1}{2}, 1]$ and $w \in \mathcal{W}_p(\hat{s};\alpha)$ if α is sufficiently small. Moreover, $\bar{\pi}_p(0;\alpha)$ is independent of α and can be calculated as follows:

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

It is useful to prove the following lemma.

Lemma A11 *There exists* $\Delta_p \in (0, \frac{1}{2})$ *such that* $\pi_p(\hat{s}, w;\alpha) < \bar{\pi}_p(0;\alpha)$ *for all* $\hat{s} \in [1-\Delta_p, 1]$ *and* $w \in \mathcal{W}_p(\hat{s};\alpha)$ *as* $\alpha \to 0$*.*

Proof. It follows from Lemma 4 that fixing any $\epsilon > 0$, $1 - F_1(1-\Delta_p) < \Delta_p + \epsilon$ for all $\Delta_p \in [0,1]$

as $\alpha \to 0$. The board's expected profit after replacement can be bounded above by

$$\underline{\pi}(e_1) - k \leq \frac{1}{4}\Big(\frac{1}{2}\lambda + \frac{1-\lambda}{\lambda}e_1\Big)^2 = \frac{1}{16}\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]w\Big)^2$$
$$\leq \frac{1}{16}\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]\Big)^2, \tag{A20}$$

where the first inequality follows from the observation that $\frac{1}{4}\Big(\frac{1}{2}\lambda + \frac{1-\lambda}{\lambda}e_1\Big)^2 \geq \frac{1}{2}(1-\lambda)e_1$ for all e_1, and the second inequality follows from $w \leq 1$. Thus, for all $\hat{s} \in [1-\Delta_p, 1]$ and $w \in [0,1]$, the board's expected profit as $\alpha \to 0$ can be bounded above by

$$\begin{aligned}
\pi_p(\hat{s}, w; \alpha) &\leq \frac{1}{2}\big[1 - F_1(\hat{s})\big]\left\{\frac{1 - F_1(\hat{s})}{2}(1-\lambda)^2 + \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}\lambda^2\right\} w(1-w) \\
&\quad + \frac{1}{32}\big[F_1(\hat{s}) + F_0(\hat{s})\big]\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]\Big)^2 \\
&\leq \frac{1}{8}\big[1 - F_1(\hat{s})\big]\left\{\frac{1 - F_1(\hat{s})}{2}(1-\lambda)^2 + \frac{1 - F_1(\hat{s})}{[1 - F_1(\hat{s})] + [1 - F_0(\hat{s})]}\lambda^2\right\} \\
&\quad + \frac{1}{16}\Big(\lambda + \frac{(1-\lambda)^2}{\lambda}\big[1 - F_1(\hat{s})\big]\Big)^2 \\
&< \frac{1}{8}(\Delta_p + \epsilon)\Big[\frac{\Delta_p + \epsilon}{2}(1-\lambda)^2 + \lambda^2\Big] + \frac{1}{16}\Big[\lambda + \frac{(1-\lambda)^2}{\lambda}(\Delta_p + \epsilon)\Big]^2,
\end{aligned}$$

where the first inequality follows from (A20), the second inequality follows from $w(1-w) \leq \frac{1}{4}$ and $F_1(\hat{s}) + F_0(\hat{s}) \leq 2$, and the third inequality follows from Lemma 4. Note that the last expression is strictly increasing in $\Delta_p + \epsilon$, and is approaching $\frac{1}{16}\lambda^2$ as $\Delta_p \to 0$ and $\epsilon \to 0$. Hence we can always find sufficiently small Δ_p and ϵ such that

$$\frac{1}{8}(\Delta_p + \epsilon)\Big[\frac{\Delta_p + \epsilon}{2}(1-\lambda)^2 + \lambda^2\Big] + \frac{1}{16}\Big[\lambda + \frac{(1-\lambda)^2}{\lambda}(\Delta_p + \epsilon)\Big]^2 < \frac{1}{16}\big[(1-\lambda)^2 + \lambda^2\big] \equiv \bar{\pi}_p(0; \alpha).$$

This completes the proof. ■

We can now prove the entrenchment result. Lemma A11 states that $\hat{s} \in [1-\Delta_p, 1]$ cannot be the equilibrium replacement cutoff under the optimal contract as $\alpha \to 0$, and hence it suffices to show that this is also the case for $\hat{s} \in [\frac{1}{2}, 1-\Delta_p]$. Note that $\bar{\pi}_p(\hat{s}; \alpha)$ is the maximum expected profit without the limited liability constraint for k. Thus $\pi_p(\hat{s}, w; \alpha) \leq \bar{\pi}_p(\hat{s}; \alpha)$ for all $\hat{s} \in [0,1]$.

Fixing $\epsilon > 0$, Lemma 4 and the definition of completely uninformative information structure imply that $1 - F_1(\hat{s}) \geq \hat{s} - \epsilon$ for all $\hat{s} \in [0,1]$, and $f_1(\hat{s}) \leq 1 + \epsilon$ for all $\hat{s} \in [\frac{1}{2}, 1-\Delta_p]$ as $\alpha \to 0$.

Therefore, $\pi_p(\hat{s}, w; \alpha)$ can be bounded above by

$$\begin{aligned}
\pi_p(\hat{s}, w; \alpha) \leq \bar{\pi}_p(\hat{s}; \alpha) &\equiv \frac{1}{16}\Big[1 - F_1(\hat{s})\Big]\Big[(1-\lambda)^2 + \lambda^2 \frac{1}{1-\hat{s}}\Big]\Big[[1 - F_1(\hat{s})] + \hat{s} f_1(\hat{s})\Big] \\
&\leq \frac{1}{16}(1 - \hat{s} + \epsilon)\Big[(1-\lambda)^2 + \lambda^2 \frac{1}{1-\hat{s}}\Big]\Big[(1 - \hat{s} + \epsilon) + \hat{s}(1+\epsilon)\Big] \\
&= \frac{1}{16}\frac{1 - \hat{s} + \epsilon}{1-\hat{s}}\Big[(1-\hat{s})(1-\lambda)^2 + \lambda^2\Big]\Big[(1 - \hat{s} + \epsilon) + \hat{s}(1+\epsilon)\Big] \\
&\leq \frac{1}{16}\Big(1 + \frac{\epsilon}{\Delta_p}\Big)\Big[\frac{1}{2}(1-\lambda)^2 + \lambda^2\Big](1 + 2\epsilon),
\end{aligned}$$

where the last inequality follows from $\hat{s} \in [\frac{1}{2}, 1 - \Delta_p]$. Note that the last expression is approaching $\frac{1}{16}\Big[\frac{1}{2}(1-\lambda)^2 + \lambda^2\Big]$ as $\epsilon \to 0$, and hence it remains to show 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. This completes the proof of entrenchment.

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

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

where the last inequality follows from $1 - F_1(\hat{s}) \leq 1$ and $f_1(\hat{s}) \leq 1$ for $\hat{s} \in [0, \frac{1}{2}]$. Therefore, it remains to construct a feasible contract that induces an equilibrium cutoff $\hat{s} > \frac{1}{2}$ and yields a profit no less than $\frac{3}{32}[(1-\lambda)^2 + 2\lambda^2]$. For notational convenience, let $\psi' := \big(\frac{\lambda}{1-\lambda}\big)^2$. It follows from $\lambda < \sqrt{2} - 1$ that $\psi' < \frac{1}{2}$. From the limited liability constraint for k, we have that

$$\frac{f_1(\hat{s})}{f_1(\hat{s}) + f_0(\hat{s})}\big[(1-\lambda)e_1 + \lambda e_2\big](1 - w) \leq \underline{\pi}(e_1).$$

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

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

and

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

The second inequality comes from the construction whereby 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 the tuple (κ', ι') that yields a higher expected profit than $\frac{3}{32}\Big[(1-\lambda)^2 + 2\lambda^2\Big]$ given λ. Note that Lemma 3 implies that $\frac{1 - F_1(\hat{s})}{2}$ can be arbitrarily close to $1 - \hat{s}$ when α is sufficiently large. Thus, the above two inequalities 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

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

which is equivalent to

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

Let $\iota' := \frac{\frac{1}{2}\psi'}{\frac{1}{2}-\kappa'+\psi'}$. Then ι' is well-defined if $\frac{\frac{1}{2}\psi'}{\frac{1}{2}-\kappa'+\psi'} \geq \frac{\psi'}{1-2\kappa'} - \frac{1}{2}$. Note that this condition holds when for sufficiently small κ' because

$$\lim_{\kappa'\to 0} \frac{\frac{1}{2}\psi'}{\frac{1}{2}-\kappa'+\psi'} = \frac{\frac{1}{2}\psi'}{\frac{1}{2}+\psi'} > 0 \geq \psi' - \frac{1}{2} = \lim_{\kappa'\to 0} \frac{\psi'}{1-2\kappa'} - \frac{1}{2}.$$

Next, note that the board's expected profit from offering a period-1 contract with wage $w' = \frac{1}{2} + \iota'$ that induces $\hat{s}' = \frac{1}{2} + \kappa'$ as $\alpha \to \infty$ is

$$\lim_{\alpha\to\infty} \pi_p(\hat{s}', w'; \alpha) = \left[(1-\lambda)^2 \left(\frac{1}{2}-\kappa'\right) + \lambda^2\right] \times \left[\frac{1}{4} - \left(\frac{\frac{1}{2}\psi'}{\frac{1}{2}-\kappa'+\psi'}\right)^2\right].$$

Therefore, we have that

$$\begin{aligned}\lim_{\kappa'\to 0}\lim_{\alpha\to\infty} \pi_p(\hat{s}', w'; \alpha) &= \lim_{\kappa'\to 0}\left\{\left[(1-\lambda)^2 \left(\frac{1}{2}-\kappa'\right) + \lambda^2\right] \times \left[\frac{1}{4} - \left(\frac{\frac{1}{2}\psi'}{\frac{1}{2}-\kappa'+\psi'}\right)^2\right]\right\} \\ &= \frac{1}{2}\left[(1-\lambda)^2 + 2\lambda^2\right] \times \left[\frac{1}{4} - \left(\frac{\frac{1}{2}\psi'}{\frac{1}{2}+\psi'}\right)^2\right] > \frac{3}{32}\left[(1-\lambda)^2 + 2\lambda^2\right],\end{aligned}$$

where the inequality follows from the fact that $\psi' < \frac{1}{2}$. Thus, we can find a sufficiently small κ' such that $\lim_{\alpha\to\infty} \pi_p(\hat{s}', w'; \alpha) > \frac{3}{32}\left[(1-\lambda)^2 + 2\lambda^2\right]$. That is, anti-entrenchment emerges under the optimal contract when α is sufficiently large and $\psi' < \frac{1}{2}$, or equivalently, $\lambda < \sqrt{2} - 1$. This completes the proof. ■

References

Johnson, Justin P. and David P. Myatt, "On the simple economics of advertising, marketing, and product design," *American Economic Review*, 2006, *96* (3), 756–784.