Internet Appendix for "The Labor Market for Directors and Externalities in Corporate Governance"

DORON LEVIT and NADYA MALENKO*

The Internet Appendix has three sections. Section I contains supplemental materials for the proofs of the main results. Section II contains a formal analysis of directors' welfare. Section III contains the extensions of the basic model described in Section II.C of the published article, as well as an additional extension that endogenizes directors' benefit from a new directorship.

I. Supplemental Results

Supplemental material for the proof of Lemma 1: Consider the case in which some directors are never pivotal. This case gives rise to multiple equilibria. If a director is pivotal with probability zero, his action does not affect his utility, and hence any strategy of this director is optimal given the strategies of other directors. Hence, there exist two symmetric equilibria, where all directors vote for the proposal or vote against the proposal regardless of their type, as well as asymmetric equilibria (for example, a strategy profile where $T+1$ directors always vote for the proposal and $K-T-1$ directors always vote against the proposal constitutes an equilibrium). However, as we show next, the trembling hand refinement eliminates all equilibria where directors are

*Citation format: Levit, Doron, and Nadya Malenko, Internet Appendix for "The Labor Market for Directors and Externalities in Corporate Governance," *Journal of Finance* [DOI STRING]. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

pivotal with probability zero, and the only trembling hand perfect equilibria are those where each director k in each firm i plays a threshold strategy and votes for the proposal if and only if $\theta_{ik} \geq \theta_i^*$ for some finite θ_i^*.

Consider any equilibrium and any director k_0 in firm i. Consider further any sequence $\{\sigma_n\}$, $\sigma_n \in (0,1)$, $\lim_{n\to\infty} \sigma_n = 1$. If all directors of firm i except k_0 play their respective equilibrium strategy with probability σ_n and the opposite strategy with probability $1-\sigma_n$, then director k_0's probability of being pivotal is strictly positive. Conditional on being pivotal, the director's utility from voting for shareholder control relative to voting against it is given by (A1), and hence the director votes for shareholder control if and only if (A2) is satisfied.

It follows that any equilibrium where some director is pivotal with probability zero does not satisfy the refinement. Indeed, being pivotal with probability zero requires that either at least T other directors of the firm vote for shareholder control with probability one, or at least $K-T+1$ other directors of the firm vote for shareholder control with probability zero. Consider any director k_0 whose equilibrium strategy does not depend on his type, that is, who votes for (against) shareholder control regardless of his type. Consider the best response of this director in the perturbed game. According to (A2), since $\lim_{\theta\to\infty} \Delta(\theta) = \infty$ and $\lim_{\theta\to-\infty} \Delta(\theta) = -\infty$, and since the right-hand side of (A2) does not depend on θ, there exists $\underline{\theta}$ $(\bar{\theta})$ such that for any $\sigma_n < 1$, the director finds it optimal to vote against (for) shareholder control if his type is below $\underline{\theta}$ (above $\bar{\theta}$). Hence, there exist types whose best response is not their equilibrium decision for any n, and thus the equilibrium does not survive the refinement.

Thus, the only equilibria that could survive the refinement are those where each director is pivotal with a positive probability, which implies, as shown above, that these are threshold equilibria. We next show that any threshold equilibrium survives the refinement. Indeed, conditional on being pivotal, the director's relative utility from voting for shareholder control in the perturbed game is given by (A1), which is the same as in equilibrium since τ_j and beliefs are the same. Hence, for any n, the director's best response is his equilibrium strategy, which confirms that the equilibrium survives the refinement.

Proof of Proposition 1, part (iii): Note that $\frac{\partial\beta(\theta,\alpha)}{\partial\theta} = -2\frac{\alpha\delta}{K}\tau'(\theta)\cdot d\left(\frac{\alpha\delta}{K}(1-2\tau(\theta))\right)$, where $d(\cdot) = (\Delta^{-1})'$ and

$$\tau'(\theta) = \sum_{t=T}^{K} C_t^K \left[(K-t)(1-F(\theta))^t F(\theta)^{K-t-1} - t(1-F(\theta))^{t-1} F(\theta)^{K-t}\right] f(\theta).$$

Since $\tau(\theta) \in [0,1]$ and $d(\cdot)$ is continuous, there exists $\bar{d}$ such that $\left|d\left(\frac{\alpha\delta}{K}(1-2\tau(\theta))\right)\right| < \bar{d}$ for all θ and all $\alpha < \bar{\alpha}$. Since $F(\theta) \in [0,1]$ and, by assumption, $f(\cdot)$ is bounded, there exists $\bar{f}$ such that $|\tau'(\theta)| < \bar{f}$ for all $\theta \in R$. It follows that $\frac{\partial\beta(\theta,\alpha)}{\partial\theta} < 2\frac{\alpha\delta}{K}\bar{f}\bar{d}$ and hence there exists $\underline{\alpha} \in (0,\bar{\alpha})$ such that for any $\alpha < \underline{\alpha}$, $\frac{\partial\beta(\theta,\alpha)}{\partial\theta} < 1$.

We next prove that for any $\alpha < \underline{\alpha}$, the equilibrium is unique. Suppose this is not true, and $\theta_1 < \theta_2$ are two equilibria for some $\alpha < \underline{\alpha}$. Then $\Psi(\theta_1,\alpha) = \Psi(\theta_2,\alpha) = 0$, and since $\Psi(\theta,\alpha)$ is continuous and differentiable, by the mean value theorem, there exists $\hat{\theta} \in (\theta_1,\theta_2)$ such that $\frac{\partial\Psi(\hat{\theta},\alpha)}{\partial\theta} = 0 \Leftrightarrow \frac{\partial\beta(\hat{\theta},\alpha)}{\partial\theta} = 1$. This contradicts the fact that $\frac{\partial\beta(\theta,\alpha)}{\partial\theta} < 1$ for all θ. Note that $\alpha < \underline{\alpha}$ is a sufficient, but not necessary, condition for the equilibrium to be unique.

Supplemental material for the proof of Proposition 2: We first prove that $\bar{\theta}^*$ and $\underline{\theta}^*$ exist. Since $\beta(\theta)$ is bounded, there exists B such that $\beta(\theta) \in [-B,B]$ for all θ. Since $[-B,B]$ is a complete lattice and $\beta : [-B,B] \to [-B,B]$ is increasing, by Tarski's Fixed Point Theorem β has the greatest and least fixed points $\bar{\theta}^*$ and $\underline{\theta}^*$ on $[-B,B]$. Moreover, for all $\theta \in [-B,B]$, $\beta(\theta) \geq \theta$ implies $\theta \leq \bar{\theta}^*$ and $\beta(\theta) \leq \theta$ implies $\theta \geq \underline{\theta}^*$, and hence $\bar{\theta}^* = \sup\{\theta \in [-B,B] : \beta(\theta) \geq \theta\}$ and $\underline{\theta}^* = \inf\{\theta \in [-B,B] : \beta(\theta) \leq \theta\}$. Since $\beta(\theta) \leq B < \theta$ for all $\theta > B$, and $\beta(\theta) \geq -B > \theta$ for all $\theta < -B$, $\bar{\theta}^*$ and $\underline{\theta}^*$ are the greatest and least fixed points of β on $(-\infty,\infty)$ and $\bar{\theta}^* = \sup\{\theta : \beta(\theta) \geq \theta\}$ and $\underline{\theta}^* = \inf\{\theta : \beta(\theta) \leq \theta\}$.

Next, let $\theta^* \in \{\bar{\theta}^*, \underline{\theta}^*\}$. Consider any parameter p and let $\bar{S}(p) = \{\theta : \beta(\theta,p) \geq \theta\}$ and $\underline{S}(p) = \{\theta : \beta(\theta,p) \leq \theta\}$, where $\beta(\theta,p)$ denotes the best response function for a given value of the parameter. Then the greatest and smallest fixed points of $\beta(\theta,p)$ satisfy $\bar{\theta}^*(p) = \sup\bar{S}(p)$ and $\underline{\theta}^*(p) = \inf\underline{S}(p)$. We prove that if $\beta(\theta,p)$ increases with p, then $\bar{\theta}^*(p)$ and $\underline{\theta}^*(p)$ increase with p as well. This result is used for the proofs of statements (ii), (iii), (iv.a), and (iv.b). Consider any $p' \geq p$. Since $\beta(\theta,p)$ increases with p, $\beta\left(\bar{\theta}^*(p),p'\right) - \bar{\theta}^*(p) \geq \beta\left(\bar{\theta}^*(p),p\right) - \bar{\theta}^*(p) = 0$. Hence, $\bar{\theta}^*(p) \in \bar{S}(p')$, and since $\bar{\theta}^*(p') = \sup\bar{S}(p')$, we have $\bar{\theta}^*(p') \geq \bar{\theta}^*(p)$, as needed. The proof that $\underline{\theta}^*(p)$ increases with p is similar.

Last, for the proof of part (i), we show that if $\theta^* \in \{\bar{\theta}^*, \underline{\theta}^*\}$, then $\frac{\partial\beta(\theta)}{\partial\theta}|_{\theta=\theta^*} \leq 1$. Suppose that $\frac{\partial\beta(\theta)}{\partial\theta}|_{\theta=\bar{\theta}^*} > 1$. Then there exists $\theta_1 > \bar{\theta}^*$ such that $\beta(\theta_1) - \theta_1 > 0$. Since $\beta(\theta)$ is bounded, there exists $\theta_2 > \theta_1$ such that $\beta(\theta_2) - \theta_2 < 0$. Hence, by the intermediate value theorem, there exists $\hat{\theta} \in (\theta_1,\theta_2)$ such that $\beta\left(\hat{\theta}\right) - \hat{\theta} = 0$. But then $\hat{\theta}$ is a fixed point of $\beta(\cdot)$, which is greater than $\bar{\theta}^*$, contradicting the definition of $\bar{\theta}^*$ as the greatest fixed point of $\beta(\cdot)$. The proof that $\frac{\partial\beta(\theta)}{\partial\theta}|_{\theta=\underline{\theta}^*} \leq 1$ is similar. The case $\frac{\partial\beta(\theta)}{\partial\theta}|_{\theta=\theta^*} = 1$ is a knife-edge case, in which the best response function is tangent to the 45-degree line at the equilibrium point. Since we focus on local comparative statics, when the equilibrium continues to exist upon a small change in the parameter, this case will be ignored. Formally, it can be shown that the condition $\Delta\left(\tau^{-1}(0.5)\right) \neq 0$ ensures that $\frac{\partial \beta(\theta)}{\partial \theta}|_{\theta=\theta^*} \neq 1$ for any equilibrium θ^*.

Lemma IA.1 is used for the proof of Proposition 2 and Lemma 4.

LEMMA IA.1: *Consider the function* $g(p,K)=\sum_{t=\frac{K+1}{2}}^{K} C_t^K p^t (1-p)^{K-t}$ *and let* $K' > K$. *Then* $g(p,K') > g(p,K)$ *if and only if* $g(p,K) > 0.5$.

Proof of Lemma IA.1: The lemma follows from two statements: (1) $g(p,K) > 0.5$ if and only if $p > 0.5$, and (2) $g(p,K') > g(p,K)$ if and only if $p > 0.5$. We start by proving the first statement. Note that

$$1-g(p,K)=\sum_{t=0}^{\frac{K-1}{2}} C_t^K p^t (1-p)^{K-t} \overset{s=K-t}{=} \sum_{s=K}^{\frac{K+1}{2}} C_{K-s}^K p^{K-s} (1-p)^s = \sum_{s=\frac{K+1}{2}}^{K} C_s^K p^{K-s} (1-p)^s,$$

and hence

$$g(p,K) > 0.5 \Leftrightarrow g(p,K) > 1-g(p,K) \Leftrightarrow \sum_{t=\frac{K+1}{2}}^{K} C_t^K \left[p^t (1-p)^{K-t} - p^{K-t}(1-p)^t\right] > 0.$$

For any $t > \frac{K}{2}$, $p^t(1-p)^{K-t} > p^{K-t}(1-p)^t \Leftrightarrow p > 0.5$, and hence $g(p,K) > 0.5 \Leftrightarrow p > 0.5$.

Next, consider the second statement. We prove that $g(p,K+2) > g(p,K) \Leftrightarrow p > 0.5$. Using $C_t^{K+2} = C_t^{K+1} + C_{t-1}^{K+1}$,

$$\begin{aligned}
&g(p,K+2) = p^{K+2} + \textstyle\sum_{t=\frac{K+3}{2}}^{K+1} C_t^{K+1} p^t (1-p)^{K+2-t} + \sum_{t=\frac{K+3}{2}}^{K+1} C_{t-1}^{K+1} p^t (1-p)^{K+2-t} \\
&= p^{K+1} + \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^{K+1} p^t (1-p)^{K+2-t} + \sum_{t=\frac{K+3}{2}}^{K+1} C_{t-1}^{K+1} p^t (1-p)^{K+2-t} \\
&\overset{j=t-1}{=} p^{K+1} + \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^{K+1} p^t (1-p)^{K+2-t} + \sum_{j=\frac{K+1}{2}}^{K} C_j^{K+1} p^{j+1} (1-p)^{K+1-j} \\
&= p^{K+1} + \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^{K+1} p^t (1-p)^{K+2-t} + \sum_{t=\frac{K+3}{2}}^{K} C_t^{K+1} p^{t+1} (1-p)^{K+1-t} \\
&+ C_{\frac{K+1}{2}}^{K+1} p^{\frac{K+1}{2}+1} (1-p)^{\frac{K+1}{2}} \\
&= p^{K+1} + \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^{K+1} p^t (1-p)^{K+1-t} + p C_{\frac{K+1}{2}}^{K+1} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}}.
\end{aligned}$$

Using the fact that $C_t^{K+1} = C_t^K + C_{t-1}^K$, $g(p, K+2)$ can be further rewritten as

$$
\begin{aligned}
& p^{K+1} + \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^K p^t (1-p)^{K+1-t} + \sum_{t=\frac{K+3}{2}}^{K} C_{t-1}^K p^t (1-p)^{K+1-t} + pC_{\frac{K+1}{2}}^{K+1} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}} \\
& \overset{j=t-1}{=} p^{K+1} + \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^K p^t (1-p)^{K+1-t} + \sum_{j=\frac{K+1}{2}}^{K-1} C_j^K p^{j+1} (1-p)^{K-j} \\
& + pC_{\frac{K+1}{2}}^{K+1} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}} \\
& = \textstyle\sum_{t=\frac{K+3}{2}}^{K} C_t^K p^t (1-p)^{K+1-t} + \sum_{t=\frac{K+1}{2}}^{K} C_t^K p^{t+1} (1-p)^{K-t} + pC_{\frac{K+1}{2}}^{K+1} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}} \\
& = \textstyle\sum_{t=\frac{K+1}{2}}^{K} C_t^K \left[p^t (1-p)^{K+1-t} + p^{t+1} (1-p)^{K-t} \right] \\
& + pC_{\frac{K+1}{2}}^{K+1} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}} - C_{\frac{K+1}{2}}^{K} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}} \\
& = \textstyle\sum_{t=\frac{K+1}{2}}^{K} C_t^K p^t (1-p)^{K-t} + \left[pC_{\frac{K+1}{2}}^{K+1} - C_{\frac{K+1}{2}}^{K} \right] p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}} .
\end{aligned}
$$

The first term equals $g(p, K)$, and the second term equals $(2p-1) C_{\frac{K+1}{2}}^{K} p^{\frac{K+1}{2}} (1-p)^{\frac{K+1}{2}}$. Hence, $g(p, K+2) > g(p, K) \Leftrightarrow p > 0.5$.

Supplemental material for the proof of Lemma 4: The next argument is used in the proof of part (iii) of the second statement of the lemma. Recall from the proof of Proposition 1 that $\bar{\alpha} = \inf\{A\}$, where $A = \{\hat{\alpha} \geq 0 :$ for any $\alpha \geq \hat{\alpha}$, there exists at least one shareholder-friendly and at least one management-friendly equilibrium$\}$. We prove that if $\Delta(\mu) \neq 0$, then $\bar{\alpha} \in A$. Define $\Psi(\theta, \alpha) \equiv \beta(\theta, \alpha) - \theta$. Suppose that $\bar{\alpha} \notin A$. The assumption $\Delta(\mu) \neq 0$ ensures that $\theta^* = \mu$ can never be an equilibrium. Indeed, if it were an equilibrium, (7) would imply $\mu = \beta(\mu) = \Delta^{-1}(0)$, which would contradict $\Delta(\mu) \neq 0$. Hence, if $\bar{\alpha} \notin A$, either all the roots of $\Psi(\theta, \bar{\alpha})$ are in $(\mu, +\infty)$, or all the roots of $\Psi(\theta, \bar{\alpha})$ are in $(-\infty, \mu)$. Consider the first case; the proof for the second case is similar. Consider $\bar{\alpha} + \delta$ for any $\delta > 0$. Since $\lim_{\theta \to -\infty} \Psi(\theta, \bar{\alpha} + \delta) = +\infty$, there exists $\underline{\theta} < \mu$ such that $\Psi(\theta, \bar{\alpha} + \delta) > 0$ for any $\theta \leq \underline{\theta}$. By (7), $\Psi(\theta, \alpha)$ decreases in α when $\theta < \mu$, and hence $\Psi(\theta, \alpha) > 0$ for any $\theta \leq \underline{\theta}$ and any $\alpha \leq \bar{\alpha} + \delta$. By the extreme value theorem, the function $\Psi(\theta, \alpha)$ attains its minimum on $[\underline{\theta}, \mu]$. Now consider the function $\Psi^*(\alpha) = \min_{\theta \in [\underline{\theta}, \mu]} \Psi(\theta, \alpha)$. Since $\Psi(\theta, \alpha)$ is continuous in both arguments and $[\underline{\theta}, \mu]$ is a compact set, then by Berge's Maximum theorem, $\Psi^*(\alpha)$ is continuous in α. Note also that $\Psi^*(\bar{\alpha}) > 0$. Indeed, since all the roots of $\Psi(\theta, \bar{\alpha})$ are in $(\mu, +\infty)$ and since $\lim_{\theta \to -\infty} \Psi(\theta, \bar{\alpha}) = +\infty$, $\Psi(\theta, \bar{\alpha}) > 0$ for any $\theta \leq \mu$ as otherwise, by the intermediate value theorem, there would be a root on $(-\infty, \mu]$. Hence, $\Psi(\theta, \bar{\alpha}) > 0$ for $\theta = \arg\min_{\theta \in [\underline{\theta}, \mu]} \Psi(\theta, \alpha)$, that is, $\Psi^*(\bar{\alpha}) > 0$. Since $\Psi^*(\alpha)$ is continuous and $\Psi^*(\bar{\alpha}) > 0$, there exists $\delta' \in (0, \delta)$ such that $\Psi^*(\bar{\alpha} + \varepsilon) > 0$ for any $\varepsilon < \delta'$. Hence, for any $\varepsilon < \delta'$, $\Psi(\theta, \bar{\alpha} + \varepsilon) > 0$ for any $\theta \in [\underline{\theta}, \mu]$. In addition, as shown above, $\Psi(\theta, \bar{\alpha} + \varepsilon) > 0$ for any $\theta \leq \underline{\theta}$ and any $\varepsilon < \delta'$. Thus, for any $\varepsilon < \delta'$,

$\Psi(\theta, \bar{\alpha}+\varepsilon) > 0$ for any $\theta \leq \mu$, that is, there is no shareholder-friendly equilibrium for $\bar{\alpha}+\varepsilon$. Thus, $\bar{\alpha}+\varepsilon \notin A$ for any $\varepsilon \in [0, \delta')$, which contradicts the fact that $\bar{\alpha} = \inf\{A\}$. Hence, $\bar{\alpha} \in A$.

Next, consider the proof of part (ii). Suppose $\alpha \in (\alpha_1, \alpha_2)$. According to the proof of Lemma IA.1, since $\tau(\theta) = g(1 - F(\theta), K)$, $\tau(\theta)$ increases with K if and only if $\theta < \mu$. In addition,

$$\lim_{K \to \infty} \tau(\theta, K) = \begin{cases} 0 & \text{if } \theta > \mu \\ 1 & \text{if } \theta < \mu. \end{cases}$$

Hence, $\beta(\theta, K)$ increases with K if and only if $\theta > \mu$ and

$$\lim_{K \to \infty} \beta(\theta, K) = \begin{cases} \Delta^{-1}(\alpha\lambda) & \text{if } \theta > \mu \\ \Delta^{-1}(-\alpha\lambda) & \text{if } \theta < \mu \end{cases} \tag{IA.1}$$

Since $\alpha > \alpha_1$, $\mu \in (\Delta^{-1}(-\alpha\lambda), \Delta^{-1}(\alpha\lambda))$. Equation (IA.1) therefore implies that there exist $\varepsilon > 0$ and $\hat{K} > 3$ such that if $K \geq \hat{K}$, then $\beta(\mu+\varepsilon, K) > \mu+\varepsilon$ and $\beta(\mu-\varepsilon, K) < \mu-\varepsilon$. Thus, the function $\Psi(\theta, K) \equiv \beta(\theta, K) - \theta$ satisfies $\Psi(\mu+\varepsilon, K) > 0 > \Psi(\mu-\varepsilon, K)$ for any $K \geq \hat{K}$. Since $\lim_{\theta\to\infty} \Psi(\theta, K) = -\infty$ and $\lim_{\theta\to-\infty} \Psi(\theta, K) = \infty$, then by the intermediate value theorem, for any $K \geq \hat{K}$, there exist $\theta_1^* < \mu - \varepsilon$ and $\theta_2^* > \mu+\varepsilon$ such that $\Psi(\theta_1^*, K) = \Psi(\theta_2^*, K) = 0$. Hence, θ_1^* (θ_2^*) is a shareholder-friendly (management-friendly) equilibrium, that is, both types of equilibria exist for any $K \geq \hat{K}$. Since $\alpha < \alpha_2 = \bar{\alpha}(3)$, Proposition 1 implies that only one type of equilibrium exists for $K = 3$. The first statement of the lemma then implies that there exists $\hat{K} > 3$ such that both types of equilibria exist if and only if $K \geq \hat{K}$, which completes the proof of part (ii).

II. Directors' Utility

LEMMA IA.2: *(i) The aggregate expected utility of directors is given by*

$$W_{Directors}(\theta^*) = 2\delta\alpha + 2K \times \begin{pmatrix} \tau(\theta^*)\,\mathbb{E}[v(1,\theta)] + (1-\tau(\theta^*))\,\mathbb{E}[v(0,\theta)] \\ -C_{T-1}^{K-1}(1-F(\theta^*))^{T-1} F(\theta^*)^{K-T} \times \int_{-\infty}^{\theta^*} \Delta(\theta) f(\theta)\, d\theta \end{pmatrix}. \tag{IA.2}$$

(ii) Suppose $T = \frac{K+1}{2}$ and $\Delta(\theta_{median}) = 0$, where θ_{median} is the median of the distribution F. Then $\frac{\partial W_{Directors}(\theta^*)}{\partial \theta^*} > 0 \Leftrightarrow \tau(\theta^*) < \frac{1}{2}$.

The aggregate expected utility of all incumbent directors and outside candidates consists of two components – directors' utility from the labor market and incumbent directors' utility from the allocation of control in their firms.

First, directors get utility from being appointed to the board of a firm that was hit by a resignation shock. Recall that each director, whether he is an incumbent director or an outside candidate, gets utility α if he is invited to the board at the second stage. Therefore, the aggregate utility from the labor market is zero if no director resigns, α if only one director resigns, and 2α if two directors resign, regardless of whether an incumbent director or an outside candidate fills the vacancy. Overall, directors' aggregate expected utility from the labor market does not depend on θ^* and is given by $\delta^2 2\alpha + 2\delta(1-\delta)\alpha = 2\delta\alpha$.

The second term in (IA.2) is the expected utility of each incumbent director from the allocation of control in his firm, multiplied by the number of incumbent directors, $2K$. Intuitively, the first two terms in parentheses represent the director's expected value from the allocation of control as if the director's own type did not matter for the outcome. Because of the assumption $\mathbb{E}[\Delta(\theta)] = 0$, the sum of these two terms is $\mathbb{E}[v(1,\theta)]$, which is independent of θ^*.[1] The last term in parentheses, which is always positive ($\int_{-\infty}^{\theta^*} \Delta(\theta) f(\theta) d\theta < 0$ since $\mathbb{E}[\Delta(\theta)] = 0$ and $\Delta(\theta)$ is increasing), reflects the correlation between the allocation of control in the firm and the type of the director. In particular, in situations in which the director is pivotal for the outcome (which happens with probability $C_{T-1}^{K-1}(1-F(\theta^*))^{T-1} F(\theta^*)^{K-T}$), control of the firm is given to shareholders if and only if the director's type satisfies $\theta \geq \theta^*$. Since $v(1,\theta)$ increases in θ and $v(0,\theta)$ decreases in θ, this creates additional expected value to the director. Note that this term can increase or decrease in θ^* depending on the probability of a director being pivotal as well as on whether $\Delta(\theta^*) > 0$ (the integral $\int_{-\infty}^{\theta^*} \Delta(\theta) f(\theta) d\theta$ increases in θ^* if and only if $\Delta(\theta^*) > 0$).[2] Thus, the overall effect of θ^* is generally ambiguous. The second part of Lemma IA.2 shows that under certain conditions, the welfare of directors is increasing in θ^* if and only if the equilibrium is management-friendly.

Proof of Lemma IA.2: We start by proving part (i). Consider director k in firm i. If $\theta_{ik} \geq \theta^*$,

[1] If $\mathbb{E}[\Delta(\theta)] \neq 0$, there is an additional effect of θ^* on directors' welfare. In particular, θ^* negatively (positively) affects directors' welfare if $\mathbb{E}[\Delta(\theta)] > 0$ ($\mathbb{E}[\Delta(\theta)] < 0$).

[2] To illustrate the intuition behind this point, suppose that the director is pivotal with probability one, for example, if he is the only director. Consider an increase in θ^*, that is, an increase in the management-friendliness of the equilibrium. If $\Delta(\theta^*) < 0$, the marginal type, θ^*, benefits from allocating control to management. Therefore, the director's expected utility increases when the equilibrium becomes more management-friendly. Similarly, if $\Delta(\theta^*) > 0$, the marginal type benefits from allocating control to shareholders, and hence the director's expected utility decreases with θ^*.

the director expects to vote for the proposal. The proposal is passed if and only if at least $T-1$ of the other $K-1$ directors in firm i support the proposal. Similarly, if $\theta_{ik} < \theta^*$, the director expects to vote against the proposal. The proposal is passed if and only if at least T of the other $K-1$ directors support the proposal as well. Thus, the director's expected value from the allocation of control is given by

$$U_D = \Pr\left[\theta_{ik} \geq \theta^*\right] \begin{pmatrix} \mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] \sum_{t=T-1}^{K-1} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \\ +\mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] \sum_{t=0}^{T-2} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \end{pmatrix}$$
$$+\Pr\left[\theta_{ik} < \theta^*\right] \begin{pmatrix} \mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \sum_{t=T}^{K-1} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \\ +\mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \sum_{t=0}^{T-1} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \end{pmatrix}.$$

Rearranging the terms, this utility can be rewritten as

$$\begin{aligned} U_D &= \begin{pmatrix} \Pr\left[\theta_{ik} \geq \theta^*\right]\mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] + \\ \Pr\left[\theta_{ik} < \theta^*\right]\mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \end{pmatrix} \sum_{t=T}^{K-1} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \\ &\quad + \begin{pmatrix} \Pr\left[\theta_{ik} \geq \theta^*\right]\mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] \\ +\Pr\left[\theta_{ik} < \theta^*\right]\mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \end{pmatrix} \sum_{t=0}^{T-2} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \\ &\quad + \begin{pmatrix} \Pr\left[\theta_{ik} \geq \theta^*\right]\mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] \\ +\Pr\left[\theta_{ik} < \theta^*\right]\mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \end{pmatrix} C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-1} F\left(\theta^*\right)^{K-1-(T-1)} \\ &= \mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] \sum_{t=T}^{K-1} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \\ &\quad +\mathbb{E}\left[v\left(0,\theta_{ik}\right)\right] \sum_{t=0}^{T-2} C_t^{K-1}\left(1-F\left(\theta^*\right)\right)^t F\left(\theta^*\right)^{K-1-t} \\ &\quad + \begin{pmatrix} \Pr\left[\theta_{ik} \geq \theta^*\right]\mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] \\ +\Pr\left[\theta_{ik} < \theta^*\right]\mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \end{pmatrix} C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-1} F\left(\theta^*\right)^{K-1-(T-1)}. \end{aligned}$$

The assumption $\mathbb{E}\left[\Delta\left(\theta_{ik}\right)\right] = 0$ implies $\mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] = \mathbb{E}\left[v\left(0,\theta_{ik}\right)\right]$, and hence the director's utility U_D satisfies

$$
\begin{aligned}
& U_D - \mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] = \\
& \begin{pmatrix} \Pr\left[\theta_{ik} \geq \theta^*\right] \mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} \geq \theta^*\right] \\ +\Pr\left[\theta_{ik} < \theta^*\right] \mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \\ -\mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] \end{pmatrix} C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-1} F\left(\theta^*\right)^{K-1-(T-1)} \\
& = \Pr\left[\theta_{ik} < \theta^*\right] \begin{pmatrix} \mathbb{E}\left[v\left(0,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \\ -\mathbb{E}\left[v\left(1,\theta_{ik}\right)|\theta_{ik} < \theta^*\right] \end{pmatrix} C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-1} F\left(\theta^*\right)^{K-1-(T-1)} \\
& = -\Pr\left[\theta_{ik} < \theta^*\right] \mathbb{E}\left[\Delta\left(\theta_{ik}\right)|\theta_{ik} < \theta^*\right] C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-1} F\left(\theta^*\right)^{K-1-(T-1)} \\
& \Leftrightarrow U_D = \mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] - C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-1} F\left(\theta^*\right)^{K-T} \times \int_{-\infty}^{\theta^*} \Delta\left(\theta_{ik}\right) f\left(\theta_{ik}\right) d\theta_{ik}.
\end{aligned}
$$

Finally, since $\mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] = \mathbb{E}\left[v\left(0,\theta_{ik}\right)\right]$, it follows that $\mathbb{E}\left[v\left(1,\theta_{ik}\right)\right]$ can be rewritten as $\tau\left(\theta^*\right)\mathbb{E}\left[v\left(1,\theta_{ik}\right)\right] + \left(1-\tau\left(\theta^*\right)\right)\mathbb{E}\left[v\left(0,\theta_{ik}\right)\right]$, which gives (IA.2). Noting that there are $2K$ incumbent directors and combining this expression with directors' aggregate utility from the labor market, $2\delta\alpha$, concludes the proof.

We next prove part *(ii)*. Using (IA.2),

$$
\begin{aligned}
\frac{\partial W_{Directors}}{\partial \theta^*} &= -2Kf\left(\theta^*\right) C_{T-1}^{K-1}\left(1-F\left(\theta^*\right)\right)^{T-2} F\left(\theta^*\right)^{K-T-1} \\
& \times \begin{pmatrix} \left[-\left(T-1\right)F\left(\theta^*\right) + \left(K-T\right)\left(1-F\left(\theta^*\right)\right)\right] \times \int_{-\infty}^{\theta^*} \Delta\left(\theta_{ik}\right) f\left(\theta_{ik}\right) d\theta_{ik} \\ +\left(1-F\left(\theta^*\right)\right) F\left(\theta^*\right) \times \Delta\left(\theta^*\right) \end{pmatrix}.
\end{aligned}
$$

Suppose $T = \frac{K+1}{2}$. Then

$$
\frac{\partial W_{Directors}}{\partial \theta^*} > 0 \Leftrightarrow \left(\frac{K-1}{2}\right)\left(\frac{1-2F\left(\theta^*\right)}{1-F\left(\theta^*\right)}\right) \times \mathbb{E}\left[\Delta\left(\theta_{ik}\right)|\theta_{ik} < \theta^*\right] + \Delta\left(\theta^*\right) < 0.
$$

Since $\mathbb{E}\left[\Delta\left(\theta_{ik}\right)\right] = 0$ and $\Delta\left(\cdot\right)$ is increasing, we have $\mathbb{E}\left[\Delta\left(\theta_{ik}\right)|\theta_{ik} < \theta^*\right] < 0$. Also, note that $1-2F\left(\theta^*\right) < 0$ if and only if $\theta^* > \theta_{median}$. Hence, if $\Delta\left(\theta_{median}\right) = 0$, then

$$
\frac{\partial W_{Directors}}{\partial \theta^*} > 0 \Leftrightarrow \theta^* > \theta_{median}.
$$

Finally, note that when $T = \frac{K+1}{2}$, $\tau\left(\theta^*\right) > \frac{1}{2}$ if and only if $F\left(\theta^*\right) < \frac{1}{2}$. Hence, $\frac{\partial W_{Directors}}{\partial \theta^*} > 0 \Leftrightarrow \tau\left(\theta^*\right) < \frac{1}{2}$.

III. Extensions of the Basic Model

A. Value of Shareholder Control

If managers have high expertise or if they need to be given incentives to make firm-specific investments, shareholders may be better off delegating control to management and having some management-friendly directors on the board (e.g., Grossman and Hart (1986), Hart and Moore (1990), Adams and Ferreira (2007), Harris and Raviv (2008)). The relative value of management control can differ across firms. To capture this cross-sectional heterogeneity, we extend the model and assume that in each firm, shareholder control is optimal only with some probability. In particular, there is a random variable $\zeta_i \in \{SH, M\}$ such that shareholders of firm i are better off having control if $\zeta_i = SH$ and are better off delegating control to management if $\zeta_i = M$. Formally, we assume that shareholders' utility function is still $u_{SH}(\chi_i, \theta_{i1}, ..., \theta_{iK})$ when $\zeta_i = SH$, but coincides with management's utility function, $u_M(\chi_i, \theta_{i1}, ..., \theta_{iK})$ when $\zeta_i = M$, where u_{SH} and u_M are given by (2). In particular, this implies that if $\zeta_i = M$, then regardless of the allocation of control, the firm will demand management-friendly directors.

We assume that ζ_i are independent across firms and independent of directors' types, and that ζ_i is the private information of directors of firm i and the party that controls firm i at the second stage. The prior probability that $\zeta_i = SH$ is equal to $\psi \in (0, 1]$. Directors' relative utility from shareholder control, $\Delta_{\zeta_i}(\theta) \equiv v_{\zeta_i}(1, \theta) - v_{\zeta_i}(0, \theta)$, depends on ζ_i. Specifically, we assume that $\Delta_{SH}(\theta) \geq \Delta_M(\theta)$, that $\Delta_{SH}(\theta)$ has the same properties as $\Delta(\theta)$ in the basic model (that is, $\frac{\partial \Delta_{SH}(\theta)}{\partial \theta} > 0$, $\lim_{\theta \to \infty} \Delta_{SH}(\theta) = \infty$, and $\lim_{\theta \to -\infty} \Delta_{SH}(\theta) = -\infty$), and that $\frac{\partial \Delta_M(\theta)}{\partial \theta} > 0$, $\lim_{\theta \to \infty} \Delta_M(\theta) = \Delta_0 \leq 0$, and $\lim_{\theta \to -\infty} \Delta_M(\theta) = -\infty$.

Similar to the basic model, directors follow threshold strategies, which are now conditional on the realized ζ_i. In particular, any symmetric equilibrium is characterized by two thresholds $(\theta^*_{SH}, \theta^*_M)$ such that director k of firm i with signal ζ_i and type θ_{ik} votes for shareholder control if and only if $\theta_{ik} > \theta^*_{\zeta_i}$. The proof of Proposition IA.1 shows that the extended model exhibits strategic complementarity as well.

The ex-ante probability that firm j demands shareholder-friendly directors is the probability that $\zeta_j = SH$ and $\chi_j = 1$, which equals $\psi\tau(\theta^*_{SH})$, where $\tau(\cdot)$ is given by (5). We call an equilibrium shareholder-friendly if a shareholder-friendly reputation is more valuable in the labor market, that is, if $\psi\tau(\theta^*_{SH}) > 0.5$. The next proposition characterizes the equilibria.

PROPOSITION IA.1: *A symmetric equilibrium always exists. In any symmetric equilibrium, θ^*_{SH} is finite. If $\alpha = 0$ or if the equilibrium is management-friendly, θ^*_M is infinite.*[3]

[3]If $\Delta_0 = 0$, these conditions are also necessary for θ^*_M to be infinite.

The proposition shows that equilibria feature two potential types of inefficiency. The first type of inefficiency is similar to the basic model and arises when directors do not transfer control to shareholders even when shareholder control is optimal. This inefficiency is always present because θ^*_{SH} is always finite. In addition, if θ^*_M is finite, the equilibrium features another type of inefficiency, where directors allocate control to shareholders even though management control is optimal. The only reason directors do this is to signal their shareholder-friendliness to the other firm, which is beneficial if a shareholder-friendly reputation is more valuable. Hence, the second type of inefficiency does not arise (θ^*_M is infinite) when directors have no reputational concerns ($\alpha = 0$) or when the labor market rewards a management-friendly reputation ($\psi\tau\left(\theta^*_{SH}\right) < 0.5$).

Proof of Proposition IA.1: Conditional on being pivotal, a director's relative utility from voting for shareholder control relative to voting against it is

$$v_{\zeta_i}\left(1, \theta_{ik}\right) - v_{\zeta_i}\left(0, \theta_{ik}\right) + \frac{\alpha\delta}{K}\left[\psi\tau\left(\theta^*_{SH}\right) - \left(1 - \psi\tau\left(\theta^*_{SH}\right)\right)\right]. \qquad \text{(IA.3)}$$

It is immediate that the extended model exhibits strategic complementarity as well. The equilibrium θ^*_{SH} is determined by the condition

$$\Delta_{SH}\left(\theta^*_{SH}\right) = \frac{\alpha\delta}{K}\left(1 - 2\psi\tau\left(\theta^*_{SH}\right)\right).$$

Since the right-hand side is bounded and the left-hand side takes all values in $(-\infty, \infty)$, this equation has at least one solution $\theta^*_{SH} < \infty$. In addition, (IA.3) implies that for a given θ^*_{SH}, the threshold θ^*_M is unique and is given by $\theta^*_M = \infty$ if $\Delta_{SH}\left(\theta^*_{SH}\right) \geq \Delta_0$ and by $\theta^*_M = \Delta_M^{-1}\left(\Delta_{SH}\left(\theta^*_{SH}\right)\right)$ otherwise. Hence, if $\alpha = 0$ or if $\psi\tau\left(\theta^*_{SH}\right) \leq 0.5$, then $\Delta_{SH}\left(\theta^*_{SH}\right) \geq 0 \geq \Delta_0$, and thus $\theta^*_M = \infty$. Note also that since $\psi > 0$, there is no need to specify off-equilibrium events. These arguments imply that a symmetric equilibrium always exists.

B. Board Independence

In this extension, we study how the number of insiders on the board affects corporate governance. Suppose ι out of K directors in each firm are insiders, where $\iota \in \{0, ..., K-1\}$. Whether a director is an insider is common knowledge. If a director is an insider, he always votes against shareholder control (effectively, his type is $-\infty$). The remaining $K - \iota$ directors are independent, and their type is their private information, as in the basic model. Because insiders are the most management-friendly among all directors, a firm that is controlled by management will always prefer to hire an insider over both incumbent independent directors and outside candidates, as long as at least one insider has the capacity to serve on another board. Because companies often restrict their executives from serving on too many board seats, we assume that after the resignation shocks are realized, there is probability φ that at least one of the remaining insiders participates in the labor market for directors. Hence, if some insiders remain on the board of firm i and firm j is controlled by management, firm j will hire one of the insiders of firm i with probability φ and will choose between independent directors of firm i and outside candidates with probability $1-\varphi$.

If $\iota > K-T$, insiders have enough power to block any attempt to transfer control to shareholders. Therefore, the $K-\iota$ independent directors are never pivotal for the outcome, and management retains control with probability one. Suppose $\iota \leq K-T$. Similar to the basic model, it can be shown that every equilibrium is a symmetric threshold equilibrium. The expressions for $\pi_i(\chi_i;\theta^*)$ and $\tau(\theta^*)$ are given by (A3) and (5), respectively, with the exception that K is replaced by $K-\iota$. The following lemma characterizes the best response strategies.

LEMMA IA.3: *The best response of the director is given by*

$$\beta_\iota(\theta) = \Delta^{-1}\left(\alpha\frac{\delta}{K-\iota}\left[\left(1-\tau(\theta,\iota)\right)\left(1-\hat{\varphi}(\iota)\right)-\tau(\theta,\iota)\right]\right), \tag{IA.4}$$

where

$$\tau(\theta,\iota) = \sum_{t=T}^{K-\iota} C_t^{K-\iota}\left(1-F(\theta)\right)^t F(\theta)^{K-\iota-t}$$

and

$$\hat{\varphi}(\iota) = \begin{cases} 0 & \text{if } \iota = 0 \\ \varphi - \varphi\frac{\delta}{K} & \text{if } \iota = 1 \\ \varphi & \text{if } \iota > 1. \end{cases}$$

Note that if $\varphi = 0$, that is, insiders do not participate in the labor market for directors, then the model is similar to the basic model, with the exception that K is replaced by $K-\iota$ (while T is unchanged). Thus, an increase in the number of insiders has a similar effect as a decrease in K while keeping T constant. Hence, based on Proposition 2, part (iv.a), if $\frac{\delta}{K-\iota}$ is fixed, then an increase in the number of insiders decreases $\tau(\theta^*,\iota)$, that is, shareholders are less likely to obtain control. However, if δ is fixed, then there is an additional supply effect:

the more insiders there are, the lower is the competition between the independent directors for a position in the other firm. This effect is similar to an increase in α and amplifies corporate governance. The two effects act in the same (opposite) direction in a management-friendly (shareholder-friendly) equilibrium. Hence, a management-friendly equilibrium becomes more management-friendly as ι increases, while the effect of ι in a shareholder-friendly equilibrium is ambiguous.

Interestingly, when $\varphi = 1$, that is, insiders always participate in the labor market for directors, independent directors can be more likely to vote for shareholders in the presence of insiders. There are two reasons. First, directors do not have incentives to create a management-friendly reputation because if the other firm is controlled by management, it will always prefer to invite insiders rather than independent directors. Second, if the other firm is controlled by shareholders, then each independent director faces competition from only $K - \iota - 1$ fellow independent directors, which increases his incentives to create a shareholder-friendly reputation. Both effects induce directors to vote for shareholder control. This positive effect of insiders is counteracted by the more intuitive negative effect: shareholders obtain control only if T out of $K - \iota$ directors support the proposal, and the likelihood of that decreases with ι. Thus, generally, the effect of the number of insiders on the likelihood of shareholder control is ambiguous. Following the proof of Lemma IA.3, we present an example that shows that the first effect can dominate, and the presence of insiders may result in a higher probability of shareholder control. Thus, even when stronger governance increases shareholder welfare, regulators should be cautious about imposing restrictions on the proportion of independent directors on the board.

Proof of Lemma IA.3: When shareholders obtain control in both firms, an independent director of firm i has probability

$$\delta\left[(1-\delta)\frac{1}{K-\iota}+\delta\frac{\iota}{K}\frac{1}{K-\iota}+\delta\frac{K-\iota-1}{K}\frac{1}{K-\iota-1}\right]=\frac{\delta}{K-\iota}$$

of being hired by the other firm. When management obtains control in both firms, an independent director of firm i has probability

$$\delta(1-\varphi)\left[(1-\delta)\frac{1}{K-\iota}+\delta\frac{\iota}{K}\frac{1}{K-\iota}+\delta\frac{K-\iota-1}{K}\frac{1}{K-\iota-1}\right]=\frac{\delta}{K-\iota}(1-\varphi)$$

of being hired by the other firm if $\iota > 1$,

$$\delta\left[(1-\delta)\frac{1-\varphi}{K-1}+\delta\frac{1}{K}\frac{1}{K-1}+\delta\frac{K-2}{K}\frac{1-\varphi}{K-2}\right]=\frac{\delta}{K-1}\left(1-\varphi+\varphi\frac{\delta}{K}\right)$$

if $\iota = 1$, and $\frac{\delta}{K}$ if $\iota = 0$. If the allocation of control in the two firms is different, the director is never hired. Thus, similar to the basic model, the best response of the director is given by (IA.4).

Example. In this example, we show that the presence of insiders could lead to a higher probability of shareholder control. Suppose $T = 2$, $K = 3$, and $\varphi = 1$. Consider the most shareholder-friendly equilibrium, which is defined in Section II.A of the published article as the greatest fixed point of the best response function. First, we show that if δ is sufficiently small, then the most shareholder-friendly equilibrium for $\iota = 1$ is more shareholder-friendly than for $\iota = 0$. To prove this, it is sufficient to show that $\beta_1(\theta) < \beta(\theta)$ for all θ, that is, $\frac{1}{K-1}\left[(1-\tau(\theta,1))\frac{\delta}{K}-\tau(\theta,1)\right] < \frac{1}{K}\left[1-2\tau(\theta,0)\right]$. According to the formulas above, this holds if and only if

$$\begin{aligned}
\frac{3}{2}\left[\begin{array}{c}\left(1-\sum_{t=2}^{2}C_t^2\left(1-F(\theta)\right)^t F(\theta)^{2-t}\right)\frac{\delta}{3}\\ -\sum_{t=2}^{2}C_t^2\left(1-F(\theta)\right)^t F(\theta)^{2-t}\end{array}\right] &< 1-2\sum_{t=2}^{3}C_t^3\left(1-F(\theta)\right)^t F(\theta)^{3-t} \Leftrightarrow\\
\delta-(3+\delta)\left(1-F(\theta)\right)^2 &< 2-4\left[\left(1-F(\theta)\right)^3+3F(\theta)\left(1-F(\theta)\right)^2\right] \Leftrightarrow\\
\left(1-\delta+8F(\theta)\right)\left(1-F(\theta)\right)^2 &< 2-\delta.
\end{aligned}$$

Consider the function $g(x) = (1-x)^2(1-\delta+8x)$. Since $g'(x) = 2(1-x)(3+\delta-12x)$, we have that $(1-\delta+8F(\theta))(1-F(\theta))^2$ has a maximum when $F(\theta) = \frac{3+\delta}{12}$, where it achieves the maximum value of $\frac{(9-\delta)^3}{3^3\times 4^2}$. Note that $\frac{(9-\delta)^3}{3^3\times 4^2} < 2-\delta$ for δ sufficiently close to zero. Thus, in this case, the most shareholder-friendly equilibrium for $\iota = 1$ is more shareholder-friendly than for $\iota = 0$.

Next, we show that a higher propensity of independent directors to vote for shareholder control can also translate into a higher probability of shareholder control, even though insiders never vote for shareholder control. Consider the above parameters, that is, $T = 2$, $K = 3$, and $\varphi = 1$, and suppose that $\delta = 0.01$, $\alpha = 100$, $v(1,\theta) = e^{0.2\theta}$, $v(0,\theta) = e^{-0.2\theta}$, and the distribution of types is normal with mean μ and variance one. Solving for the equilibrium numerically, we show that for a large range of μ (particularly, in the negative part of the real line) the equilibrium is unique for both $\iota = 1$ and $\iota = 0$, and that τ^* for $\iota = 1$ is strictly higher than for $\iota = 0$. For example, if $\mu = -0.1$, then $\theta^*_{\iota=1} = -0.58$ and $\tau^*_{\iota=1} = 0.47$, while $\theta^*_{\iota=0} = 0.57$ and $\tau^*_{\iota=0} = 0.16$.

C. Boardroom Transparency

While the board's decision-making process is generally opaque, recent regulations have increased boardroom transparency. For example, as we discuss in Section III of the published article, the 2004 SEC law requires firms to disclose any director departure that is due to a disagreement, and the 2004 law in China requires firms to disclose the names of directors who vote in dissent. Our setting allows us to study whether increased boardroom transparency is beneficial for corporate governance.

To study the effect of transparency, we consider a variation of the basic model in which not only the allocation of control but also the individual votes of all directors are observable. When making their decisions, directors know that their votes will be observable. All other assumptions remain unchanged.

Recall from the proof of Lemma 1 that when individual directors' voting decisions are not observable, a pooling equilibrium that survives small perturbations in directors' strategies never exists. However, as the next result demonstrates, when individual votes are observed, both a shareholder-friendly and a management-friendly pooling equilibrium always exist and are robust to small perturbations.

PROPOSITION IA.2: *Suppose $T \notin \{1, K\}$ and $\alpha > 0$. Then there exists a pooling equilibrium in which all directors vote for shareholder control and a pooling equilibrium in which all directors vote against shareholder control, regardless of their types.*

In particular, Proposition IA.2 implies that with transparency, the most shareholder-friendly equilibrium becomes more shareholder-friendly ($\underline{\theta}^*$ becomes infinitely small) and the least shareholder-friendly equilibrium becomes less shareholder-friendly ($\bar{\theta}^*$ becomes infinitely large). In this sense, boardroom transparency amplifies corporate governance: strong governance systems become stronger and weak governance systems become weaker.[4] Hence, trans-

[4]More precisely, the amplification effect of transparency holds when $\alpha > \bar{\alpha}$, that is, when both types of equilibria coexist without transparency. When $\alpha < \bar{\alpha}$, and hence only one type of equilibrium exists (for example, shareholder-friendly), it is still true that due to transparency the most (least) shareholder-friendly equilibrium becomes more (less) shareholder-friendly. However, since in this case the least shareholder-friendly equilibrium is shareholder-friendly, transparency weakens governance in the least shareholder-friendly equilibrium.

parency can be harmful by increasing the level of management control. Note that this result is similar in spirit to Proposition 2, part (i), which shows that directors' reputational concerns improve governance only if the equilibrium is shareholder-friendly. This similarity is not a coincidence: as we explain below, transparency effectively magnifies directors' reputational concerns.

The connection between boardroom transparency and directors' reputational concerns can be best explained by understanding why pooling equilibria exist when individual votes are observable, but do not exist when votes are unobservable. Consider, for example, the pooling equilibrium in which directors vote for shareholder control regardless of their types. In this equilibrium a director is never pivotal for the outcome of the vote, and hence his decisions are driven solely by his reputational concerns. When individual votes are unobservable, his reputation is unaffected by his vote, and hence the director is completely indifferent between voting for and against shareholder control. However, with any – even arbitrarily small – positive probability of being pivotal (upon a perturbation of the game), the director will put a strictly positive weight on his intrinsic preferences, represented by $v(\chi_i, \theta_i)$. If the director is sufficiently management-friendly, he will then have incentives to deviate from his equilibrium strategy and vote for management. Hence, without transparency, this pooling equilibrium does not survive small perturbations of the game. By contrast, when the director's vote is observable, the director has strict incentives to vote for shareholder control even though he is not pivotal for the outcome of the vote. The reason is that the other firm is expected to be controlled by shareholders, and hence voting for shareholder control, due to transparency, strictly increases the likelihood of being invited to the other board relative to voting for management control. It follows that with transparency, this pooling equilibrium survives small perturbations of the game: even if a director is pivotal with a positive but sufficiently small probability, he still strictly prefers to vote for shareholder control in order to maintain his reputational gains. In this sense, boardroom transparency magnifies directors' reputational concerns.

In addition to pooling equilibria, partial pooling equilibria may exist, as in the basic model. In a partial pooling equilibrium, there exists a finite threshold θ^* such that a director of type θ votes for shareholder control if and only if $\theta > \theta^*$. To analyze the properties of partial pooling equilibria, it is convenient to distinguish between the threshold of directors in the peer firm, θ_j^*, and the threshold of directors in the same firm, θ_i^*. Lemma IA.5 below shows that the best response threshold of a director in firm i to the threshold strategies θ_i^* and θ_j^* of directors in firm i and j, respectively, is given by

$$\beta\left(\theta_i^*, \theta_j^*\right) = \Delta^{-1}\left(\alpha\frac{\delta}{K}\frac{\Upsilon\left(\theta_i^*, \theta_j^*\right)}{C_{T-1}^{K-1}\left(1 - F\left(\theta_i^*\right)\right)^{T-1} F\left(\theta_i^*\right)^{K-T}}\right), \tag{IA.5}$$

where

$$\begin{aligned}\Upsilon\left(\theta_i^*, \theta_j^*\right) = \left(1 - \tau\left(\theta_j^*\right)\right) \times &\left((1-\delta)\left(1 - F\left(\theta_i^*\right)\right)^{K-1} + \tfrac{1-\left(1-F\left(\theta_i^*\right)\right)^{K-1}}{F\left(\theta_i^*\right)}\right) \\ -\tau\left(\theta_j^*\right) \times &\left((1-\delta) F\left(\theta_i^*\right)^{K-1} + \tfrac{1-F\left(\theta_i^*\right)^{K-1}}{1-F\left(\theta_i^*\right)}\right).\end{aligned}$$

It follows that any symmetric partial pooling equilibrium is a solution to $\beta\left(\theta^*, \theta^*\right) = \theta^*$.

The expression (IA.5) shows that there are three distinct effects of how the voting decision of a director is affected by the strategies of his peers. First, since $\tau(\cdot)$ is a decreasing function, $\beta\left(\theta_i^*, \theta_j^*\right)$ increases with θ_j^*, the threshold of directors in the other firm. Hence, there are strategic *complementarities* between directors' decisions *across* firms, by the same intuition as in the basic model. Second, since the function $\frac{1-x^K}{1-x}$ is increasing, $\Upsilon\left(\theta_i^*, \theta_j^*\right)$ decreases with θ_i^*. Therefore, ignoring the effect of θ_i^* on the probability of being pivotal, $C_{T-1}^{K-1}\left(1 - F\left(\theta_i^*\right)\right)^{T-1} F\left(\theta_i^*\right)^{K-T}$, the best response $\beta\left(\theta_i^*, \theta_j^*\right)$ decreases with θ_i^*. This suggests that there is strategic *substitutability* between directors' decisions *within* firms. The reason is that directors within a firm compete for the board seat of the other firm. If θ_i^* decreases, more directors on the same board are likely to take shareholder-friendly actions, and hence the probability of being invited to a shareholder-friendly board conditional on having a shareholder-friendly reputation is lower due to a higher supply of shareholder-friendly directors. On the margin, this effect increases the incentives of each director to create a management-friendly reputation by voting against shareholder control. It follows that the labor market for directors creates incentives for *nonconformity* within the boardroom. Overall, these two effects, across and within firms, work in opposite directions, with one effect dominating the other depending on the parameters of the model. The final effect of θ_i^*, *amplification*, is related to the probability of being pivotal: the lower is the probability of being pivotal, the more important, in relative terms, are directors' reputational concerns. Both the substitution effect and the amplification effect exist only if individual votes are observable and do not arise in the basic model.

To prove Proposition IA.2, we start by proving two auxiliary results, Lemma IA.4 and Lemma IA.5.

LEMMA IA.4: *For any $\lambda \in (0,1)$ and integer Y,*

$$\sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \frac{1}{Y+1-j} = \frac{1}{Y+1} \frac{1-\lambda^{Y+1}}{1-\lambda}. \tag{IA.6}$$

Proof of Lemma IA.4: We first prove that

$$\sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \times \frac{j}{Y+1-j} = \frac{\lambda}{1-\lambda} \left[1-\lambda^Y\right]. \tag{IA.7}$$

Since $C_j^Y \times \frac{j}{Y+1-j} = C_{j-1}^Y$ and since for $j=0$ the expression in the sum on the left-hand side is zero, we have

$$\sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \times \frac{j}{Y+1-j} = \sum_{j=1}^{Y} C_{j-1}^Y \lambda^j (1-\lambda)^{Y-j} = \sum_{h=0}^{Y-1} C_h^Y \lambda^{h+1} (1-\lambda)^{Y-h-1} = \\ \frac{\lambda}{1-\lambda} \left[\sum_{h=0}^{Y} C_h^Y \lambda^h (1-\lambda)^{Y-h} - \lambda^Y\right] = \frac{\lambda}{1-\lambda} \left[1-\lambda^Y\right].$$

Next, note that

$$\begin{aligned} \sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \times \frac{j}{Y+1-j} &= \sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \times \left(\frac{Y+1}{Y+1-j} - 1\right) \\ &= (Y+1) \sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \frac{1}{Y+1-j} - 1. \end{aligned} \tag{IA.8}$$

Combining (IA.7) and (IA.8), we get (IA.6).

LEMMA IA.5:

1. *Consider a symmetric equilibrium. If a director is pivotal with a positive probability, he votes for the proposal if and only if his type exceeds some threshold.*

2. *Consider a symmetric equilibrium with a threshold θ^*. The expected utility of a director of type θ_{ik} from voting for shareholder control relative to voting against it is given by*

$$\begin{aligned} &\mathbb{E}U(1) - \mathbb{E}U(0) = \Delta(\theta_{ik}) \times C_{T-1}^{K-1} (1-F(\theta^*))^{T-1} F(\theta^*)^{K-T} \\ &+\alpha \frac{\delta}{K} \times \tau(\theta^*) \times \left[(1-\delta) F(\theta^*)^{K-1} + \frac{1-F(\theta^*)^{K-1}}{1-F(\theta^*)}\right] \\ &-\alpha \frac{\delta}{K} \times (1-\tau(\theta^*)) \times \left[(1-\delta)(1-F(\theta^*))^{K-1} + \frac{1-(1-F(\theta^*))^{K-1}}{F(\theta^*)}\right]. \end{aligned} \tag{IA.9}$$

3. *If directors in firm i (j) follow a threshold strategy with a threshold θ_i^* (θ_j^*), then the best response of a director in firm i is given by (IA.5).*

Proof of Lemma IA.5: Proof of part 1. Consider any director's decision of whether to vote to transfer control to shareholders. We focus on symmetric equilibria, where all directors, within and across firms, have the same strategy. Denote by $p^*(t,K)$ the equilibrium probability that exactly t of K directors vote for shareholder control, and by τ^* the likelihood that shareholders obtain control. If a director of type θ_i in firm i votes to transfer control to shareholders, his expected utility is given by

$$\begin{aligned}&\mathbb{E}U(1) = v(0,\theta_i)\textstyle\sum_{t=0}^{T-2} p^*(t,K-1) + v(1,\theta_i)\sum_{t=T-1}^{K-1} p^*(t,K-1)\\&+\alpha\delta\times\tau^*\times\left[(1-\delta)\textstyle\sum_{t=0}^{K-1} p^*(t,K-1)\frac{1}{t+1} + \delta\frac{K-1}{K}\sum_{t=0}^{K-2} p^*(t,K-2)\frac{1}{t+1}\right].\end{aligned}$$

The first two terms represent the direct expected utility from the allocation of control in firm i. The third term is the expected utility of being invited to the board of firm j, with the first component corresponding to the case of no resignation shock in firm i, and the second component corresponding to a resignation shock occurring. Since individual decisions are observed, the director is never invited to the board of the other firm if that firm is controlled by management, regardless of the allocation of control in his own firm. This is because by voting for shareholder control, the director signals his shareholder-friendliness, and thus the manager of the other firm is better off hiring the outside candidate or a director who voted against shareholder control. Note that with transparency, not only the allocation of control but also the number of directors voting for each option begins to matter. If more directors have voted for shareholder control, the likelihood of being invited to the board of the other firm decreases. Similarly,

$$\begin{aligned}&\mathbb{E}U(0) = v(0,\theta_i)\textstyle\sum_{t=0}^{T-1} p^*(t,K-1) + v(1,\theta_i)\sum_{t=T}^{K-1} p^*(t,K-1)\\&+\alpha\delta\times(1-\tau^*)\times\left[(1-\delta)\textstyle\sum_{t=0}^{K-1} p^*(t,K-1)\frac{1}{K-t} + \delta\frac{K-1}{K}\sum_{t=0}^{K-2} p^*(t,K-2)\frac{1}{K-1-t}\right].\end{aligned}$$

Hence,

$$\begin{aligned}&\mathbb{E}U(1) - \mathbb{E}U(0) = \Delta(\theta_i)\times p^*(T-1,K-1)\\&+\alpha\delta\times\tau^*\times\left[(1-\delta)\textstyle\sum_{t=0}^{K-1}\frac{p^*(t,K-1)}{t+1} + \delta\frac{K-1}{K}\sum_{t=0}^{K-2}\frac{p^*(t,K-2)}{t+1}\right]\\&-\alpha\delta\times(1-\tau^*)\times\left[(1-\delta)\textstyle\sum_{t=0}^{K-1}\frac{p^*(t,K-1)}{K-t} + \delta\frac{K-1}{K}\sum_{t=0}^{K-2}\frac{p^*(t,K-2)}{K-1-t}\right].\end{aligned}$$

Since only the first term depends on θ_i and $\Delta(\cdot)$ is increasing and takes all values on $(-\infty, +\infty)$, it follows that as long as $p^*(T-1, K-1) > 0$ (the probability of being pivotal is positive), the director follows a threshold strategy and votes for shareholder control if and only if his type exceeds some threshold θ^*. Hence, the equilibrium probability that exactly t of K directors vote for shareholder control is given by

$$p(t, K; \theta^*) = C_t^K (1 - F(\theta^*))^t F(\theta^*)^{K-t}, \tag{IA.10}$$

and the likelihood that shareholders obtain control is given by

$$\tau(\theta^*) = \sum_{t=T}^{K} p(t, K; \theta^*). \tag{IA.11}$$

Proof of part 2. Note that

$$\sum_{t=0}^{K-2} \frac{p(t, K-2; \theta^*)}{K-1-t} = \sum_{t=0}^{K-2} C_t^{K-2} (1 - F(\theta^*))^t F(\theta^*)^{K-2-t} \frac{1}{K-1-t}.$$

Applying Lemma IA.4 to $Y \in \{K-1, K-2\}$ and $\lambda = 1 - F(\theta^*)$, we get

$$\begin{array}{c} \sum_{t=0}^{K-2} \frac{p(t,K-2;\theta^*)}{K-1-t} = \frac{1}{K-1} \frac{1-(1-F(\theta^*))^{K-1}}{F(\theta^*)} \\ \sum_{t=0}^{K-1} \frac{p(t,K-1;\theta^*)}{K-t} = \frac{1}{K} \frac{1-(1-F(\theta^*))^{K}}{F(\theta^*)}. \end{array}$$

Note also that

$$\begin{array}{c} \sum_{t=0}^{K-1} \frac{p(t,K-1;\theta^*)}{t+1} = \sum_{t=0}^{K-1} C_t^{K-1} (1 - F(\theta^*))^t F(\theta^*)^{K-1-t} \frac{1}{t+1} \\ \overset{j=K-1-t}{=} \sum_{t=0}^{K-1} C_j^{K-1} (1 - F(\theta^*))^{K-1-j} F(\theta^*)^j \frac{1}{K-j}. \end{array}$$

Therefore, applying Lemma IA.4 to $Y \in \{K-1, K-2\}$ and $\lambda = F(\theta^*)$, we get

$$\sum_{j=0}^{Y} C_j^Y \lambda^j (1-\lambda)^{Y-j} \frac{1}{Y+1-j} = \frac{1}{Y+1} \frac{1-\lambda^{Y+1}}{1-\lambda}$$

and

$$\begin{array}{c} \sum_{t=0}^{K-1} \frac{p(t,K-1;\theta^*)}{t+1} = \frac{1}{K} \frac{1-F(\theta^*)^K}{1-F(\theta^*)} \\ \sum_{t=0}^{K-2} \frac{p(t,K-2;\theta^*)}{t+1} = \frac{1}{K-1} \frac{1-F(\theta^*)^{K-1}}{1-F(\theta^*)}. \end{array}$$

It follows that for a director of type θ_i,

$$\mathbb{E}U(1) - \mathbb{E}U(0) = \Delta(\theta_i) \times p(T-1, K-1; \theta^*)$$
$$+\alpha\delta \times \tau(\theta^*) \times \left[(1-\delta)\tfrac{1}{K}\tfrac{1-F(\theta^*)^K}{1-F(\theta^*)} + \delta\tfrac{K-1}{K}\tfrac{1}{K-1}\tfrac{1-F(\theta^*)^{K-1}}{1-F(\theta^*)}\right]$$
$$-\alpha\delta \times (1-\tau(\theta^*)) \times \left[(1-\delta)\tfrac{1}{K}\tfrac{1-(1-F(\theta^*))^K}{F(\theta^*)} + \delta\tfrac{K-1}{K}\tfrac{1}{K-1}\tfrac{1-(1-F(\theta^*))^{K-1}}{F(\theta^*)}\right],$$

or equivalently,

$$\mathbb{E}U(1) - \mathbb{E}U(0) = \Delta(\theta_i) \times p(T-1, K-1; \theta^*)$$
$$+\alpha\tfrac{\delta}{K} \times \tau(\theta^*) \times \left[(1-\delta)\tfrac{1-F(\theta^*)^K}{1-F(\theta^*)} + \delta\tfrac{1-F(\theta^*)^{K-1}}{1-F(\theta^*)}\right]$$
$$-\alpha\tfrac{\delta}{K} \times (1-\tau(\theta^*)) \times \left[(1-\delta)\tfrac{1-(1-F(\theta^*))^K}{F(\theta^*)} + \delta\tfrac{1-(1-F(\theta^*))^{K-1}}{F(\theta^*)}\right].$$

Finally, using the property $(1-\delta)\frac{1-x^K}{1-x} + \delta\frac{1-x^{K-1}}{1-x} = (1-\delta)x^{K-1} + \frac{1-x^{K-1}}{1-x}$ for $x = F(\theta^*)$ and $x = 1 - F(\theta^*)$ for the second and third terms, respectively, we can rewrite the above expression as (IA.9). The equilibrium threshold θ^* is then determined by the indifference condition $\mathbb{E}U(1)|_{\theta_i=\theta^*} = \mathbb{E}U(0)|_{\theta_i=\theta^*}$.

Proof of part 3. Repeating the arguments in the proof of parts 1 and 2 for a director of type θ_{ik} in firm i, given the thresholds θ_i^* and θ_j^* of directors in firm i and j, respectively, we get the analog of (IA.9) as a function of θ_i^* and θ_j^*. The best response $\beta\left(\theta_i^*, \theta_j^*\right)$ is then the level of θ_{ik} that sets $\mathbb{E}U(1) - \mathbb{E}U(0)$ to zero, which gives (IA.5).

Proof of Proposition IA.2: To prove the proposition, we prove that if $T \notin \{1, K\}$, the following two statements hold:

1. A shareholder-friendly (management-friendly) pooling equilibrium exists if and only if the off-equilibrium beliefs $\hat{\pi}$ satisfy $\hat{\pi} \leq (\geq)\, \mathbb{E}[\theta]$.

2. A shareholder-friendly (management-friendly) pooling equilibrium is trembling hand perfect if and only if the off-equilibrium beliefs satisfy $\hat{\pi} < \mathbb{E}[\theta]$ ($\hat{\pi} > \mathbb{E}[\theta]$).

Proof of statement 1: Suppose the pooling equilibrium is shareholder-friendly (management-friendly). Then the reputational utility from playing the equilibrium strategy and voting for shareholders (management) is $\alpha\frac{\delta}{K}$. The reason is that all directors have the same reputation $\mathbb{E}[\theta]$ and are treated symmetrically, and hence the probability of being invited to a shareholder-friendly (management-friendly) board of firm j is $\frac{\delta}{K}$. The reputational utility from deviating from the equilibrium strategy and voting for management (shareholders) is

$$\alpha\frac{\delta}{K}1_{\{\hat{\pi}=\mathbb{E}[\theta]\}}+\alpha\delta\left[(1-\delta)+\delta\frac{K-1}{K}\right]1_{\{\hat{\pi}>(<)\mathbb{E}[\theta]\}},$$

which depends on the off-equilibrium beliefs $\hat{\pi}$. It follows that this equilibrium exists if and only if

$$\begin{aligned}\alpha\frac{\delta}{K} &\geq \alpha\frac{\delta}{K}1_{\{\hat{\pi}=\mathbb{E}[\theta]\}}+\alpha\delta\left[(1-\delta)+\delta\frac{K-1}{K}\right]1_{\{\hat{\pi}>(<)\mathbb{E}[\theta]\}}\Leftrightarrow \\ 1 &\geq 1_{\{\hat{\pi}=\mathbb{E}[\theta]\}}+(K-\delta)\,1_{\{\hat{\pi}>(<)\mathbb{E}[\theta]\}}.\end{aligned} \tag{IA.12}$$

Therefore, this pooling equilibrium exists if and only if $\hat{\pi}\leq(\geq)\,\mathbb{E}[\theta]$, that is, if deviation to voting for management (shareholders) builds a reputation for being management-friendly (shareholder-friendly).

Proof of statement 2: Consider a pooling equilibrium in which each director in both firms votes for shareholders with probability one. The same argument holds for the pooling equilibrium in which each director votes for management. Consider any director k_0 in firm i and any sequence $\{\sigma_n\}$, $\sigma_n\in(0,1)$, $\lim_{n\to\infty}\sigma_n=1$. In the perturbed game, all directors of firm i except k_0 play their respective equilibrium strategy with probability σ_n and the opposite strategy with probability $1-\sigma_n$. Then each director's probability of voting for shareholder control is strictly between zero and one, and hence director k_0's probability of being pivotal is strictly positive and converges to zero as $n\to\infty$. Denote by $p\,(t,K-1;n)$ the probability that exactly t of $K-1$ directors vote for shareholders in the perturbed game. Note that all directors except k_0 vote for shareholders with probability σ_n. If any director votes for shareholders, his reputation is $\mathbb{E}\,[\theta]$, and if he votes for management, his reputation is $\hat{\pi}$. Also, note that firm j is controlled by shareholders for sure. Consider director k_0's best response strategy.

Note that based on statement 1, the off-equilibrium beliefs $\hat{\pi}$ must satisfy $\hat{\pi}\leq\mathbb{E}\,[\theta]$. There are two cases: $\hat{\pi}=\mathbb{E}\,[\theta]$ and $\hat{\pi}<\mathbb{E}\,[\theta]$. First, suppose $\hat{\pi}=\mathbb{E}\,[\theta]$. In this case, regardless of the voting decisions of directors in firm i, all directors have the same reputation, $\mathbb{E}\,[\theta]$. Therefore, regardless of whether the director votes for management, he is hired by the other firm with probability $\frac{\delta}{K}$. Hence, for each n, director k_0's utility from voting for shareholders is

$$\mathbb{E}U\,(1)=v\,(0,\theta_i)\sum_{t=0}^{T-2}p\,(t,K-1;n)+v\,(1,\theta_i)\sum_{t=T-1}^{K-1}p\,(t,K-1;n)+\alpha\frac{\delta}{K}$$

and his utility from not voting for management is

$$\mathbb{E}U(0) = v(0, \theta_i) \sum_{t=0}^{T-1} p(t, K-1; n) + v(1, \theta_i) \sum_{t=T}^{K-1} p(t, K-1; n) + \alpha \frac{\delta}{K}.$$

Hence,

$$\mathbb{E}U(1) - \mathbb{E}U(0) > 0 \Leftrightarrow \Delta(\theta_i) p(T-1, K-1; n) > 0.$$

Since $\lim_{\theta \to -\infty} \Delta(\theta) = -\infty$, there exists θ_i such that $\Delta(\theta_i) < 0$, and hence voting for shareholders is not the best response of the director if his type is θ_i. This implies that the equilibrium is not trembling hand perfect.

Second, suppose $\hat{\pi} < \mathbb{E}[\theta]$. If the director votes for management, he is never hired by firm j. On the other hand, if the director votes for shareholders, he has a chance of being hired, depending on the decisions of other directors. In this case, the relative benefit from voting for shareholders is the same as the expression in (IA.9), where we substitute $p(T-1, K-1; \theta^*)$ with $p(T-1, K-1; n)$, $\tau(\theta^*)$ with one, and $F(\theta^*)$ with $1 - \sigma_n$. Hence,

$$\begin{aligned} &\mathbb{E}U(1) - \mathbb{E}U(0) = \Delta(\theta_i) \times p(T-1, K-1; n) \\ &+\alpha \tfrac{\delta}{K} \times \left[(1-\delta)(1-\sigma_n)^{K-1} + \tfrac{1-(1-\sigma_n)^{K-1}}{\sigma_n} \right]. \end{aligned}$$

Note that

$$\lim_{n \to \infty} \mathbb{E}_n U(1) - \mathbb{E}_n U(0) = \alpha \frac{\delta}{K} > 0.$$

Hence, for any θ_i, there exists n_0 such that if $n > n_0$, director k_0 is strictly better off voting for shareholders if his type is θ_i. Therefore, a pooling equilibrium with $\hat{\pi} < \mathbb{E}[\theta]$ is trembling hand perfect.

D. Multiple Firms

We consider the extension of the model to $N \geq 2$ firms. As in the basic model, each firm is hit by a resignation shock with probability δ, and the shocks are independent across firms. Based on the allocation of control across firms after the first stage, the market is divided into two sets: firms controlled by shareholders search among directors with a shareholder-friendly reputation, and firms controlled by managers search among directors with a management-friendly reputation. We assume that the labor market allocation is efficient in the following sense. In equilibrium, no firm that is controlled by shareholders (management) hires an outside candidate if a director serving on one of the boards has a shareholder-friendly (management-friendly) reputation and has the capacity to serve on another board. If such a situation occurred, both the firm and the director would be better off matching with each other.

Our first observation is that in most cases, there is an excess supply of directors in each segment of the market. If there are n firms with a demand for shareholder-friendly directors, there are at least n firms controlled by shareholders. Hence, there are at least $(K-1) \times n \geq n$ directors with a shareholder-friendly reputation who can fill these board seats. The only exception is when there is only one firm with a demand for shareholder-friendly directors, and this is the only firm in the economy with shareholder control. In this case, that firm will have to hire the outside candidate. Given the excess supply of directors, we assume that directors are limited to one additional directorship. In practice, the number of board seats that a director can hold is often limited by the director's time constraints or by regulation (for example, many countries impose a limit on the number of directorships).[5] In addition, given that the number of firms that can supply a director with a certain reputation is at least as large as the number of firms that demand a director with this reputation, we assume that only one, randomly chosen, director from each firm has the capacity to get an additional directorship. This assumption can be motivated by the restrictions that companies impose on the number of directorships their board members can hold. Finally, we assume that all incumbent directors with the same reputation have equal probability of being invited to other firms.

We search for symmetric equilibria, which are characterized by a threshold θ^*. Consider the value of reputation π for a director in firm i. Let n be the number of shareholder-controlled firms out of the other $N-1$ firms. Among these firms, let d_{SH} be the number of firms that were hit by a resignation shock and thus have demand for a new (shareholder-friendly) director. Similarly, let d_M be the number of firms with demand for a management-friendly director out of the other $N-1-n$ management-controlled firms. Finally, let $\Lambda(n, d_{SH}, d_M, \pi)$ denote the probability that a director from firm i with reputation π gains an additional directorship.

LEMMA IA.6:

If $\pi > \mathbb{E}[\theta]$, the probability that a director from firm i gains an additional directorship is

$$\Lambda(n, d_{SH}, d_M, \pi) = \frac{1}{K} \times \begin{cases} \frac{d_{SH}}{n} & \text{if } n \geq 1 \\ 0 & \text{if } n = 0. \end{cases}$$

[5] The assumption that each director can add at most one additional directorship also keeps this setup compatible with the basic model, where $N = 2$.

If $\pi < \mathbb{E}[\theta]$, the probability that a director from firm i gains an additional directorship is

$$\Lambda(n, d_{SH}, d_M, \pi) = \frac{1}{K} \times \begin{cases} \frac{d_M}{N-1-n} & \text{if } n \leq N-2 \\ 0 & \text{if } n = N-1. \end{cases}$$

Taking the expectation over possible realizations of n, d_{SH}, and d_M, we derive the expected value of reputation as follows.

LEMMA IA.7: *If τ is the probability that each firm is controlled by shareholders, the expected value of reputation is $\alpha\frac{\delta}{K}\left(1-(1-\tau)^{N-1}\right)$ if $\pi > \mathbb{E}[\theta]$, and $\alpha\frac{\delta}{K}\left(1-\tau^{N-1}\right)$ if $\pi < \mathbb{E}[\theta]$.*

It follows from Lemma IA.7 that, similar to the basic model, a reputation for being shareholder-friendly generates a higher payoff than a reputation for being management-friendly if and only if $\tau(\theta^*) > 0.5$. Therefore, Definition 3 of shareholder-friendly and management-friendly equilibria can be extended to any number of firms.

By Lemma IA.7, an equilibrium with a threshold θ^* exists if and only if $\theta^* = \beta_N(\theta^*)$, where

$$\beta_N(\theta) = \Delta^{-1}\left(\alpha\frac{\delta}{K}\left((1-\tau(\theta))^{N-1} - \tau(\theta)^{N-1}\right)\right).$$

The best response function $\beta_N(\theta)$ coincides with (7) for $N = 2$ and is strictly increasing in θ for any $N \geq 2$. Hence, the extended model exhibits strategic complementarity as well. Moreover, using the arguments in the proof of Proposition 2, it is immediate that the comparative statics of θ^* (and hence of $\tau(\theta^*)$) with respect to the parameters of the model is the same as in the basic model. Note also that as the number of firms becomes infinitely large, the externalities due to reputational effects disappear. However, given that the labor market for directors is somewhat segmented both by industry and by geographical location (see the discussion in Section III of the published article), we think of N as representing the number of firms in the relevant segment and hence not being very large.

Proof of Lemma IA.6: The probability that a director from firm i does not resign and is chosen as the director within his firm who can take another board seat is $(1-\delta)\frac{1}{K} + \delta\frac{K-1}{K}\frac{1}{K-1} = \frac{1}{K}$. Suppose that firm i is controlled by shareholders and hence the director's reputation satisfies $\pi > \mathbb{E}[\theta]$. If $n = 0$, the director will not gain an additional directorship. Suppose $n \geq 1$ and consider any firm j that has demand for a shareholder-friendly director. Then a director from firm i faces competition from $n-1$ directors, one from each of the $n-1$ firms that are controlled by shareholders, excluding firm j (existing directors of firm j cannot fill the vacancy in firm j). Thus, the director from firm i has probability $\frac{1}{n}$ of gaining a directorship in firm j. With probability $\frac{n-1}{n}$, the director from firm i does not gain a directorship in firm j and is competing with $n-2$ directors for the other $d_{SH}-1$ available directorships. Let $p\left(d_{SH},n\right)$ denote the probability that a director from firm i fills one of the d_{SH} directorships. The above argument implies that $p\left(1,n\right)=\frac{1}{K}\frac{1}{n}$ and that

$$p\left(d_{SH},n\right)=\frac{1}{K}\frac{1}{n}+\frac{n-1}{n}\times p\left(d_{SH}-1,n-1\right)$$

for $d_{SH}>1$. Conjecture that $p\left(d_{SH},n\right)=\frac{1}{K}\frac{d_{SH}}{n}$. By induction,

$$p\left(d_{SH},n\right)=\frac{1}{K}\frac{1}{n}+\frac{n-1}{n}\times\frac{1}{K}\frac{d_{SH}-1}{n-1}=\frac{1}{K}\frac{d_{SH}}{n},$$

which confirms the conjecture. The proof for the case in which firm i is controlled by management is similar.

Proof of Lemma IA.7: The expected value of reputation is given by

$$\begin{aligned}\Upsilon\left(\pi\right)=\alpha\textstyle\sum_{n=0}^{N-1}&\left[C_n^{N-1}\tau^n\left(1-\tau\right)^{N-1-n}\times\right.\\ &\left.\times\textstyle\sum_{i=0}^{n}\sum_{j=0}^{N-1-n}\left[C_i^n\delta^i\left(1-\delta\right)^{n-i}\times C_j^{N-1-n}\delta^j\left(1-\delta\right)^{N-1-n-j}\times\Lambda\left(n,i,j,\pi\right)\right]\right],\end{aligned}$$

where $\Lambda\left(n,i,j,\pi\right)$ is given by Lemma IA.6. Therefore, for $\pi>\mathbb{E}\left[\theta\right]$,

$$\begin{aligned}\Upsilon\left(\pi\right)&=\alpha\tfrac{1}{K}\textstyle\sum_{n=1}^{N-1}\left[C_n^{N-1}\tau^n\left(1-\tau\right)^{N-1-n}\times\right.\\ &\quad\left.\times\textstyle\sum_{i=0}^{n}\sum_{j=0}^{N-1-n}\left[C_i^n\delta^i\left(1-\delta\right)^{n-i}\times C_j^{N-1-n}\delta^j\left(1-\delta\right)^{N-1-n-j}\times\tfrac{i}{n}\right]\right]\\ &=\alpha\tfrac{1}{K}\textstyle\sum_{n=1}^{N-1}\left[\begin{array}{c}C_n^{N-1}\tau^n\left(1-\tau\right)^{N-1-n}\frac{1}{n}\times\sum_{i=0}^{n}C_i^n\delta^i\left(1-\delta\right)^{n-i}i\\ \times\sum_{j=0}^{N-1-n}C_j^{N-1-n}\delta^j\left(1-\delta\right)^{N-1-n-j}\end{array}\right]\\ &=\alpha\tfrac{1}{K}\textstyle\sum_{n=1}^{N-1}\left[C_n^{N-1}\tau^n\left(1-\tau\right)^{N-1-n}\frac{1}{n}\times\sum_{i=0}^{n}C_i^n\delta^i\left(1-\delta\right)^{n-i}i\right]\\ &=\alpha\tfrac{1}{K}\textstyle\sum_{n=1}^{N-1}\left[C_n^{N-1}\tau^n\left(1-\tau\right)^{N-1-n}\frac{1}{n}\times n\delta\right]\\ &=\alpha\tfrac{\delta}{K}\textstyle\sum_{n=1}^{N-1}C_n^{N-1}\tau^n\left(1-\tau\right)^{N-1-n}=\alpha\tfrac{\delta}{K}\left(1-\left(1-\tau\right)^{N-1}\right).\end{aligned}$$

Similarly, for $\pi < \mathbb{E}[\theta]$,

$$\begin{aligned}
\Upsilon(\pi) &= \alpha \tfrac{1}{K} \textstyle\sum_{n=0}^{N-2} \Big[C_n^{N-1} \tau^n (1-\tau)^{N-1-n} \times \\
&\times \textstyle\sum_{j=0}^{N-1-n} \sum_{i=0}^{n} \Big[C_i^n \delta^i (1-\delta)^{n-i} \times C_j^{N-1-n} \delta^j (1-\delta)^{N-1-n-j} \times \tfrac{j}{N-1-n} \Big] \Big] \\
&= \alpha \tfrac{1}{K} \textstyle\sum_{n=0}^{N-2} \left[\begin{array}{c} C_n^{N-1} \tau^n (1-\tau)^{N-1-n} \frac{1}{N-1-n} \times \sum_{j=0}^{N-1-n} C_j^{N-1-n} \delta^j (1-\delta)^{N-1-n-j} j \\ \times \sum_{i=0}^{n} C_i^n \delta^i (1-\delta)^{n-i} \end{array} \right] \\
&= \alpha \tfrac{1}{K} \textstyle\sum_{n=0}^{N-2} \Big[C_n^{N-1} \tau^n (1-\tau)^{N-1-n} \tfrac{1}{N-1-n} \times \sum_{j=0}^{N-1-n} C_j^{N-1-n} \delta^j (1-\delta)^{N-1-n-j} j \Big] \\
&= \alpha \tfrac{1}{K} \textstyle\sum_{n=0}^{N-2} \Big[C_n^{N-1} \tau^n (1-\tau)^{N-1-n} \tfrac{1}{N-1-n} \times (N-1-n)\,\delta \Big] \\
&= \alpha \tfrac{\delta}{K} \textstyle\sum_{n=0}^{N-2} C_n^{N-1} \tau^n (1-\tau)^{N-1-n} = \alpha \tfrac{\delta}{K} \left(1 - \tau^{N-1}\right).
\end{aligned}$$

E. Endogenous Benefit from Additional Directorships

In the basic model, we take the benefit from an additional directorship as given. The goal of this section is to microfound α. For this purpose, we simplify the model on one dimension and extend it on another dimension. In particular, we abstract from the collective decision-making within the board by assuming $K = 1$ (and hence $T = 1$). That is, each board consists of one director, who solely determines the allocation of control in his firm. As in the basic model, each firm is hit by a resignation shock with probability δ, in which case the director resigns and the firm searches for a replacement, and the resignation shocks are independent across firms.

We extend the model in two ways. First, we introduce a third stage, during which directors make a decision about the firm's strategy. In particular, after the labor market for directors clears, the board of each firm chooses between strategy s_{SH} and strategy s_M. As specified below, shareholders prefer strategy s_{SH}, management prefers strategy s_M, and more shareholder-friendly directors get a higher (lower) utility from strategy s_{SH} (s_M) relative to less shareholder-friendly directors. Second, we allow the party making the director appointment decision to offer the new director a contract that is contingent on his action at the third stage. In particular, when a new director joins firm j and before he chooses strategy s_j at the third stage, the controlling party of firm j offers him a take-it-or-leave-it contract, which specifies payment

$\alpha(s_j)$ upon the choice of strategy s_j.[6,7] Directors have limited liability, normalized at zero.[8] Hence, the director always accepts the offer regardless of his type.

The utility functions of the players are as follows. First, as in the basic model, shareholders and management of the firm derive direct utility from the allocation of control. However, different from the basic model, they do not get utility from the composition of the board, but rather from the strategy chosen by the board. Finally, we assume that shareholders and managers at least partly internalize the cost of hiring directors. In particular, given strategy $s_i \in \{s_M, s_{SH}\}$ at firm i, the utility functions of shareholders and management of firm i are given by

$$\begin{aligned} u_{SH}(\chi_i, s_i) &= v_{SH}(\chi_i) + \hat{g}_{SH} \times 1\{s_i = s_{SH}\} - \rho_{SH}\alpha(s_i), \\ u_M(\chi_i, s_i) &= v_M(\chi_i) + \hat{g}_M \times 1\{s_i = s_M\} - \rho_M\alpha(s_i), \end{aligned}$$

where $\rho_{SH} > 0$ and $\rho_M > 0$ capture the extent to which shareholders and managers internalize the cost of hiring the director.[9] Similar to the basic model, we assume $v_{SH}(0) = v_M(1) = 0$, $v_{SH}(1) \geq v_M(0) > 0$, and $\hat{g}_{SH} > 0$, $\hat{g}_M > 0$. In other words, shareholders always prefer shareholder control over management control and strategy s_{SH} over strategy s_M, while management has the opposite preferences.

The utility of director i is a sum of his utility from the position at firm i and his utility from the position at firm j if he is invited there at the second stage. The director's utility from the position at firm i consists of two components. The first is $v(\chi_i, \theta_i)$, his direct utility from the allocation of control, which has the same properties as in the basic model. The second is given by a function $g(s_i, \theta_i)$, which can be thought of as the director's private benefits from

[6] The assumption that the choice of strategy is contractible is made for simplicity. Alternatively, we could consider a model in which the firm's output is a function of the director's chosen strategy and a random shock, and output is contractible. In the extreme case, when the distribution of output perfectly reveals the chosen strategy, such a model is equivalent to the one we analyze.

[7] Note that since the only thing that happens after the director accepts the contract is his choice of strategy, offering a menu of contracts instead of a single contract will not make the controlling party better off. Indeed, suppose the controlling party offers a menu of contracts $\left(\alpha^{(t)}(s_{SH}), \alpha^{(t)}(s_M)\right)$, indexed by t. Let Θ_{SH} and Θ_M denote the sets of types who, after choosing one of the contracts, take action s_{SH} and s_M, respectively. Then any type $\theta \in \Theta_{SH}$ will choose contract $t_{SH} \in \arg\max_t \alpha^{(t)}(s_{SH})$, and any type $\theta \in \Theta_M$ will choose contract $t_M \in \arg\max_t \alpha^{(t)}(s_M)$. Effectively, the director chooses between $\max_t \alpha^{(t)}(s_{SH})$ and $\max_t \alpha^{(t)}(s_M)$ and implements strategy s_{SH} if and only if $g(s_{SH}, \theta) + \max_t \alpha^{(t)}(s_{SH}) > g(s_M, \theta) + \max_t \alpha^{(t)}(s_M)$. Hence, this menu of contracts is equivalent to a single contract $\left(\max_t \alpha^{(t)}(s_{SH}), \max_t \alpha^{(t)}(s_M)\right)$.

[8] The results would continue to hold if directors' minimum salary were $\alpha_0 > 0$.

[9] For example, the case $\rho_{SH} = 1$ and $\rho_M > 0$ captures the situation in which the director is paid out of shareholders' funds, and the manager partly internalizes this cost because he cares about shareholder value, for example, due to reputational considerations. All the results hold for any $\rho_{SH} > 0$ and $\rho_M > 0$.

choosing strategy s_i (which he gets whenever he does not resign from firm i). We assume that $g(s_{SH},\theta)$ is nonnegative and increasing in θ and that $g(s_M,\theta)$ is nonnegative and decreasing in θ. We define $\gamma(\theta) \equiv g(s_{SH},\theta) - g(s_M,\theta)$ and assume that $\lim_{\theta\to\infty}\gamma(\theta) > 0 > \lim_{\theta\to-\infty}\gamma(\theta)$.

If director i is invited to firm j, his utility from this position is the sum of $g(s_j,\theta_i)$, his utility from the choice of strategy s_j at firm j, and $\alpha(s_j)$, his payment from the contract. Overall, director i's utility is given by $v(\chi_i,\theta_i)$ if he resigns, by $v(\chi_i,\theta_i) + g(s_i,\theta_i)$ if he does not resign but is not invited to the board of firm j, and by

$$v(\chi_i,\theta_i) + g(s_i,\theta_i) + g(s_j,\theta_i) + \alpha(s_j)$$

if he does not resign and is invited to the board of firm j.

Consider the third stage and the decision of a director who was invited to firm j and offered a contract $(\alpha(s_{SH}),\alpha(s_M))$. If $\alpha(s_{SH}) - \alpha(s_M) = \alpha$, the director will choose strategy s_{SH} if and only if $\gamma(\theta) + \alpha > 0$. Since $\gamma(\theta)$ is increasing, there exists a unique, potentially infinite cutoff $y(\alpha)$ such that a director of type θ chooses strategy s_{SH} over strategy s_M if and only if $\theta > y(\alpha)$. In particular, $y(\alpha)$ decreases in α and is given by

$$y(\alpha) = \begin{cases} -\infty & \textit{if } -\alpha \leq \lim_{\theta\to-\infty}\gamma(\theta) \\ \gamma^{-1}(-\alpha) & \textit{if } \lim_{\theta\to-\infty}\gamma(\theta) < -\alpha < \lim_{\theta\to\infty}\gamma(\theta) \\ \infty & \textit{if } -\alpha \geq \lim_{\theta\to\infty}\gamma(\theta). \end{cases} \tag{IA.13}$$

In what follows, we focus on symmetric threshold equilibria, where at the first stage, a director votes for shareholder control if and only if his type exceeds some threshold θ^*.[10] The next proposition characterizes the equilibria of the game.

PROPOSITION IA.3: *An equilibrium with a threshold θ^* exists if and only if θ^* satisfies the equation*

$$\theta^* = \Delta^{-1}\left(\ \delta(1-\delta)(2F(\theta^*)-1) \times \max\{g(s_{SH},\theta^*), g(s_M,\theta^*)\}\right). \tag{IA.14}$$

In this equilibrium, if firm j is controlled by shareholders (management), it hires the director

[10] The proof of Proposition IA.3 shows that if $\frac{\partial}{\partial\theta}\Delta(\theta) > \max\left\{\frac{\partial}{\partial\theta}g(s_{SH},\theta), -\frac{\partial}{\partial\theta}g(s_M,\theta)\right\}$, then any equilibrium is a threshold equilibrium. This assumption implies that the director's utility is more sensitive to the allocation of control than to his private benefit from the selection of the strategy at the third stage. Note that this assumption holds naturally in the basic model, where the benefit from joining the board of firm j is independent of the director's type.

of firm i if and only if $\chi_i = 1$ ($\chi_i = 0$) and offers him the contract $\alpha(s_{SH}) = \alpha_{SH}(\theta^*)$; $\alpha(s_M) = 0$ ($\alpha(s_{SH}) = 0$; $\alpha(s_M) = \alpha_M(\theta^*)$), where

$$\begin{array}{r} \alpha_{SH}(\theta^*) = \arg\max_{\alpha_{SH}\geq 0} \Pr[\theta > y(\alpha_{SH}) | \theta \geq \theta^*](\hat{g}_{SH} - \rho_{SH}\alpha_{SH}) \\ \alpha_M(\theta^*) = \arg\max_{\alpha_M \geq 0} \Pr[\theta < y(-\alpha_M) | \theta \leq \theta^*](\hat{g}_M - \rho_M\alpha_M). \end{array} \tag{IA.15}$$

The payment $\alpha_{SH}(\theta^*)$ decreases in θ^* and equals zero when $\theta^* \geq y(0)$. The payment $\alpha_M(\theta^*)$ increases in θ^* and equals zero when $\theta^* \leq y(0)$.

Even though shareholders and management no longer derive direct utility from the director's type, as was assumed in the basic model, the equilibrium still features the property that shareholders (management) of firm j will only prefer a director from firm i over the outside candidate if that director created a shareholder-friendly (management-friendly) reputation by allocating control to shareholders (managers). Intuitively, a director with a shareholder-friendly (management-friendly) reputation is more (less) likely to choose the shareholder-preferred strategy s_{SH} than the outside candidate, given the same compensation.

The intuition behind the equilibrium contracts is as follows. Consider the contract offered by shareholders (the intuition for contracts offered by management is similar). As expected, shareholders never compensate the director for choosing their least preferred strategy. The payment for choosing their most preferred strategy, $\alpha_{SH}(\theta^*)$, decreases in θ^* for the following reason. When θ^* increases, the equilibrium becomes more management-friendly, and thus the likelihood of being hired upon giving control to shareholders decreases. Hence, the fact that a director gave control to shareholders becomes a stronger signal that he is shareholder-friendly: his reputation $\mathbb{E}[\theta|\theta > \theta^*]$ increases. A more shareholder-friendly reputation implies that the director's interests are more aligned with those of shareholders, and hence he needs smaller incentives to choose the shareholder-preferred strategy. Interestingly, this logic leads to the result that shareholders do not give any compensation for choosing their most preferred strategy if the equilibrium is sufficiently management-friendly: $\alpha_{SH}(\theta^*) = 0$ if $\theta^* \geq y(0)$. In other words, the contract offered by shareholders is not contingent on the director's choice of strategy. This is because when the equilibrium is sufficiently management-friendly, allocating control to shareholders is a signal that the director is very shareholder-friendly ($\theta_i > \theta^* \geq y(0)$) and hence will choose strategy s_{SH} even if he is not compensated for doing so. Similarly, managers do not offer a contingent contract if the equilibrium is sufficiently shareholder-friendly. Note that if $F(y(0)) = \frac{1}{2}$, that is, the average type is indifferent between the two strategies, then $\theta^* \geq y(0)$ if and only if the equilibrium is management-friendly. This leads to the following corollary.

COROLLARY IA.1: *Suppose $F(y(0)) = \frac{1}{2}$. Then the contract offered by shareholders is contingent on the director's chosen strategy if and only if the equilibrium is shareholder-friendly, and the contract offered by management is contingent on the director's chosen strategy if and only if the equilibrium is management-friendly.*

We next examine how the equilibrium of the game changes if contingent contracts are not allowed and the only utility of a director from an additional board seat is $g(s_j, \theta_i)$. Surprisingly, we show that the allocation of control is unaffected by contingent contracts. Indeed, if contracts are not allowed, the director votes for shareholder control if and only if

$$\begin{aligned} & v(1, \theta_i) + \delta(1-\delta)\tau_j \times \max\{g(s_M, \theta_i), g(s_{SH}, \theta_i)\} \\ > \ & v(0, \theta_i) + \delta(1-\delta)(1-\tau_j) \times \max\{g(s_M, \theta_i), g(s_{SH}, \theta_i)\}. \end{aligned} \tag{IA.16}$$

Hence, any equilibrium threshold satisfies (IA.14) as well. Thus, when contracts are allowed and are designed optimally by the controlling party, the set of equilibrium thresholds θ^* is the same as when contracts are not allowed: contracts change directors' decisions about strategy at the third stage, but not their decisions about the allocation of control at the first stage.

The intuition is as follows. Suppose shareholders (management) hire the director of the other firm, knowing that his type is above (below) θ^*, and offer him incentives to implement their preferred strategy: $\alpha_{SH}(\theta^*) > 0$ ($\alpha_M(\theta^*) > 0$). Then they will never offer enough to induce even the threshold type θ^* to strictly prefer strategy s_{SH} (s_M); otherwise, since all types they hire are above (below) θ^*, they could reduce the director's compensation without changing any of these types' decision about the strategy. Hence, the threshold type implements the same third-stage strategy as if he were not offered the equilibrium contract, which implies that his incentives at the first stage are not affected by the contract either.

The result that the set of equilibrium thresholds θ^* is invariant to the availability of contingent contracts allows us to study the welfare implications of contingent contracts. Similar to the basic model, we assume that the combined utility of shareholders and managers is higher under a shareholder-preferred than under a management-preferred strategy, that is, $\hat{g}_{SH} > \hat{g}_M$.

LEMMA IA.8: *Suppose $\hat{g}_{SH} > \hat{g}_M$ and consider an equilibrium with a threshold θ^*. There exists $\overline{\theta} < \infty$ such that if $\theta^* > \overline{\theta}$, then contingent compensation contracts decrease the aggregate* *utility of shareholders and managers.*

Intuitively, when the governance system is weak (θ^* is high), management is likely to have control, and therefore contingent compensation contracts induce directors to choose the management-preferred strategy more often, decreasing the aggregate welfare of shareholders and managers. In this respect, contingent compensation contracts amplify the adverse effect of weak governance systems.

Proof of Proposition IA.3: First, suppose that shareholders control firm j and offer a contract to the director of firm i. Clearly, $\alpha(s_M) = 0$, regardless of the director's reputation. Define $\alpha_{SH} \equiv \alpha(s_{SH})$. Then $\alpha_{SH} = \arg\max_{\alpha_{SH}\geq 0} V_i(\alpha_{SH}, \chi_i)$, where

$$V_i(\alpha_{SH}, \chi_i) = \Pr(\theta > y(\alpha_{SH})|\chi_i)(\hat{g}_{SH} - \rho_{SH}\alpha_{SH}). \tag{IA.17}$$

Similarly, a contract offered to an outside candidate will feature $\alpha(s_M) = 0$ and $\alpha(s_{SH}) = \arg\max_{\alpha_{SH}\geq 0} V_{out}(\alpha_{SH})$, where

$$V_{out}(\alpha_{SH}) = \Pr(\theta > y(\alpha_{SH}))(\hat{g}_{SH} - \rho_{SH}\alpha_{SH}). \tag{IA.18}$$

We next prove that shareholders prefer to hire the director of firm i if $\chi_i = 1$ and the outside candidate if $\chi_i = 0$. For any α_{SH},

$$\Pr(\theta > y(\alpha_{SH})|\chi_i = 1) = \tfrac{\Pr(\theta > y(\alpha_{SH}),\theta>\theta^*)}{\Pr(\theta>\theta^*)}$$
$$= \begin{cases} 1 & \text{if } y(\alpha_{SH}) < \theta^* \\ \frac{\Pr(\theta>y(\alpha_{SH}))}{\Pr(\theta>\theta^*)} & \text{if } y(\alpha_{SH}) > \theta^* \end{cases} \geq \Pr(\theta > y(\alpha_{SH})),$$
$$\Pr(\theta > y(\alpha_{SH})|\chi_i = 0) = 1 - \Pr(\theta < y(\alpha_{SH})|\theta < \theta^*) = 1 - \tfrac{\Pr(\theta<y(\alpha_{SH}),\theta<\theta^*)}{\Pr(\theta<\theta^*)}$$
$$= \begin{cases} 1 - \frac{\Pr(\theta<y(\alpha_{SH}))}{\Pr(\theta<\theta^*)} & \text{if } y(\alpha_{SH}) < \theta^* \\ 0 & \text{if } y(\alpha_{SH}) > \theta^* \end{cases} \leq 1 - \Pr(\theta < y(\alpha_{SH})) = \Pr(\theta > y(\alpha_{SH})).$$

It follows that for any α_{SH}, $V_i(\alpha_{SH}, 1) \geq V_{out}(\alpha_{SH})$ and $V_i(\alpha_{SH}, 0) \leq V_{out}(\alpha_{SH})$. Hence, $\max_{\alpha_{SH}\geq 0} V_i(\alpha_{SH}, 1) \geq \max_{\alpha_{SH}\geq 0} V_{out}(\alpha_{SH})$ and $\max_{\alpha_{SH}\geq 0} V_i(\alpha_{SH}, 0) \leq \max_{\alpha_{SH}\geq 0} V_{out}(\alpha_{SH})$, which implies that shareholders prefer the director of firm i over the outside candidate if and only if $\chi_i = 1$.

Next, suppose that management controls firm j. By the same arguments as above, management will prefer the director of firm i over the outside candidate if and only if $\chi_i = 0$. If management hires director i, it will offer $\alpha(s_{SH}) = 0$ and $\alpha(s_M) = \alpha_M$, where α_M solves

$$\max_{\alpha_M \geq 0} \Pr[\theta < y(-\alpha_M)|\chi_i](\hat{g}_M - \rho_M \alpha_M).$$

Given (α_M, α_{SH}) and the expectation that the director will be hired by the other firm if and only if the allocation of control across the firms is the same, director of firm i gives control to shareholders if and only if

$$\begin{gathered} v(1,\theta_i) + (1-\delta)\max\{g(s_M,\theta_i), g(s_{SH},\theta_i)\} \\ +\delta(1-\delta)\tau_j \times \max\{g(s_M,\theta_i), g(s_{SH},\theta_i) + \alpha_{SH}\} > \\ v(0,\theta_i) + (1-\delta)\max\{g(s_M,\theta_i), g(s_{SH},\theta_i)\} \\ +\delta(1-\delta)(1-\tau_j) \times \max\{g(s_M,\theta_i) + \alpha_M, g(s_{SH},\theta_i)\}. \end{gathered}$$

or equivalently, if and only if

$$\Delta(\theta_i) > \delta(1-\delta)\left[\begin{array}{c} (1-\tau_j) \times \max\{g(s_M,\theta_i) + \alpha_M, g(s_{SH},\theta_i)\} \\ -\tau_j \times \max\{g(s_M,\theta_i), g(s_{SH},\theta_i) + \alpha_{SH}\}. \end{array}\right] \tag{IA.19}$$

Since $y(\cdot)$ is decreasing, $y(\alpha_{SH}) \leq y(0) \leq y(-\alpha_M)$. Consider three ranges of the director's type θ_i:

1. If $\theta_i \geq y(-\alpha_M) \Leftrightarrow \alpha_M \leq g(s_{SH},\theta_i) - g(s_M,\theta_i)$, (IA.19) is equivalent to

$$\Delta(\theta_i) > \delta(1-\delta)[(1-\tau_j) \times g(s_{SH},\theta_i) - \tau_j \times (g(s_{SH},\theta_i) + \alpha_{SH})]. \tag{IA.20}$$

2. If $\theta_i \in (y(\alpha_{SH}), y(-\alpha_M)) \Leftrightarrow -\alpha_{SH} \leq g(s_{SH},\theta_i) - g(s_M,\theta_i) \leq \alpha_M$, (IA.19) is equivalent to

$$\Delta(\theta_i) > \delta(1-\delta)[(1-\tau_j) \times (g(s_M,\theta_i) + \alpha_M) - \tau_j \times (g(s_{SH},\theta_i) + \alpha_{SH})]. \tag{IA.21}$$

3. If $\theta_i \leq y(\alpha_{SH}) \Leftrightarrow g(s_{SH},\theta_i) - g(s_M,\theta_i) \leq -\alpha_{SH}$, (IA.19) is equivalent to

$$\Delta(\theta_i) > \delta(1-\delta)[(1-\tau_j) \times (g(s_M,\theta_i) + \alpha_M) - \tau_j \times g(s_M,\theta_i)]. \tag{IA.22}$$

Consider the equilibrium threshold θ^*. Note that there is no threshold equilibrium in which $\theta^* \in (y(\alpha_{SH}), y(-\alpha_M))$. Suppose first that $\theta^* < y(-\alpha_M)$ and consider a realization $\chi_i = \chi_j = 0$. Management of firm j understands that $\theta_i \leq \theta^*$, and hence $\theta_i < y(-\alpha_M)$.

Hence, given α_M, director i will choose strategy s_M with probability one. If $\alpha_M > 0$, then by lowering α_M by a sufficiently small $\varepsilon > 0$, firm j can still guarantee that the director will choose s_M with probability one. Thus, this cannot be an equilibrium unless $\alpha_M = 0$. By a similar argument, $\theta^* > y(\alpha_{SH})$ can only be an equilibrium if $\alpha_{SH} = 0$. It follows that if $\theta^* \in (y(\alpha_{SH}), y(-\alpha_M))$, then $\alpha_M = \alpha_{SH} = 0$, and hence the interval $(y(\alpha_{SH}), y(-\alpha_M))$ is empty, which is a contradiction.

Therefore, either $\theta^* \geq y(-\alpha_M)$ or $\theta^* \leq y(\alpha_{SH})$. Suppose first that $\theta^* \geq y(-\alpha_M)$. Then either 1) $y(-\alpha_M) > y(\alpha_{SH})$, in which case $\theta^* > y(\alpha_{SH})$ and hence $\alpha_{SH} = 0$, or 2) $y(-\alpha_M) = y(\alpha_{SH})$, which requires $\alpha_M = \alpha_{SH} = 0$. In any case, $\alpha_{SH} = 0$ and $\theta^* \geq y(0)$. Since (IA.19) must be satisfied as an equality for θ^*, (IA.20) implies

$$\Delta(\theta^*) = \delta(1-\delta)(1-2\tau_j) \times g(s_{SH}, \theta^*).$$

Suppose next that $\theta^* \leq y(\alpha_{SH})$. Then either 1) $y(-\alpha_M) > y(\alpha_{SH})$, in which case $\theta^* < y(-\alpha_M)$ and hence $\alpha_M = 0$, or 2) $y(-\alpha_M) = y(\alpha_{SH})$, which requires $\alpha_M = \alpha_{SH} = 0$. In any case, $\alpha_M = 0$ and $\theta^* \leq y(0)$. Since (IA.19) must be satisfied as an equality for θ^*, (IA.22) implies

$$\Delta(\theta^*) = \delta(1-\delta)(1-2\tau_j) \times g(s_M, \theta^*).$$

Combining the two cases together, we get

$$\Delta(\theta^*) = \delta(1-\delta)(1-2\tau_j) \times \max\{g(s_{SH}, \theta^*), g(s_M, \theta^*)\}.$$

Since in equilibrium $\tau_j = 1 - F(\theta^*)$, any equilibrium threshold θ^* satisfies (IA.14). In addition, by construction, if (IA.14) is satisfied, then an equilibrium with θ^* exists, which proves the first statement of the proposition. Since the right-hand side of (IA.14) is continuous and bounded, an equilibrium always exists.

The argument above proves that $\alpha_{SH} = 0$ if $\theta^* \geq y(0)$ and $\alpha_M = 0$ if $\theta^* \leq y(0)$. Thus, to prove the proposition, it remains to prove that α_{SH} decreases with θ^* and α_M increases with θ^*. Recall that if $\theta^* \leq y(0)$, $\alpha_{SH} = \arg\max_{\alpha_{SH} \geq 0} H(\alpha_{SH}, \theta^*)$, where

$$H(\alpha_{SH}, \theta^*) = \begin{cases} \frac{1}{1-F(\theta^*)}(1 - F(y(\alpha_{SH})))(\hat{g}_{SH} - \rho_{SH}\alpha_{SH}) & \text{if } \theta^* < y(\alpha_{SH}) \Leftrightarrow \alpha_{SH} < y^{-1}(\theta^*) \\ \hat{g}_{SH} - \rho_{SH}\alpha_{SH} & \text{if } \theta^* \geq y(\alpha_{SH}) \Leftrightarrow \alpha_{SH} \geq y^{-1}(\theta^*). \end{cases}$$

Denote by $h(\alpha_{SH}) = (1 - F(y(\alpha_{SH})))(\hat{g}_{SH} - \rho_{SH}\alpha_{SH})$. In the region $\alpha_{SH} \geq y^{-1}(\theta^*)$, we have $H(\alpha_{SH}, \theta^*) = \hat{g}_{SH} - \rho_{SH}\alpha_{SH}$, and thus $H(\alpha_{SH}, \theta^*)$ decreases in α_{SH}. Hence, the maximum of $H(\alpha_{SH},\theta^*)$ is achieved on $\alpha_{SH} \in [0, y^{-1}(\theta^*)]$ and coincides with the maximum of the function $h(\alpha_{SH})$, which does not depend on θ^*. For any function $h(\alpha)$, $\arg\max_{[0,y]} h(\alpha)$ is weakly increasing in y. Hence, $\arg\max_{[0,y^{-1}(\theta^*)]} h(\alpha_{SH})$ is weakly increasing in $y^{-1}(\theta^*)$ and thus weakly decreasing in θ^*. Overall, $\alpha_{SH}(\theta^*)$ is nonnegative, is weakly decreasing in θ^* for $\theta^* \leq y(0)$, and equals zero for $\theta^* \geq y(0)$. This proves that $\alpha_{SH}(\theta^*)$ is everywhere weakly decreasing in θ^*.

Similarly, $\alpha_M(\theta^*) = 0$ if $\theta^* \leq y(0)$ and $\alpha_M(\theta^*) = \arg\max_{\alpha_M \geq 0} \tilde{H}(\alpha_M, \theta^*)$ if $\theta^* \geq y(0)$, where

$$\tilde{H}(\alpha_M,\theta^*) = \begin{cases} \hat{g}_M - \rho_M \alpha_M & \text{if } \theta^* < y(-\alpha_M) \Leftrightarrow \alpha_M > -y^{-1}(\theta^*) \\ \frac{1}{F(\theta^*)} F(y(-\alpha_M))(\hat{g}_M - \rho_M \alpha_M) & \text{if } \theta^* \geq y(-\alpha_M) \Leftrightarrow \alpha_M \leq -y^{-1}(\theta^*). \end{cases}$$

In the region $\alpha_M > -y^{-1}(\theta^*)$, the function is decreasing and hence the maximum is achieved on $[0, -y^{-1}(\theta^*)]$ and coincides with the maximum of $F(y(-\alpha_M))(\hat{g}_M - \rho_M \alpha_M)$. Thus, $\alpha_M(\theta^*)$ is weakly increasing in $-y^{-1}(\theta^*)$ and hence is weakly increasing in θ^*, as required. Since $\alpha_M(\theta^*)$ is nonnegative and equals zero for $\theta^* \leq y(0)$, it is everywhere weakly increasing in θ^*.

Sufficient condition for threshold equilibria. We conclude by proving that the condition $\frac{\partial}{\partial\theta}\Delta(\theta) > \max\left\{\frac{\partial}{\partial\theta} g(s_{SH},\theta), -\frac{\partial}{\partial\theta} g(s_M,\theta)\right\}$ ensures that any equilibrium takes a threshold form. As shown above, the director of firm i gives control to shareholders if and only if (IA.19) is satisfied. Denote the right-hand side of (IA.19) by $J(\theta_i)$. To prove that any equilibrium is a threshold equilibrium, it is sufficient to show that the function $\Delta(\cdot)$, which is increasing, crosses the function $J(\cdot)$ at exactly one point θ_i^*, and that $\Delta'(\theta_i^*) > J'(\theta_i^*)$. This would imply that $\Delta(\theta_i) > J(\theta_i)$ if and only if $\theta_i > \theta_i^*$, that is, that the director follows a threshold strategy with a threshold θ_i^*.

First, since $J(\theta)$ is bounded, $\Delta(\theta) > (<)\, J(\theta)$ for a sufficiently large (small) θ, and hence at least one intersection point always exists. Note also that $J(\theta)$ is always decreasing in the range $\theta \in (y(\alpha_{SH}), y(-\alpha_M))$. There are two cases: $\tau_j < 0.5$ and $\tau_j > 0.5$. We only present the proof for the case $\tau_j < 0.5$; the proof for $\tau_j > 0.5$ is similar.

If $\tau_j < 0.5$, $J(\theta)$ is decreasing in the ranges $\theta \leq y(\alpha_{SH})$ and $\theta \in (y(\alpha_{SH}), y(-\alpha_M))$, and is increasing in the range $\theta \geq y(-\alpha_M)$. For $\theta > y(-\alpha_M)$, $J'(\theta) = \delta(1-\delta)(1-2\tau_j)\frac{\partial g(s_{SH},\theta)}{\partial\theta} < \frac{\partial g(s_{SH},\theta)}{\partial\theta} < \Delta'(\theta)$. There are two cases. First, suppose $\Delta(y(-\alpha_M)) \geq J(y(-\alpha_M))$. Since $\Delta'(\theta) > J'(\theta)$ for $\theta > y(-\alpha_M)$, we have $\Delta(\theta) > J(\theta)$ for all $\theta > y(-\alpha_M)$. For $\theta \leq y(-\alpha_M)$, Δ is increasing, J is decreasing, and $\Delta(\theta) < J(\theta)$ for a sufficiently small θ. Hence, there exists a unique intersection point $\theta_i^* \leq y(-\alpha_M)$, and $\Delta'(\theta_i^*) > 0 > J'(\theta_i^*)$. Second, suppose

$\Delta(y(-\alpha_M)) < J(y(-\alpha_M))$. Since for $\theta \le y(-\alpha_M)$, Δ is increasing, and J is decreasing, the two functions do not intersect in that range. Since $\Delta(\theta) > J(\theta)$ for a sufficiently large θ, there exists at least one intersection point for $\theta > y(-\alpha_M)$. In addition, since $\Delta'(\theta) > J'(\theta)$ for $\theta > y(-\alpha_M)$, this intersection point is unique and $\Delta'(\theta) > J'(\theta)$ at this point.

Note also that $J(\theta_i)$ is decreasing in τ_j and hence increasing in θ_j^*, the threshold of the director in the other firm. The best response of director i, $\theta_i^*\left(\theta_j^*\right)$, is given by the intersection point of $\Delta(\theta_i)$ and $J(\theta_i)$. Since $\Delta'(\theta_i)|_{\theta_i=\theta_i^*\left(\theta_j^*\right)} > J'(\theta_i)|_{\theta_i=\theta_i^*\left(\theta_j^*\right)}$, then, as J increases with θ_j^*, the best response threshold $\theta_i^*\left(\theta_j^*\right)$ increases as well. Hence, the game exhibits strategic complementarity, as in the basic model.

Proof of Lemma IA.8: Consider an equilibrium with a threshold θ^*. Let $(\alpha_{SH}(\theta^*), \alpha_M(\theta^*))$ characterize the compensation contract when an incumbent director is hired (given by Proposition IA.3), and let $(\alpha_{SH}^{out}, \alpha_M^{out})$ characterize the compensation contract when the outside candidate is hired. When contingent compensation contracts are allowed, the expected third-stage aggregate utility of shareholders and managers (not including the cost of directors' compensation) is given by

$$\begin{aligned}
&U_{Contracts} = (1-\delta)^2 \times 2\left[F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH}\right] \\
&+\delta^2 \times 2\left[\begin{array}{c} F(\theta^*)\left(F\left(y\left(-\alpha_M^{out}\right)\right)\hat{g}_M + \left(1-F\left(y\left(-\alpha_M^{out}\right)\right)\right)\hat{g}_{SH}\right) \\ +(1-F(\theta^*))\left(F\left(y\left(\alpha_{SH}^{out}\right)\right)\hat{g}_M + \left(1-F\left(y\left(\alpha_{SH}^{out}\right)\right)\right)\hat{g}_{SH}\right) \end{array}\right] \\
&+2\delta(1-\delta) \times \left[F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH} + \right. \\
&+(1-F(\theta^*))^2\left[\Pr(\theta < y(\alpha_{SH}(\theta^*))|\theta > \theta^*)\hat{g}_M + \Pr(\theta > y(\alpha_{SH}(\theta^*))|\theta > \theta^*)\hat{g}_{SH}\right] \\
&+F(\theta^*)^2\left[\Pr(\theta < y(-\alpha_M(\theta^*))|\theta < \theta^*)\hat{g}_M + \Pr(\theta > y(-\alpha_M(\theta^*))|\theta < \theta^*)\hat{g}_{SH}\right] \\
&+F(\theta^*)(1-F(\theta^*))\left[F\left(y\left(-\alpha_M^{out}\right)\right)\hat{g}_M + \left(1-F\left(y\left(-\alpha_M^{out}\right)\right)\right)\hat{g}_{SH}\right] \\
&\left.+F(\theta^*)(1-F(\theta^*))\left[F\left(y\left(\alpha_{SH}^{out}\right)\right)\hat{g}_M + \left(1-F\left(y\left(\alpha_{SH}^{out}\right)\right)\right)\hat{g}_{SH}\right]\right].
\end{aligned}$$

The first component represents the case in which none of the directors resigns, in which case directors choose s_{SH} if and only if their type exceeds $y(0)$. The second component represents the case in which both directors resign, and hence both firms have to hire outside candidates. The last component represents the case in which only one director resigns.

Similarly, when contingent compensation contracts are not allowed, and hence $\alpha_{SH}(\theta^*) = \alpha_M(\theta^*) = 0 = \alpha_M^{out} = \alpha_{SH}^{out}$, the expected third-stage aggregate utility of shareholders and managers is given by

$$U_{NoContracts} = (1-\delta)^2 \times 2\left[F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH}\right]$$
$$+\delta^2 \times 2\begin{bmatrix} F(\theta^*)(F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH}) \\ +(1-F(\theta^*))(F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH}) \end{bmatrix}$$
$$+2\delta(1-\delta)\times\begin{bmatrix} F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH} + \\ +(1-F(\theta^*))^2\left[\Pr(\theta < y(0)\,|\theta > \theta^*)\hat{g}_M + \Pr(\theta > y(0)\,|\theta > \theta^*)\hat{g}_{SH}\right] \\ +F(\theta^*)^2\left[\Pr(\theta < y(0)\,|\theta < \theta^*)\hat{g}_M + \Pr(\theta > y(0)\,|\theta < \theta^*)\hat{g}_{SH}\right] \\ +F(\theta^*)(1-F(\theta^*))\left[F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH}\right] \\ +F(\theta^*)(1-F(\theta^*))\left[F(y(0))\hat{g}_M + (1-F(y(0)))\hat{g}_{SH}\right] \end{bmatrix}.$$

Hence, the difference in welfare between the two cases is given by

$$U_{Contracts} - U_{NoContracts} = 2\delta^2(\hat{g}_{SH} - \hat{g}_M)\begin{bmatrix} F(\theta^*)(F(y(0)) - F(y(-\alpha_M^{out})) \\ +(1-F(\theta^*))(F(y(0)) - F(y(\alpha_{SH}^{out}))) \end{bmatrix}$$
$$+2\delta(1-\delta)\times$$
$$\begin{bmatrix} (1-F(\theta^*))\begin{bmatrix} \Pr(\theta < y(\alpha_{SH}(\theta^*)), \theta > \theta^*)\hat{g}_M + \Pr(\theta > y(\alpha_{SH}(\theta^*)), \theta > \theta^*)\hat{g}_{SH} \\ -\Pr(\theta < y(0), \theta > \theta^*)\hat{g}_M - \Pr(\theta > y(0), \theta > \theta^*)\hat{g}_{SH} \end{bmatrix} \\ +F(\theta^*)\begin{bmatrix} \Pr(\theta < y(-\alpha_M(\theta^*)), \theta < \theta^*)\hat{g}_M + \Pr(\theta > y(-\alpha_M(\theta^*)), \theta < \theta^*)\hat{g}_{SH} \\ -\Pr(\theta < y(0), \theta < \theta^*)\hat{g}_M - \Pr(\theta > y(0), \theta < \theta^*)\hat{g}_{SH} \end{bmatrix} \\ +F(\theta^*)(1-F(\theta^*))(\hat{g}_{SH} - \hat{g}_M)\left[F(y(0)) - F(y(-\alpha_M^{out})) + F(y(0)) - F(y(\alpha_{SH}^{out}))\right] \end{bmatrix}. \tag{IA.23}$$

Consider the limit of $U_{Contracts} - U_{NoContracts}$ as $\theta^* \to \infty$. Then $\theta^* > y(0)$ and hence, according to the proof of Proposition IA.3, $\alpha_{SH}(\theta^*) = 0$ and $\theta^* > y(-\alpha_M(\theta^*))$. In addition, $\lim_{\theta^* \to \infty} \alpha_M(\theta^*) = \alpha_M^{out}$. Using (IA.23),

$$\lim_{\theta^* \to \infty}(U_{Contracts} - U_{NoContracts}) = 2\delta(\hat{g}_{SH} - \hat{g}_M)\left[F(y(0)) - F\left(y\left(-\alpha_M^{out}\right)\right)\right].$$

Since $0 > -\alpha_M^{out}$ and $y(\cdot)$ is decreasing, $\lim_{\theta^* \to \infty}(U_{Contracts} - U_{NoContracts}) < 0$. Thus, there exists $\bar{\theta}$ such that for $\theta^* > \bar{\theta}$, $U_{Contracts} - U_{NoContracts} < 0$. Note that if we account for the cost of directors' compensation to managers and shareholders, then $U_{Contracts}$ will weakly decrease and $U_{NoContracts}$ will not change. Therefore, $\lim_{\theta^* \to \infty}(U_{Contracts} - U_{NoContracts}) < 0$ holds even when directors' compensation is accounted for.

References

Adams, Renée B., and Daniel Ferreira, 2007, A theory of friendly boards, *Journal of Finance* 62, 217-250.

Grossman, Sanford, and Oliver Hart, 1986, The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* 94, 691-719.

Harris, Milton, and Artur Raviv, 2008, A theory of board control and size, *Review of Financial Studies* 21, 1797-1832.

Hart, Oliver, and John Moore, 1990, Property rights and the theory of the firm, *Journal of Political Economy* 98, 1119-1158.