Better to Mix than to Match: Board Composition and Firm Size *

Pierre Jinghong Liang
Tepper School of Business
Carnegie Mellon University
Tech and Frew Streets
Pittsburgh, PA 15213-3890

Madhav V. Rajan
Graduate School of Business
Stanford University
518 Memorial Way
Stanford, CA 94305-5015

Korok Ray
McDonough School of Business
Georgetown University
37th and O Streets, NW
Washington, DC 20057

Abstract

Boards of directors serve several functions, chief among them the monitoring of the firm and management. We model the monitoring activities of directors, who themselves face a moral hazard problem and need incentives to avoid shirking. The firm simultaneously sets wage incentives for all employees (directors and workers), and also chooses the size of the firm. Larger firms are more productive, but costlier to monitor. Our main result shows that directors that are more dissimilar to each other will allow firm profits to rise and the firm size to grow. Similar directors have redundant signals on firm performance and are therefore worse monitors than dissimilar directors. This creates a causal relationship between board composition and firm size, which can be tested with data.

*We’d like to thank Rohan Williamson, Lee Pinkowitz, Prem Jain for helpful conversations. The Capital Markets Research Center at Georgetown University provided financial support.
1 Introduction

Ever since the passage of the Sarbanes-Oxley of 2002, the composition of corporate boards has become central to corporate governance theory and practice. The primary vehicle in determining board composition is the independence of directors on corporate boards. In particular, the NASDAQ and NYSE proposed listing requirements states that a majority of board members must be independent directors, where an independent director is one who has no “material relationship” with the company, either directly “as a partner, shareholder, or officer of an organization that has a relationship with the company.”\footnote{An independent director or a member of his family cannot receive more than $100,000 in compensation from the company, other than director compensation. A current employee can never be an independent director, nor can a director or member of his family be affiliated with an executive officer for a company whose compensation committee employees an executive for the listed company. Nor can the director or member of his family be an executive officer for a company that accounts for 2 or $1 million of revenue of the listed company nor for a company for which the listed company accounts for 2 or $1 million of revenue.} This highly abstract and imprecise definition of independence has spawned a growing empirical literature testing the effects of board independence on firm performance (see Bhagat and Black (2002) and Hermalin and Weisbach (2000) for comprehensive surveys). But the regulatory definition of independence and the subsequent academic empirical work leave open the question of optimal board composition, namely, how similar directors should be to one another. We propose a theory of board composition and develop predications that can be tested against data. Our main result is that boards with directors who are dissimilar to each other will preside over more profitable and larger firms. Thus, it is better to mix than to match on corporate boards.

Boards of directors have many functions, but we focus on the monitoring activity of directors. We advance an agency model where a board oversees a team of workers who constitute the firm. Workers are risk averse and exert costly effort in producing output. Directors themselves face a moral hazard problem, in that they require incentives in order to exert costly effort to monitor the firm. The principal (the shareholders) simultaneously selects contracts for the directors and the workers, and also selects the size of the firm. Larger firms are more productive, but they are also harder to manage and require more monitoring effort by directors.

Directors receive signals on the performance of the firm. Directors who are similar (dissimilar) to each other receive correlated (uncorrelated) signals on firm output. We show that dissimilar directors are more efficient monitors of the firm. They effectively have orthogonal information sets, which allows them to better observe the true output of the firm, and in so doing, to reduce the volatility in profits. Because all agents are risk averse and dislike volatility, this ex-ante reduction in volatility from dissimilar directors allows the principal to increase the size of the firm. Ordinarily, increasing the size of the firm would make it harder to monitor and
manage, but the increased efficiency from better monitoring through the negative correlation in the directors’ signals compensates for this. Thus, this allows the firm to expand in size and in profits.

We also develop supplementary results on when it makes sense to divide a board into committees. Directors do not work as a single, homogenous group, but often their work takes place in specialized committees. And because boards are highly constrained in time and resources, directors often serve on multiple committees, and committees collaborate on the key monitoring function of the board. We model this complementary between committees and provide conditions under which it is optimal for the firm to split its board into specialized committees. In particular, when the price of the firm’s product is sufficiently high and there is substantial complementary between the committees, then it is optimal to organize the board according to specialized committees.

Our paper aims to provide a theory of board composition and to deliver novel empirical results between board composition and firm size. To our knowledge, this relationship has not been established before. The majority of the empirical papers examine the relationship between board independence and firm performance, where the definition of independence follows the narrow criteria of the Sarbanes-Oxley Act and the NYSE and NASDAQ listing requirements. And while the existing empirical literature may include firm size as a control in their regressions, such papers never establish a causal relationship between board composition and firm size. Like much of the corporate governance literature, the empirics on board independence are plagued by endogeneity problems, and it is not clear whether board independence drives firm performance, or the reverse. We seek to clarify this confusion by providing an explicit, causal relationship. Boards that are composed of directors who are different from one another are more effective monitors, and as such, this leads to higher profits and larger firm size.

There is an emerging literature seeking to understand the nuances of board composition. Our aim in this paper is not only to address existing empirical work, but also to stimulate the creation of new data sets that can be used to test our propositions. Our view is that the majority of empirical work suffers from a preoccupation with the legalistic definition of independence by government regulators, rather than examining qualities of boards of directors at a more granular level. That is our aim.

With that caveat, we now review the existing literature on board composition. The majority of the literature is empirical and focuses on the relationship between board independence and firm profitability. While there is still disagreement on whether independent boards leads to higher or lower firm profitability, the consensus is that independence (as defined by regulators), does not lead to higher performance.

\footnote{Fracassi, (2008) uses a novel data set to examine the social networks of corporate directors.}
Several major works indicate that director independence and firm performance are not correlated. Bhagat and Black (1999, 2000, 2002) argue that companies do not perform better with an independent board as measured by profitability. Bhagat and Black (2000) provide a comprehensive review of the literature on the subject, concluding that most studies support this lack of correlation. Hermalin and Weisbach (2000) agree that board composition does not affect company performance directly, but that more independent boards are correlated with higher quality decisions by boards in matters including CEO replacement, company acquisitions, and executive compensation. Hanson and Moon (2000) find that increasing the proportion of outside directors neither increases firm performance nor ensures that the firm does not commit illegal acts. Klein (1998) supports the lack of correlation between board composition and firm performance when the entire board is considered, but finds that increasing inside directors on certain key committees makes firms perform better. Kumar and Sivaramakrishnan (2008) argue instead that a higher proportion of independent directors can cause outside shareholders’ value to decrease because independent directors are actually more dependent on the CEO and are more likely to avoid effort. Agrawal and Knoeber (1996) agree, showing that independent directors are one of several factors that can cause reduced firm performance.

There is some support in the literature for the opposite position, that independence and performance are correlated. Gordon (2006) traces the rise of independent directors in firms in the last 50 years, arguing that independent directors perform better than insiders because they are not as influenced by managers and are more swayed by legal standards calling for more accurate disclosure. Rosenstein and Wyatt (1990) argue that outside directors are correlated with positive share-price reactions performance but that they are not themselves more valuable than other directors. Lawrence and Caylor (2004) find that although there is no Tobins Q correlation, independent boards have higher returns on equity, higher profit margins, larger dividend yields, and larger stock repurchases. Several papers address the effects of board composition on corporate mergers. Byrd and Hickman (1992) suggest that firms with more outside directors have higher abnormal returns on announcement-date. V. Subrahmanyam, N. Rangan and S. Rosenstein (1997) address the connection between board composition and bank acquisitions, specifically on abnormal returns. Cotter, Shivdasani and Zenner (1997) find that board independence is correlated with higher initial tender offer premium, bid premium revision, and the target shareholder gains over the entire tender offer period.

2 The Base Model With A Single Director

Consider a principal contracting with \( n \) workers. Let \( N = \{1, 2, \ldots, n\} \) denote the set of workers. The set of workers denotes the entire firm; the firm consists of these workers, and \( n \) is the size
of the firm. Each worker exerts effort $e_i$ at cost of effort $C_i(e_i) = 0.5c_i e_i^2$. The cost of effort parameter $c_i$ measures the quality/ability of the worker: higher quality workers have a lower $c_i$. The principal is the residual claimant of the firm, namely, the shareholders. Output of the firm is:

$$x = \sum_{i \in N} e_i + \varepsilon_x.$$ 

We assume that $x$ is either unverifiable or is realized too late to be used in contracting. The principal employs a director to monitor the firm.\footnote{Some may argue that directors monitor management, not the entire firm. It is possible to interpret $n$ as the size of the management firm, but we prefer to interpret it as the size of the firm, because the workers are producing output. The main point is that the directors monitor the output-producing agents in the firm. In that sense, we do not draw a distinction between management and labor, and instead maintain the assumption that management as well as labor produce output for the firm.} We model the director as an effort- and risk-averse economic agent who can produce a signal about the collective efforts of the firm under his supervision. In particular, the director exerts monitoring effort $m$ at personal cost $C(m) = km$, where $k$ is the director’s cost of effort parameter. This parameter $k$ measures the director’s quality/ability: higher quality (or lower cost) directors have a lower $k$.

The firm does not contract on $x$, but rather on $y$. Assume $y$ is the only variable on which contracts can be written. We specify that:

$$y = \sum_{i \in N} e_i + \epsilon, \quad \text{where } \epsilon \sim N\left(0, \frac{G(n)\sigma^2}{m}\right).$$

We can think of $y$ as the accounting profit of the firm.\footnote{We do not model the capital market and therefore do not take a position on the relationship between accounting profits and stock returns. Under market efficiency, profit returns should move closely together.} To be consistent with the managerial accounting literature, we will sometimes call this the firm’s performance measure. Note that $\sigma$ measures the uncertainty in performance measurement. Observe that the directors’ effort reduces the variance on the firm’s output measure. In particular, as the director works harder, he gets a better signal of the firm’s true performance. This captures the moral hazard of the director. His effort $m$ is not contractable, so the firm must give him incentives on the observable measure $y$. The function $G$ captures the difficulties of monitoring that are specific to the firm or industry. We normalize $G(0)$ to equal 0, and assume that $G(n)$ is increasing and strictly convex in $n$, so that measurement becomes noisier as the firm size grows. This reflects a basic assumption that the performance of larger firms are harder to measure because of problems with coordination, communication difficulties, etc. In addition, we assume that
$G(\cdot)$ satisfies the following technical condition:\(^5\)

$$\frac{d}{dn} \left( \frac{n[G'(n)]}{G(n)} \right) \geq 0.$$

(1)

The director’s effort reduces the variance of $y$ while the worker’s effort increases its mean.\(^6\)

The firm offers a linear contract to each worker, so each worker receives a wage $w_i = a_i + b_i y$. All $n$ workers are equally risk averse with exponential utility and a common coefficient of risk aversion parameter $r$. Thus, workers’ preferences assume a mean-variance representation. In certainty equivalent terms, each worker receives $Ew_i - \frac{r}{2} \text{Var}(w_i) - C_i(e_i)$: expected wages less a risk premium less the cost of effort. Each worker solves

$$\max_{e_i} a_i + b_i E_y - \frac{r}{2} b_i^2 \frac{G(n)\sigma^2}{m} - C_i(e_i),$$

yielding the standard incentive constraint, $e_i = b_i / c_i$.\(^7\) Suppose all workers have outside options, denoted by $\bar{u}$, which is normalized to zero. The firm will set the salaries $(a_i)$ such that each worker’s individual rationality constraint binds, given that all other workers select their equilibrium effort levels.

Assume $m$ is unobservable so the director must be motivated to supply monitoring efforts. The firm offers a linear contract to the director consisting of a salary and a bonus on the firm’s output, i.e., $\omega = \alpha + \beta y$. The director thus solves:

$$\max_m \alpha + \beta E_y - \frac{r}{2} \frac{\beta^2 G(n)\sigma^2}{m} - C(m).$$

The first-order condition yields the incentive constraint for the director:

$$m^* = \beta \sigma \left( \frac{G(n)r}{2k} \right)^{1/2}.$$

(2)

The director’s problem is concave even when his personal cost $C(m)$ is linear because his personal benefit of monitoring effort is increasing and concave in $m$.\(^8\) Even though all workers

---

\(^5\)This condition is not too restrictive, and is satisfied by a wide range of functional forms. In particular, it holds for all power functions ($G(n) = An^\gamma$, $A > 0$, $\gamma > 1$), as well as any general polynomial function with positive coefficients. It is also satisfied by any function in the exponential class ($G(n) = A(e^n - 1)$, $A > 0$). Finally, note that the condition is implied by, and is therefore weaker than, a requirement that $G(n)$ be weakly log-convex.

\(^6\)By this modeling choice, we choose to emphasize the monitoring role of the director, leaving their other functions (such as fulfilling some direct advisory tasks) as second-order effects. Within this monitoring role, the director can, for example, exert effort to coordinate multiple agents within the firm or facilitate communication between such agents. As a reduced form, we model this as directly reducing the variance on the firm performance measure, as opposed to explicitly framing the exact consequence of the managerial action, in order to keep the level of complexity manageable.

\(^7\)Notice that for each worker, the effort choice is also a dominant strategy response in the workers’ subgame.

\(^8\)A benefit of linear personal cost is that it provides tractability. Other functional forms of $C(m)$ provide similar economic intuition, but result in mathematical expressions of far greater complexity.
and the director are under linear contracts, the provision of incentives to exert effort comes through in different ways. The director’s effort does not affect the expectation of his wage $E\omega$, but such effort does decrease its variance. Given the director is risk-averse, he has an incentive to reduce the variance.

Assume each unit of output $(x)$ is sold in a competitive market at an exogenous price $q$. The firm maximizes the expectation of revenues less wage payments, and substitutes the binding (IR) constraints into its optimization. The firm sets the salary levels $(a_i, \alpha)$ such that the individual rationality constraint binds for every worker $f$ and the director. Assume the opportunity wage for the workers and the director are $\bar{u}_M$, which is also normalized to zero. The firm therefore maximizes total surplus,

$$\max_{b_i, \beta} \sum_{i \in \mathcal{N}} \left[ qe_i - C_i(e_i) - \frac{r}{2} \text{Var}(w_i) \right] - C(m) - \frac{r}{2} \text{Var}(\omega),$$

where

$$\text{Var}(w_i) = b_i^2 G(n)\frac{\sigma^2}{m} \quad \text{and} \quad \text{Var}(\omega) = \beta^2 G(n)\frac{\sigma^2}{m}.$$

Substituting in each worker’s incentive constraint gives

$$\max_{\beta, b_i} \sum_{i \in \mathcal{N}} \left[ qb_i - \frac{b_i^2}{2c_i} - \frac{r}{2} b_i G(n)\frac{\sigma^2}{m^*} \right] - km^* - \frac{r}{2} \beta^2 G(n)\frac{\sigma^2}{m^*},$$

where $m^*$ is given by (2). The first-order conditions yield the optimal incentives for the worker:

$$b_i^*(n) = \frac{q}{1 + \frac{rc_iG(n)\sigma^2}{m^*}}.$$

Not surprisingly, an increase in risk aversion $(r)$, cost of effort $(c_i)$, or noise $(\sigma^2)$ causes the firm to cut the worker’s incentives, delivering the standard trade-off. What is different now is that the director’s effort $(m)$ matters in choosing optimal incentives for workers. Of course, the firm affects $m$ by choosing the director’s incentive $\beta$ carefully. Thus, the incentives for the worker and the director have opposite effects on the risk premium: worker incentives raise the risk premium, while director incentives lower the risk premium. The two incentives move together in the same direction, because they are the effective response to the firm eliciting more effort out of the workers. In this sense, the incentives of the director serve as a countervailing force on the risk premium, relative to the incentives for the workers.

Optimizing the firm’s objective function with respect to $\beta$ results in a first-order condition, from which we can see the principal’s tradeoff in choosing the director’s incentive:

$$\left[ \frac{r}{2} \frac{G(n)\sigma^2}{m^*} \sum_{i=1}^{n} b_i^2 - k + \frac{r}{2} \frac{G(n)\sigma^2}{m^*} \beta \right] \frac{\partial}{\partial \beta} m^* - r \beta \frac{G(n)\sigma^2}{m^*} = 0.$$
Providing more managerial incentive (i.e., increasing $\beta$) induces more managerial effort (see (2)). The benefit is two-fold. First, it reduces the variance of the $y$ and (thus) the risk-premium needed for the director, because the director’s contract is written on $y$. Second, increasing $\beta$ also reduces the needed risk-premium of all workers because the worker contracts are also written on $y$. Notice here that the firm exploits a spill-over effect from the monitoring effort. The director is self-interested and desires to exert effort to reduce the variation in his own compensation. But in doing so, he reduces the wage variances of all workers (e.g., a spill-over effect). This is beneficial to the firm because it reduces salary costs ($a_i^*$ and $\alpha^*$). These two (marginal) benefits are shown as the two positive terms in (4). The cost of providing managerial incentives is also two-fold. First, a higher incentive leads to more personally costly monitoring effort, for which the director requires compensation. Second, a higher incentive leads to a managerial pay more sensitive to the signal variations, which increases the risk-premium. These two (marginal) costs are shown as the two negative terms in (4).9

Substituting in the director’s incentive constraint (2) into (4), and rearranging terms yields

$$\beta = \left( \frac{1}{2} \sum_{i \in N} b_i^2 \right)^{1/2}.$$  

The director’s incentives thus increase with each worker’s incentives. Intuitively, as the firm increases $b_i$, it induces more effort from the workers. The firm increases $\beta$ which induces the director to spend more effort monitoring the workers; in turn, this reduces the risk premium for everyone. But these larger incentives amplify output risk, which workers dislike because they are risk averse. Thus, the firm must compensate the workers for this additional risk and this effectively lifts the risk premium. To forestall these expensive risk payments, the firm compensates by increasing incentives for the director.

Combining the previous expressions for $b_i$, $\beta$, and the director’s incentive constraint yields the following expression for $b_i^*$ as an implicit function of the model parameters:

$$b_i^*(n) = \frac{q}{1 + 2c_1 \sqrt{\frac{\sigma G(n) a^2 k}{\sum_{j \in N} b_j^2}}}.$$

Consider a slight alteration of the model where the director exerts a productive effort, say $e_m$, which, similar to workers’ efforts $e_i$, increases the output and the aggregate performance measure $y$. We can interpret this productive effort as the director advising management, which leads to higher production. When choosing the optimal $\beta$, the principal now must consider the marginal benefit and cost of inducing $e_m$, in addition to the four effects consider in equation (4). However, so long as these additional marginal effects are well-behaved (i.e., smooth and bounded), their only impact is to make the optimal $\beta$ quantitatively different. The fundamental tradeoff in the managerial monitoring activities, which is the focus of our paper, persists and leads to the same qualitative results.
Despite the fact that equation (5) is an implicit function of \( b_i \), it can be shown that, as in the setting with a mechanical monitor, the strength of incentives \( (b_i) \) falls as the cost of effort \( (c_i) \) increases.

### 2.1 Optimal Incentives and Firm Size

Assuming identical workers \( (c_i = c) \) leads to \( b_i^* = b^* \), which in turn allows us to solve (5) and derive a closed-form solution for the equilibrium worker incentive for a given \( n \):

\[
b^*(n) = q - 2c\sigma\sqrt{rkg(n)/n}.
\]

(6)

Recall that the equilibrium director incentive \( (\beta^*) \), monitoring effort \( (m^*) \), and worker effort \( (e^*) \) are all functions of \( b^* \).\(^{10}\) From (3), we can now write the profit function with a director for a given \( n \), denoted by \( \Pi(n) \), as follows:

\[
\Pi(n) = n \left[ \frac{qb^*}{c} - \frac{b^*^2}{2c} - \frac{r}{2} b^*^2 \frac{G(n)\sigma^2}{\sqrt{rnG(n)/k}} \right] - 2k \frac{b^*\sigma}{2} \sqrt{\frac{rnG(n)}{k}} = \frac{n (b^*(n))^2}{2c}.
\]

(7)

Also, observe that firm size \( n \) moves inversely with the workers’ incentives. As the size of the firm grows, the firm becomes harder to monitor and the performance measure becomes more noisy (recall that \( G \) increases in \( n \)). Because workers are risk averse, they dislike this increase in volatility. To compensate, the principal reduces the worker’s incentive and in so doing, lowers the risk premium. Of course, note that there is a simultaneous effect of firm size on the monitoring effort, as mentioned earlier. As firm size grows and the risk premium rises, the director himself works harder at monitoring in order to reduce this risk premium, since the director himself is risk averse and dislikes volatility in output. Therefore, the director works harder. Therefore, there are two effects of firm size on worker incentives: a direct effect through increasing the risk premium via \( G(n) \) and an indirect effect through the director’s efforts. A comparison of these two effects will determine the ultimate effect of the workers’ incentives. Now we characterize the optimal \( n \) in this setting.

**Lemma 1** Assume \( c_i = c \) for all \( i \in N \). There exists a unique optimal firm size, \( n^* \), which is characterized by the following implicit function:

\[
h(n^*) + 2n^* h'(n^*) = \frac{q}{2c\sigma\sqrt{rk}},
\]

\(^{10}\)The maintained assumption here is that \( b^* > 0 \). We rule out combinations of parameters such that the above equation gives a negative \( b^* \), because that implies equilibrium worker incentives would be set to zero, effectively shutting down production.
where \( h(n) = \sqrt{\frac{G(n)}{n}} \).

Implicit differentiation of the equation defining \( n^* \) above enables us to study how the firm adjusts firm size in response to changes in the exogenous parameters of interest.

**Proposition 1** The optimal firm size is decreasing in measurement uncertainty \( \sigma \), worker risk aversion \( r \), and worker and director costs of effort, \( c \) and \( k \).

When either \( c \), \( k \), \( r \), or \( \sigma \) increases, this inflates the risk premium. As in the mechanical monitor setting, the principal responds by choosing a smaller firm, and thus deflating the risk premium somewhat. As the quality of the director improves (\( k \) shrinks), the optimal firm size grows, for two reasons. The first is the reverse of the risk-premium effect mentioned above. Second, better directors are better able to measure performance, and can thus handle larger firms.

This expression, while seemingly complex, illustrates the basic tradeoff in firm size. As the firm grows the number of workers rises and output increases. But this larger firm also becomes difficult to monitor, and increases the volatility of output. Because \( G \) is increasing and strictly convex in \( n \), eventually the cost from poor performance measurement will outweigh the gains in output. Of course, the incentives for the worker and the director endogenously vary with \( n \), making the expression for the profit as a function of \( n \) slightly complex. But the basic intuition rests between a tradeoff in firm size; larger firms are more productive but harder to monitor. This will eventually lead to a closed-form solution with the optimal firm size.

### 3 Optimal Firm Size with Two Directors

Thus far the firm hires a single director to assess the firm’s performance. A natural question is whether the owner would benefit from employing more than one supervisor for a given firm and, if so, how the choice of other organizational design instruments are adjusted accordingly. In this section, we analyze the effects of expanding the size of the firm on worker incentives as well as on the optimal firm size.

We proceed in two steps. We first extend our basic model to incorporate a second (self-interested) director who generates an additional, contractible signal regarding worker performance. We allow for the noise in this signal to be correlated with the noise in the signal produced by the original director. For an exogenous firm size, we analyze the costs and benefits of adding the second director. Subsequently, we characterize the optimal firm size when two directors are employed. We examine its relation to the correlation in the directors’ evaluations, and also compare it to the optimal firm size in the single-director setting.
This leads to an implication on firm size and board quality. In this model, the main function of the director is to monitor the firm, and the quality of the director reflects the efficiency of the monitoring activity. Better directors are better monitors and therefore they can monitor a firm of fixed size with less effort than a less able director. Said differently, a better director is a more efficient monitor, and therefore expending the same amount of resources can oversee a larger firm. As shareholders select the quality of the directors, if shareholders choose better directors, they will simultaneously allow the directors to oversee a larger, more complex firm.

### 3.1 Adding a Second Director

Suppose each director works independently to produce a separate, verifiable signal about the performance of the entire firm. With two directors, this results in a total of two contractible signals, denoted by $y_1$ and $y_2$ respectively. Of course, in practice, boards of directors do not consist of only two people, but rather many. However, our analysis is still relevant. We can simply think of the two directors as two groups of directors, each of which is compensated in the same way. And even if we consider these directors as individuals, the intuition for the two-person board generalizes for the multi-person board. As is often the case with economic models, binary outcomes are tractable and useful for generating intuition in explaining a non-binary world.

$$y_1 = \sum_{i \in N} e_i + \epsilon_1, \text{ and } y_2 = \sum_{i \in N} e_i + \epsilon_2$$

where

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{G(n)\sigma^2}{m_1} & \rho G(n)\sigma^2 \\ \rho G(n)\sigma^2 & \frac{G(n)\sigma^2}{m_2} \end{pmatrix} \right).$$

These two signals can be any two signals drawn from the wide array of measures used to track firms’ output. For example, consider a large financial institution, like a global investment bank. The boards of these institutions have several committees, with each committee possibly overseeing a different operation of the firm. The director of risk management may receive a signal on the quality of the risk management of the firm, while a director on the audit committee may observe the quality of the internal auditing. The two signals can also represent different measures of the effectiveness of the firm’s management. A director who comes from the deal-making side of investment banking may receive a signal on the number of mergers and acquisition deals that the financial institution brokered, while a director who comes from a trading background may receive a signal on the profitability of the trading book in the financial institution’s proprietary trading group. Our point is simply that there is a wide variety of performance measures for the firm, and not all directors receive the same signal. The correlation between the two signals captures the idea that the two directors monitor the
same worker environment and, as such, their errors may be correlated. Such correlation may be positive, as when the directors are subject to similar shocks in their observations of worker performance, or negative, as when biases or errors may reverse or differ across the two directors’ evaluations.\footnote{Note that the covariance of the signals is independent of either director’s effort. This specification can be recovered from the following, more detailed, construction: for director $j \in 1, 2$, signal $y_j$ is the mean of a sample drawn from a distribution of a more primitive signal or data, denoted $z_j$. At a personal cost of $km_j$ (or $k$ dollars per draw), the sample size is $m_j$. Suppose population $z_j$ is normally distributed with a population mean of $\sum_{i \in N} e_i$, variance $\text{Var}(z_j) = G(n)\sigma^2$ and $\text{Cov}(z_1, z_2) = \rho G(n)\sigma^2$. From here, it is straightforward to show that $y_j = \frac{\sum_{i \in N} e_i}{m_j} \text{ follow the joint distribution specified in the text.}$}

With the availability of a second signal, the principal modifies the contracts for the workers because the firm now has additional incentive coefficients at its disposal. In keeping with the LEN framework, we assume that the worker contracts are structured to be linear in the two performance signals. For worker $i \in N$, wages are thus given by:

$$w_i = a_i + b_{1i} y_1 + b_{2i} y_2.$$

For director $j$, $j = 1, 2$, we restrict attention to contracts that are linear in the signal produced by the director’s efforts, i.e., each director’s compensation is given by: \footnote{We do this for reasons of tractability. More generally, the owner could compensate each director on the signals generated by both directors. There is no incentive gain to doing so (because the monitoring effort only affects the variance of the director’s own signal); however, a potential second-order benefit arises via better risk-sharing because of the correlation in the signals. It can be shown that each director would be induced to work harder at monitoring if his contract were based on both signals. The reason is that the improved risk-sharing allows the principal to place a higher weight on the director’s own signal, leading to an increase in induced monitoring activity. Unfortunately, the optimal incentive weights can no longer be derived in closed-form. Nevertheless, as long as the correlation levels are not extreme, Proposition 3 continues to hold and sufficient conditions can be generated for Proposition 4 to hold as well. Moreover, after extensive numerical analysis, we have not uncovered a single instance in which either result is violated under the more general contracting arrangement. Details are available on request.}$12$

$$\omega_j = \alpha_j + \beta_j y_j.$$ 

Adding a second director introduces three separate effects to the base model with a single-director. First, there is a benefit from gaining an additional performance signal of workers, which would improve contracting. This effect strictly increases firm profit. Second, there is an incentive cost of acquiring the signal. This effect changes the endogenous monitoring activities ($m$) and strictly decreases firm profit. The third effect comes from the ability to adjust firm size after acquiring the additional director. We analyze these three effects in sequence.
3.2 Cost and Benefit of the Second Director

First, we consider the benefit of acquiring another signal. We operationalize this by considering a slightly altered setting where the second director is not subject to moral hazard. As a result, the principal is able to choose the amount of monitoring and simply pays the director his/her personal cost. We compute firm profit for this setting, denoted by \( \Pi_{II}(\rho, n) \), and compare the profit with that of the signal-director setting (i.e., \( \Pi(n) \) in equation 7).

**Proposition 2** For any fixed firm size \( n \) and all \( \rho \),

\[
\Pi_{II}(\rho, n) > \Pi(n) \quad \text{and} \quad \frac{\partial}{\partial \rho} \Pi_{II}(\rho, n) < 0.
\]

We thus find that the second signal is always beneficial and its value is decreasing in its correlation (\( \rho \)) with the first signal. This result is reminiscent of findings in Feltham and Xie (1994) and Christensen and Sabac (2002), where both signals are assumed to be available for free. We next incorporate the incentive cost of acquiring the second signal by motivating a self-interested (second) director. For each pair of values (\( \rho, n \)), denote the equilibrium profit in the setting where both directors are subject to moral hazard by \( \Pi_{II}(\rho, n) \). We provide a stark characterization of the difference in profits relative to the single-director setting, as a function of the correlation (\( \rho \)) in the directors' signals.

**Proposition 3** Assume \( c_i = c \) for all \( i \in N \), and \( k_j = k \) for all \( j = 1, 2 \). For any fixed firm size \( n \),

\[
\begin{align*}
&\text{if } \rho > 0, \quad \Pi_{II}(\rho, n) < \Pi(n); \\
&\text{if } \rho = 0, \quad \Pi_{II}(\rho, n) = \Pi(n); \\
&\text{if } \rho < 0, \quad \Pi_{II}(\rho, n) > \Pi(n).
\end{align*}
\]

In the proof, we demonstrate an even stronger result - the marginal profit from adding a second director is monotonically decreasing in the level of correlation. Together, propositions 2 and 3 illustrate the twin effects of hiring a second director. It is instructive to review the zero-correlation case (\( \rho = 0 \)) as a point of reference. When an additional informative signal (\( y_2 \)) is available for contracting, it may seem this would lead the principal to a solution closer to the “first-best.” However, hiring a second director increases the costs to the firm. In addition to the direct costs (i.e., the personal cost of monitoring and the risk-premium due to the second director), there are indirect costs of imposing the risk of the second signal on each worker whose pay depends on the signal. The principal takes both the direct and indirect costs into
account when designing contracts. When signals are uncorrelated, the benefit is exactly offset by the costs.\footnote{Further, it is easy to see how monitoring activity by the second director changes between the two settings. When there is no moral hazard with the second director (in Proposition 2), the principal requests more monitoring from the second director, which makes the second signal more precise. As a result, the principal places greater incentive weight on the second signal in the workers’ contracts. In Proposition 3, the principal induces the same monitoring effort from each director and places identical weights on the two signals in the workers’ contracts.}

Now consider the negative correlation cases. Two additional factors come into play. First, with negative correlation, the overall risk-premium is lowered for any given level of bonus rates. Second, lowered risk lowers the cost of inducing worker effort, so the principal can now increase the incentive rate for the workers to increase production. As a result, negative correlation increases firm profit. Parallel (but opposite) arguments follow in the positive correlation case.\footnote{Rajan and Sarath (1997) demonstrate the value of negatively correlated signals in a single-agent setting where the signals are available for free (i.e., no direct or indirect incentive costs). Christensen and Sabac (2002) conduct a similar analysis within the LEN framework and show that having two free signals, regardless of correlation, is always preferred to having just one signal. The key reason is that the marginal effect of increasing one incentive coefficient (such as $b_1$) on the risk-premium due to the correlation is smaller (in the scale of $\rho \sigma_1 \sigma_2$) than that on the risk-premium due to the variance (proportional to $rb_1 \sigma_1^2$). In addition, their analysis implies that profit is strictly decreasing in the level of correlation.}

As Proposition 3 demonstrates, the key to the value of having two directors is the negative correlation between the two signals, which is effectively a form of complementarity in monitoring. Such negative correlation may be a natural result of different monitoring styles brought forth by different directors. Alternatively, if we think of the noise in each director’s signal as the type of the director, then the correlation in the noise measures the similarity of the two directors. Intuitively, two directors are similar if they receive similar signals on the firm’s output. In other words, there is overlap in their information sets. Similar to a gain from diversification, the net effect is to reduce the total error (or noise) in the monitoring system.

Proposition 3 is our first main result, and states that it is better to mix than to match: boards with directors who are dissimilar from each other will preside over more profitable companies than boards with directors who are similar to each other. Matching directors (by placing similar directors on the board) is inefficient in terms of the efficiency of monitoring. Mixing directors (composing a board of dissimilar directors) is better because there is little overlap between the information sets of the directors. As such, the directors receive a cleaner signal on the firm’s performance, and this reduces the volatility in the output measure for both the workers and the directors. Because both the directors and the workers are risk averse, this reduction in volatility reduces their disutility, and therefore the firm can reduce its risk premium payments. This reduction in cost raises the profits of the firm.
This proposition gives a precise empirical proposition that can be tested against data. There is a growing body of literature on boards of directors and social networks, which constructs new data sets of the personal and professional characteristics of directors and their relationships to one another (see Fracassi (2008) on early work in this area).

3.3 Optimal Firm Size with a Second Director

The third effect of adding another director is that on firm size. As in earlier sections, we return to the question of optimal firm size. In particular, we are interested in comparing the optimal firm size in the two-director setting, which we denote by \( n_{II}^* \), to that in the single-director case, denoted by \( n^* \). In other words, we wish to understand whether hiring an additional director who provides yet another evaluative signal on worker performance induces the principal to hire more or fewer workers. Based on the analysis above, it is clear that the answer to this question must depend on the correlation between the two signals.

It is not obvious at the outset how the correlation should affect optimal firm size. Recall the basic tradeoff from the earlier section that larger firms are more productive but they are also more difficult to monitor and manage, captured by the \( G \) function. On top of this, the principal selects the firm size and compensation parameters jointly. Therefore, even if directors with uncorrelated signals are more efficient in monitoring, it is not clear how the principal will take advantage of this. For example, it is plausible that the negative correlation makes monitoring more efficient, thereby decreasing the risk premium, therefore allowing the principal to raise incentives and elicit more effort from the workers, keeping firm size fixed. It turns out, however, that the principal will, in fact, adjust firm size as well to take advantage of the negative correlation between the signals.

**Proposition 4** Assume \( c_i = c \) for all \( i \in N \), \( k_j = k \) for all \( j = 1, 2 \). When firm size \( n \) is endogenous, the optimal \( n_{II}^* \) satisfies:

\[
\begin{align*}
\text{if } \rho > 0, & \quad n_{II}^* < n^*; \\
\text{if } \rho = 0, & \quad n_{II}^* = n^*; \\
\text{if } \rho < 0, & \quad n_{II}^* > n^*.
\end{align*}
\]

To understand this result, it helps to again consider the zero-correlation case. Recall that for every fixed firm size, the profit function with two directors is identical to that with a single director. It is not surprising then that the optimal firm sizes are identical as well. Now consider the case of positive correlation. A positive correlation lowers the incentive for each worker \( (\frac{\partial b_i}{\partial \rho} < 0) \), leading to lower worker efforts \( (e_i) \). Lower worker productivity reduces the marginal benefit of maintaining a larger firm. Another effect of the positive correlation
is the loss in profit due to the additional risk premium needed to compensate each agent. The marginal effect of increasing $n$ on the loss in profit is thus substantial. Accordingly, in comparison to the single-director case (or equivalently, the setting with two directors and zero correlation), the optimal firm size is smaller.

With negatively correlated signals, the risk premiums for a given bonus rate are lower, bringing in an additional benefit of enlarging the firm. As before, negative correlation also affects worker incentives. The difference here is the presence of a subtle interaction between the variance effect and the covariance effect on worker efforts ($e_i$). On one hand, a larger firm induces more negative covariance between the two signals, making a higher bonus rate more attractive; on the other hand, a larger firm induces higher variances in each signal, making a lower bonus rate more attractive. The overall effect is to drive the optimal firm size to be larger than in the single-director setting.

Finally, we note that the profit comparison across the single-director and two-director settings is unaffected by the adjustment of firm size to its optimal levels. As in Proposition 3, even when firm sizes are attuned to be efficient in light of the correlation in performance signals, it is strictly valuable for the owner to employ a second director when, and only when, the signals he generates are negatively correlated with those of the existing director.

4 Committees

Directors do not monitor the firm in an unorganized fashion, but rather in committees, such as an audit committee, a compensation committee, a risk management committee, and so on. What are the costs and benefits of organizing directors into these different committees? We now turn to this question.

Let’s return to our first principles on the tradeoffs between a single and multiple directors. Employing an additional director introduces two possible incentive costs: a free-riding cost among the two directors (if the two work jointly to generate one signal) and additional compensation risk placed onto the workers (if the new director generates an additional signal about the workers). These two incentive costs lead to a desirability of concentrating monitoring activities to a single director. But synergies may allow the firm to overcome these costs. For example, the two directors may work jointly to produce two signals, each informative about a different part of the firm. This speaks to the issue not only of designing committees but designing the scope of each committee, and under what conditions it makes sense for committees to specialize on certain functions of the firm.

To begin, suppose the principal hires a second director and both directors work together on a single committee. Their objective is to improve monitoring of the firm, namely, to reduce
the variance on the firm’s output. The directors pool their efforts together and observe a single performance measure of firm output. Thus, we ask whether it makes sense for two directors to form a committee, when both directors effectively serve the same role in monitoring. Is this any better than a single director operating independently? Our next result shows the answer to this is no. We refer to this as Joint Monitoring (JM). The second directors is also work- and risk-averse, with a personal cost function \((k_2 m)\) and an exponential utility function. Letting \(m_1\) and \(m_2\) represent the monitoring efforts of director 1 and 2, respectively, the distribution of the realized signal is given by:

\[
y = \sum_{i \in N} e_i + \epsilon \quad \text{where} \quad \epsilon \sim N \left(0, \frac{G(n)\sigma^2}{m_1 + m_2}\right),
\]

The firm offers each director a linear contract, \(\omega_j = \alpha_j + \beta_j y\), for \(j = 1, 2\). Following the similar solution technique used earlier, we can calculate the equilibrium profit in this two-director setting for a given firm size \(n\), denoted by \(\Pi_{JM}(n)\), and compare this with the profit in the one-director setting (equation 7). Assume workers and directors are identical.

**Lemma 2** Assume firm size \(n\) is exogenous, \(c_i = c\) for all \(i \in N\), and \(k_j = k\) for all \(j = 1, 2\). Then, assuming symmetric managerial contracts, the value of the second director is negative, i.e.,

\[
\Pi_{JM}(n) < \Pi(n).
\]

This lemma implies that, with joint monitoring, the efficient contracting arrangement involves offering one director the optimal contract in the single-director setting, and offering the other zero for all \(y\). If effect, this reverts to the solution to the single-director setting.\(^{15}\)

The basic reason is a hidden cost of joint monitoring production. That is, the monitoring technology invites free-riding. In the director’s optimization problem, each director’s effort choice is subjected to a shirking incentive at the margin (i.e., director 1’s monitoring effort is negatively related to director 2’s effort). In other words, the precision of firm performance signal is a public good and each director internalizes the marginal (personal) cost but not all the marginal benefit; this leads to shirking. In additional, given the linear cost function for managerial effort, there is no social cost-savings in employing two directors to produce a signal of a given precision. Thus, the director’s incentive is less effective in equilibrium.\(^{16}\)

---

\(^{15}\)Even though we conjectured a symmetric solution in presenting this result, it can be shown that if the two directors have identical talent, \(k_1 = k_2\), there are only two equilibria in which non-trivial monitoring exists: (1) the symmetric solution where \(\beta_1 = \beta_2 > 0\); and (2) the single-director solution where either \(\beta_1 > 0\) or \(\beta_2 > 0\) is zero.

\(^{16}\)Huddart and Liang (2005) reached the same conclusion with a similar monitoring technology except in a partnership, as opposed to a principal-agent setting.
This suggests that committees are useful only if they serve a monitoring function different than a single director operating on his own. Given that each director has its own moral hazard problem, the principal faces an incentive cost of hiring additional directors. In order for it to be profitable to hire these directors, there must be a benefit. This benefit can come from defining more explicitly the production function of the committee. This benchmark leads us to consider one possible scenario.

4.0.1 Specializing Oversight of Committees

By their very definition, committees specialize in certain functions. For example, the compensation committee designs and structures compensation for management, while the audit committee oversees internal controls, and the risk committee monitors the risk management system of the firm. These committees have two primary economic features. First, the oversight of each committee is not the entire firm, but rather sections of the firm. For example, the risk management committee spends its resources overseeing all employees and management concerned with risk management practices, which may not include other areas of the firm that are excluded from risk management. Second, because the board is a small firm relative to the size of the entire organization, directors on different committees often work together to monitor the firm. In fact, sometimes directors serve on multiple committees explicitly.

For example, consider a major financial institution like an investment bank. Though the compensation committee and the risk management committee are separate and each deal with different functions of the firm, directors on each committee provide input and analysis toward both compensation and risk management. Because the two functions overlap and because the resources of the directors are constrained, each director on the two different committees will jointly monitor the firm. It is impossible to determine whether a bank is taking excessive risk without considering its compensation structure, and vice versa, any understanding of executive compensation must take into account the quality of risk management. Because the monitoring function of the board is complex and directors’ time is scarce, directors often do “double duty” in terms of monitoring the firm.

We now specify that the two directors can work together to produce two signals, where each signal is informative about the collective effort of a separate sub-group of workers. For example, the risk management committee oversees all employees working in risk management, while the audit committee oversees all employees working in internal audit. We label this “Joint Specialized Monitoring” assignment or simply JSM. The benefit of having the directors work jointly is that their interaction allows them to produce signals that are more focused (e.g., have lower variance). This can be achieved when directors communicate with each other and are able to isolate the contribution of each sub-group within the firm. As a result, each signal
is more precise than if the directors had worked independently. The key to the success of this monitoring assignment thus lies in the tradeoff between the synergy benefit and the shirking incentive inherent in joint monitoring activities.

The monitoring technology behind the JSM assignment is characterized as follows:

\[ y_1 = \sum_{i=1}^{n} e_i + \epsilon_1 \quad \text{where } \epsilon_1 \sim N\left(0, \frac{H(n)\sigma^2}{m_1 + m_2}\right); \]

\[ y_2 = \sum_{i=\frac{n}{2}+1}^{n} e_i + \epsilon_2 \quad \text{where } \epsilon_2 \sim N\left(0, \frac{H(n)\sigma^2}{m_1 + m_2}\right). \]

Note that each signal is now focused on a sub-group of the firm that is half in size. Dividing the firm in half eases calculations without changing the results; more general divisions of the firm are possible, but yield the same qualitative outcomes. The precision of each signal is affected by the monitoring effort by both directors, capturing the idea of joint managerial activities. Function \( H(n) \) captures the synergy generated by asking the two directors to work together to produce specialized signals. If \( H(n) \leq G(n) \), a synergy is present. In the following analysis, we sometimes assume \( H(n) = G\left(\frac{n}{2}\right) \). This specification represents substantial synergy, especially when \( G \) is convex (such as \( G(n) = n^2 \)).\(^{17}\)

Using the solution techniques developed earlier, we can solve the principal’s problem for the JSM setting. We next show that under the JSM assignment, the profit with two directors exceeds that with a single director if synergy and the value of output (\( q \)) are high enough.

**Proposition 5** Assume firm size \( n \) is exogenous, \( c_i = c \) for all \( i \in N \), and \( k_j = k \) for all \( j = 1, 2 \). Let \( G(n) = n^2 \) and \( H(n) = G\left(\frac{n}{2}\right) \). If \( q \) is sufficiently large, the firm prefers to organize the board into specialized committees.

The comparison in the proof of Proposition 5 is between profits under this Joint Specialized Monitoring setting of this section against profits under a single director from the benchmark model. Of course, we urge readers not to take the number of directors too literally, since we can easily think of these as groups of directors. In that interpretation, this proposition compares the structure of the board and argues that if the market price of output is sufficiently large, there is an economic benefit to organizing the board in two separate committees, where each committee specializes in certain functions but collaborate in their monitoring activity.

In deciding whether to expand the number of directors, the usual direct and indirect cost associated with hiring a second director must be considered, as well as the hidden cost of

\(^{17}\)A recent study by Marino and Zabojnik (2004) also considers the optimality of splitting a firm into two sub-groups. While we consider the resulting monitoring synergies, their key driving force is the incentive tradeoff of holding a tournament between two groups of risk-neutral workers.
shirking between the two directors. Notice that with the presence of synergy, each signal is now focused and thus more precise for a given managerial effort than the single-director case. When such synergy (indexed by the $H(n)$ function) is substantial, hiring a second director with a $JSM$ assignment overcomes not only the usual direct and indirect cost of the additional monitoring activity but also the hidden cost due to the induced free-riding incentives associated with multiple directors. In that case, the $JSM$ assignment allows the firm higher profits than the single-director setting.

The analysis of the two-director setting (Section 4) indicates that the desirability of hiring more directors lies in the synergy/complementarity that exists with group monitoring (such as the low-variance specialized signals in the $JSM$ setup). Given the free-riding and indirect costs associated with the expansion of the size of the management firm, it is essential for organizations to realize substantial synergy for group monitoring to be viable.$^{18}$

5 Conclusion

The corporate scandals of Enron and WorldCom in 2002 refocused world attention to the importance of corporate governance. These events gave rise to the Sarbanes-Oxley Act of 2002, as well as massive shifts in the regulatory architecture for corporations. The recent global financial crisis has further highlighted the importance of good corporate governance. But the regulatory notice of board independence that has taken hold in both the Sarbanes-Oxley Act, as well as the listing requirements of the major United States stock exchanges, is just the beginning. Their definition of independence is vague and subject to broad interpretation. And thus the empirical literature that seeks to establish a relationship between independence and performance has failed to uncover a clear relationship. Part of the problem is that the regulatory notion of independence establishes a relationship only between the director and the firm, and not between different directors. We seek to provide theoretical guidance on the effects of differences in board composition. This takes the level of analysis on optimal board design one level deeper than the simple regulatory rubrics of board independence.

Critics will no doubt remark that boards serve several functions, and we model only one, the monitoring function. For sure, boards also serve important functions in terms of helping set the strategic direction of the company. We leave for future research a broader role for the board and confine our attention in our paper to purely the monitoring aspect.

Another critique is the lack of the CEO in the model. Many will argue that the primary responsibility of the board is to pick the CEO and decide when and whether to fire him. Again,

---

$^{18}$In fact, we can show in the $JSM$ setting that if synergy exists but is not large, say $H(n) = \frac{G(n)}{2}$ while $G(n) = n^2$, the $JSM$ assignment is inferior to the single-monitor setting for any value of $q$. 

19
we seek to write the simplest model that captures the essential function of the board and do not draw a distinction between labor and management. This is also an area of future work, and in particular, we would like to model more explicitly the actual monitoring function at an even more microeconomic level between the CEO and the board. Nonetheless we hope that our results on board composition and firm size will entice empirical researchers to collect data sets that can test our model. The next frontier of research on corporate governance and board design will involve thinking carefully about the actual economic relationships within and between directors. We leave for future work continuing steps in this direction.

6 Appendix

Proof of Lemma 1. From equations (6) and (7), the owner’s profit function is

\[ \Pi(n) = \frac{n}{2c} \left( q - 2cn\sqrt{rkG(n)/n} \right)^2, \]

and this is defined over the relevant range where \( b^*(n) = q - 2cn\sqrt{rkG(n)/n} \geq 0 \). Note that \( h(n) = \sqrt{G(n)/n} \) is an increasing function of \( n \). This implies \( n \in [0, \pi] \), where \( \pi \) is such that \( q - 2cn\sqrt{rk}(\pi) = 0 \). Evaluating \( \Pi(n) \) at the two bounds, it is clear that \( \Pi(0) = 0 \). Moreover,

\[ \Pi(0) = \frac{1}{2c} \cdot \lim_{n \to 0} \left( nq^2 + 4c^2\sigma^2rkG(n) - 4qc\sigma\sqrt{rknG(n)} \right) = 0. \]

By Rolle’s theorem, we are therefore assured that an interior \( n^* \) exists with \( \Pi'(n^*) = 0 \). Now, the derivative of \( \Pi(n) \) with respect to firm size is given by:

\[ \Pi'(n) = \frac{b^*(n)}{2c} \left( q - 2cn\sqrt{rk}(n) - 4nc\sigma\sqrt{rkh(n)} \right). \]

Since \( b^*(n) > 0 \) in the interior, any \( n^* \) that satisfies \( \Pi'(n^*) = 0 \) is characterized by:

\[ h(n^*) + 2n^*h'(n^*) = \frac{q}{2cn\sqrt{rk}}. \]

The expression on the left is an increasing function of \( n \) if and only if \( 3h'(n) + 2nh''(n) > 0 \). As \( h(n) = \sqrt{G(n)/n} \), we have

\[ h'(n) = \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)} \left( \frac{1}{n} - \frac{G(n)}{G'(n)n^2} \right)} \]

and

\[ h''(n) = \frac{\sqrt{n}G'(n)}{4n\sqrt{G(n)} \left( \frac{2G''(n)}{G'(n)} + \frac{3G(n)}{n^2G'(n)} - \frac{G'(n)}{G(n)} - \frac{2}{n} \right)}. \]
Together, (10) and (11) imply that
\[
3h'(n) + 2nh''(n) = \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left( 2 \frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} \right) \\
> \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left( \frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} \right) \\
\geq 0,
\]
where the strict and weak inequalities follow from \( G''(\cdot) > 0 \) and condition (1), respectively.
Therefore, a unique \( n^* \) exists satisfying (9). To verify the second-order conditions at this point, we differentiate \( \Pi'(n) \) with respect to \( n \) to obtain:
\[
\Pi''(n) = \frac{b^*(n)}{2c} \left( q - 2c\sigma\sqrt{rk}h(n) - 4nc\sigma\sqrt{rk}h'(n) \right) \\
+ \frac{b^*(n)}{2c} \left( -2c\sigma\sqrt{rk}h'(n) - 4c\sigma\sqrt{rk}h'(n) - 4nc\sigma\sqrt{rk}h''(n) \right).
\]
Evaluating this expression at \( n^* \), we have
\[
\Pi''(n^*) = 0 + \frac{b^*(n^*)}{2c} \left( -2c\sigma\sqrt{rk} \right) \left( 3h'(n^*) + 2n^*h''(n^*) \right) < 0,
\]
where the last inequality follows from (12) and \( b^*(n^*) > 0 \). We can therefore conclude that \( n^* \) is the globally optimal firm size.

**Proof of Proposition 1.** Let \( \star \) denote any of \( c, \sigma, \sqrt{r}, \) and \( \sqrt{k} \), and let \( \check{\star} \) denote the product of the other three. Implicitly differentiating the condition for \( n^* \) in (9) shows that
\[
\frac{\partial n^*}{\partial \star} = (3h'(n^*) + 2n^*h''(n^*))^{-1} \left( -\frac{q}{2\star^2} \right) < 0,
\]
as the first term is strictly positive from (12).

**Proof of Proposition 2.** The worker’s problem in the altered two-directors setting is:
\[
\max_{e_i} a_i + b_1 \sum_j e_j + b_2 \sum_j e_j - \frac{c_i}{2} e_i^2 - \frac{r}{2} b_1^2 \frac{G(n)\sigma^2}{m_1} - \frac{r}{2} b_2^2 \frac{G(n)\sigma^2}{m_2} - \frac{r}{2} \rho b_1 b_2 G(n)\sigma^2,
\]
which leads to a first-order condition on effort, \( e_i^* = \frac{b_1 + b_2}{c_i} \).
The original director’s choice of monitoring effort solves the following problem:
\[
\max_{m_1} \alpha_1 + \beta_1 \sum_{i=1}^{n} e_i - k_1 m_1 - \frac{r}{2} \beta_1^2 \frac{G(n)\sigma^2}{m_1},
\]

leading to the first-order condition, \(m_1^* = \beta_1 \sigma \sqrt{\frac{G(n)r}{2k_1}}\).

With these findings, we can calculate firm profit and the principal’s optimization problem:

\[
\max_{\beta_1, m_2, b_i} \sum_{i=1}^{n} \left[ \frac{q (b_{1i} + b_{2i})}{c_i} - \frac{(b_{1i} + b_{2i})^2}{2c_i} - \frac{r b_{2i}^2 G(n)\sigma^2}{2m_1} - \frac{r b_{2i} G(n)\sigma^2}{m_2} - \frac{r}{2} \rho b_{1i} b_{2i} G(n)\sigma^2 \right] - k_1 m_1 - k_2 m_2 - \frac{r}{2} \beta_2^2 \frac{G(n)\sigma^2}{m_1}.
\]

Now fix \(m_2 = 0\) and \(b_{2i} = 0\). The above unconstrained maximization problem becomes the corresponding unconstrained maximization problem for a single director problem. Therefore, the single director solution is feasible in the altered two-director problem. Examining the first-order conditions reveals that the optimal \(m_2\) and \(b_{2i}\) are not zero, so we have \(\Pi^*_1(\rho, n) > \Pi(n)\).

Next we show that the equilibrium profit decreases in \(\rho\). Substituting \(m_1\) into the profit function, the resulting profit function contains only the optimizers \((\beta_1, m_2, b_i)\) and exogenous parameters. Notice that \(\rho\) only appears once, as a negative, linear term. It follows immediately from the Envelope Theorem that the equilibrium profit function is decreasing in \(\rho\). ■

**Proof of Proposition 3.** The worker’s problem in the two-director setting is given by:

\[
\max_{e_i} a_i + b_{1i} \sum_j e_j + b_{2i} \sum_j e_j - \frac{c_i}{2} e_i^2 - \frac{r b_{1i}^2 G(n)\sigma^2}{2m_1} - \frac{r b_{2i}^2 G(n)\sigma^2}{m_2} - \frac{r}{2} \rho b_{1i} b_{2i} G(n)\sigma^2,
\]

which leads to a first-order condition on effort, \(e_i^* = \frac{b_{1i} + b_{2i}}{c_i}\).

Each director’s choice of monitoring effort solves the following problem:

\[
\max_{m_j} \alpha_j + \beta_j \sum_{i=1}^{n} e_i - k m_j - \frac{r}{2} \beta_j^2 \frac{G(n)\sigma^2}{m_j},
\]

leading to the first-order condition, \(m_j^* = \beta_j \sigma \sqrt{\frac{G(n)r}{2k_j}}\).

With these findings, we can calculate firm profit in the two-director setting. The principal’s optimization problem is:
This leads to the following first-order conditions:

\[
b_{1i}^* = \frac{q - (1 + \frac{\rho c_i G(n)\sigma^2}{\sigma m_1}) b_{2i}^*}{1 + \frac{\rho c_i G(n)\sigma^2}{\sigma m_1}} \quad \text{and} \quad b_{2i}^* = \frac{q - (1 + \frac{\rho c_i G(n)\sigma^2}{\sigma m_1}) b_{1i}^*}{1 + \frac{\rho c_i G(n)\sigma^2}{\sigma m_1}}\]

and \[\frac{r}{2} G(n)\sigma^2 \sum_{i=1}^{n} b_{2i}^2 - 2k_j \frac{\partial}{\partial \beta_j} m_j^*(\beta_j) = 0.\]

Substituting \(m_j\) and \(\frac{\partial}{\partial \beta_j} m_j^*(\beta_j)\) into the last condition, we arrive at

\[\beta_j^* = \sqrt{\frac{1}{2} \sum_{i=1}^{n} b_{1i}^2}.\]

Now simplifying these conditions with identical workers and directors (i.e., \(c_i \equiv c\) and \(k_i \equiv k\)) implies \(b_i^* \equiv b^*, \beta_j^* \equiv \beta^*\) and \(m_i^* \equiv m^*\). We use these identities to solve \(\beta^*\) and \(b^*\).

In equilibrium, we have:

\[
\beta^* = b^* \sqrt{\frac{n}{2}}, \\
m^* = b^* \frac{\sigma}{\sqrt{\frac{G(n)n r}{4k}}} > 1; \\
b^* = \frac{q - (1 + \frac{\rho c G(n)\sigma^2}{\sigma m^*}) b^*}{1 + \frac{\rho c G(n)\sigma^2}{\sigma m^*}} = \frac{q - (1 + \frac{\rho c G(n)\sigma^2}{\sigma m^*}) b^*}{1 + \frac{2c\sqrt{rG(n)/n}}{b^*}} \\
\implies b^* = \frac{q - 2c\sigma \sqrt{rG(n)/n}}{2 + \frac{\sigma c G(n)\sigma^2}{b^*}}.
\]

Total profits can then be computed as follows:

\[
\Pi_{II}(\rho, n) = n \left[ \frac{q2b^*}{c} - \frac{(2b^*)^2}{2c} - \frac{r}{2}(2b^*) G(n)\sigma^2 - \frac{r}{2}\rho b^* G(n)\sigma^2 \right] - 2km^* - \frac{r}{2} \frac{(2\beta^2) G(n)\sigma^2}{m^*}
\]

To sidestep needless technical complications, we restrict attention throughout to parameter spaces for which the optimal incentive weights are strictly positive and bounded. This is primarily an issue only when correlation levels are extremely negative.
After extensive simplification, this reduces to:

\[
\Pi_{II}(\rho, n) = \Pi(n)W(\rho, n), \quad \text{where} \quad W(\rho, n) = \frac{2}{2 + \frac{r}{c}G(n)\sigma^2}.
\] (13)

Recall that \(\Pi(n)\) is the profit in the single-director case (see equation (8)) and is given by:

\[
\Pi(n) = \frac{n}{2c} \left( q - 2c\sigma \sqrt{rkG(n)/n} \right)^2.
\]

When \(\rho = 0\), \(W(0, n) = 1\), so \(\Pi_{II}(0, n) = \Pi(n)\). Combined with the fact \(\frac{\partial}{\partial \rho}W(\rho, n) < 0\), we immediately obtain \(\Pi_{II}(\rho > 0, n) < \Pi(n) < \Pi_{II}(\rho < 0, n)\).

\[\blacksquare\]

**Proof of Proposition 4.** From (13), differentiating \(\Pi_{II}(\rho, n)\) with respect to \(n\) yields the following derivative:

\[
\Pi'_{II}(\rho, n) = \Pi'(n)W(\rho, n) + \Pi(n)W_n(\rho, n),
\]

where

\[
W_n(\rho, n) = \frac{-rG'(n)\sigma^2}{(2 + \frac{r}{c}G(n)\sigma^2)^2}.
\]

Note that \(\text{sign}(W_n(\rho, n)) = \text{sign}(-\rho)\).

Consider \(\rho = 0\). Because \(W(0, n) = 1\) and \(W_n(0, n) = 0\), we have \(\Pi'_{II}(0, n^*_I) = \Pi'(n^*) = 0\), leading to \(n^*_I = n^*\). In this setting, the optimal firm size exists and is unique.

Now consider \(\rho > 0\). We have \(W(\rho > 0, n) > 0\), \(W_n(\rho > 0, n) < 0\), and \(\Pi(n) > 0\). Moreover, from the proof of Lemma 1, \(\Pi'(n) \leq 0\) for all \(n^* \leq n < \bar{n}\). This implies that \(\Pi'_{II}(\rho > 0, n) < 0\) for all \(n^* \leq n < \bar{n}\), i.e., it must be the case that \(n^*_I < n^*\). Notice that the optimal firm size exists, but is possibly not unique; however, our result establishes that every firm size that is optimal is smaller than the optimum from the single-director case.

Finally, suppose that \(\rho < 0\). For negative values of \(\rho\), \(W_n(\rho < 0, n) > 0\), \(\Pi(n) > 0\), and \(W(\rho < 0, n) > 0\). As before, from the proof of Lemma 1, \(\Pi'(n) \geq 0\) for all \(0 < n \leq n^*\), implying that \(\Pi'_{II}(\rho < 0, n) > 0\) for all \(0 < n \leq n^*\). As in the positive correlation setting, multiple optima may exist for firm size. But we have shown that when \(\rho < 0\), every optimal \(n^*_I > n^*\).

\[\blacksquare\]
Proof of Lemma 2. We first derive the equilibrium contract and action choices.

Each director maximizes the certainty equivalent of his wage, taking the other director’s effort as given. This yields the following pair of first-order conditions:

\[ m_1^* = \beta_1 \sigma \sqrt{\frac{G(n)r}{2k_1}} - m_2 \quad \text{and} \quad m_2^* = \beta_2 \sigma \sqrt{\frac{G(n)r}{2k_2}} - m_1. \]  

(14)

In turn, the firm chooses \( b_i \) and \( \beta_j \) to maximize total surplus:

\[
\max_{\beta, b_i} \sum_{i=1}^{n} \left[ \frac{qb_i}{c_i} - \frac{b_i^2}{2c_i} - \frac{r}{2} \frac{G(n)\sigma^2}{m_1^* + m_2^*} \right] + \left( k_1 m_1 + k_2 m_2 \right) - \frac{r}{2} \left( \beta_1^2 + \beta_2^2 \right) \frac{G(n)\sigma^2}{m_1^* + m_2^*},
\]

yielding a first-order condition on \( b_i \) analogous to that in the one-director setting:

\[ b_i^* = \frac{q}{1 + \frac{rcG(n)\sigma^2}{m_1^* + m_2^*}}. \]

Similarly, the first-order condition on \( \beta_j \), for \( j = 1, 2 \), is given by:

\[
\left[ \frac{r}{2} \frac{G(n)\sigma^2}{(m_1^* + m_2^*)^2} \sum_{i=1}^{n} b_i^2 - k_j + \frac{r}{2} \frac{G(n)\sigma^2}{m_1^* + m_2^*} \left( \beta_1^2 + \beta_2^2 \right) \right] \sigma \sqrt{\frac{G(n)r}{2k_j}} - r \beta_j \frac{G(n)\sigma^2}{m_1^* + m_2^*} = 0. \]  

(15)

The difference here is that the firm must consider the behavior of both directors as opposed to a single director in choosing worker and director incentives.

Assuming identical workers (\( c_i = c \) for all \( i \in N \)) and directors (\( k_j = k \) for all \( j = 1, 2 \)), we conjecture a symmetric solution where \( b_i^* = b^* \) and \( \beta_i^* = \beta^* \). The symmetry simplifies the first-order conditions and helps solve for equilibrium incentive choices. In particular, we first simplify the directors’ effort choice in equilibrium. Because \( m_1^* = m_2^* \),

\[ m_1^* = \beta^* \sigma \sqrt{\frac{G(n)r}{2k}} - m_2^* \implies m_1^* = m_2^* = \frac{\beta^* \sigma}{2} \sqrt{\frac{G(n)r}{2k}}. \]  

(16)

which leads to a simplification of worker incentive \( b_i \),

\[ b^* = \frac{q}{1 + \frac{rcG(n)\sigma^2}{2m^*}} = \frac{q}{1 + \frac{c \sigma \sqrt{2krG(n)}}{\beta^*}}. \]

Substituting (16) and \( \beta_1^* = \beta^* \) into (15), we have

\[ \beta^* = b^* \sqrt{n}. \]

Combine the last two equations and we have a closed-form for \( b^* \):

\[ b^* = q - c \sigma \sqrt{2krG(n)/n}. \]
The derivation of the profit function is as follows:

\[
\Pi_{JM}(n) = n \left[ \frac{q b^*}{c} - \frac{b^*}{2c} - \frac{b^*}{2b^*} \frac{2 \sigma \sqrt{2 r k G(n)/n}}{2 b^*} \right] - k (2 m^*) - 2 \beta^2 \sigma \sqrt{2 r k G(n)/n} - 2 b^* \sigma \sqrt{2 r k G(n)/n} \\
= \frac{nb^*}{2c} \left[ 2 q - b^* - c \sigma \sqrt{2 r k G(n)/n} \right] - k b^* \sqrt{n} \sigma \sqrt{\frac{G(n) r k}{2 k}} - 2 nb^* \sigma \sqrt{2 r k G(n)/n} \\
= \frac{nb^*}{2c} \left[ 2 q - b^* - c \sigma \sqrt{2 r k G(n)/n} \right] - 0.5 nb \sigma \sqrt{2 r k G(n)/n} - nb \sigma \sqrt{2 r k G(n)/n} \\
= \frac{nb^*}{2c} \left[ 2 q - b^* - c \sigma \sqrt{2 r k G(n)/n} - 3 c \sigma \sqrt{2 r k G(n)/n} \right] \\
= \frac{nb^*}{2c} \left[ q - 3 c \sigma \sqrt{2 r k G(n)/n} \right].
\]

As to claim (2), recall that with one director, the profit for any exogenous \( n \) is given by \( \frac{n}{2c} \left( q - 2 c \sigma \sqrt{r k G(n)} \right)^2 \). So, the value of a second director is:

\[
\frac{n}{2c} \left[ q - c \sigma \sqrt{2 r k G(n)/n} \right] \left[ q - 3 c \sigma \sqrt{2 r k G(n)/n} \right] - \frac{n}{2c} \left( q - 2 c \sigma \sqrt{r k G(n)/n} \right)^2 \\
= \frac{n}{2c} \left[ \left( q - 2 c \sigma \sqrt{r k G(n)/n} \right) + (2 - \sqrt{2}) c \sigma \sqrt{r k G(n)/n} \right] \\
\left[ \left( q - 2 c \sigma \sqrt{r k G(n)/n} \right) + (2 - 3 \sqrt{2}) c \sigma \sqrt{r k G(n)/n} \right] - \frac{n}{2c} \left( q - 2 c \sigma \sqrt{r k G(n)/n} \right)^2 \\
= \frac{n}{2c} \left[ (4 - 4 \sqrt{2}) c \sigma \sqrt{r k G(n)/n} \left( q - 2 c \sigma \sqrt{r k G(n)/n} \right) + (2 - \sqrt{2})(2 - 3 \sqrt{2}) \left( c \sigma \sqrt{r k G(n)/n} \right)^2 \right].
\]

Notice that \( b^* = \left( q - 2 c \sigma \sqrt{r k G(n)/n} \right) > 0 \) and all other parameters \((c, r, k, G(n), \text{and } \sigma)\) are positive, so the above quantity is always negative. The second director thus has negative value.

We start with each worker’s problem under \( JSM \). If \( i \in \{1, 2, ..., n/2\} \), then we have

\[
\max_{e_i} a_i + b_{1i} \sum_{j=1}^{n/2} e_j - \frac{c_i e_i^2}{2} - \frac{r}{2} b_{1i}^2 \frac{H(n) \sigma^2}{m_1 + m_2}
\]

leading to the first-order condition \( e_i^* = \frac{b_{1i}}{c_i} \). For \( j \in \{n/2 + 1, n/2 + 2, ..., n\} \), the analogous condition is \( e_j^* = \frac{b_{1j}}{c_j} \). It is easy to see that each worker’s contract should depend only on the signal about his own sub-group of the firm, i.e., \( b_{2i} = 0 \), for \( i \notin \{1, 2, ..., n/2\} \) and \( b_{1i} = 0 \), for \( i \notin \{n/2 + 1, n/2 + 2, ..., n\} \).

The director’s problem is

\[
\max_{m_j} \alpha_1 + \beta_{11} \sum_{i=1}^{n/2} e_i + \beta_{12} \sum_{i=n/2+1}^{n} e_i - k_j m_j - \frac{r}{2} \left( \beta_{11}^2 + \beta_{12}^2 \right) \frac{H(n) \sigma^2}{m_1 + m_2},
\]

leading to the first-order condition \( m_j^* = \frac{\sigma}{\sqrt{\beta_{11}^2 + \beta_{12}^2}} \frac{H(n) \sigma}{2k_j} - m_i \), for \( l \neq j \).
Substituting these two conditions, along with the usual IR constraints, the principal’s problem is now

\[
\max_{\beta, l, b_i} \sum_{i=1}^{n/2} \left[ qb_{1i} + \frac{b_{1i}^2}{2c_i} - \frac{r}{2} b_{1i} \frac{H(n)\sigma^2}{m_1 + m_2} \right] + \sum_{j=n/2+1}^{n} \left[ qb_{2j} + \frac{b_{2j}^2}{2c_j} - \frac{r}{2} b_{2j} \frac{H(n)\sigma^2}{m_1 + m_2} \right]
\]

\[-k_1 m_1 - k_2 m_2 - \frac{r}{2} \left( \beta_{11}^2 + \beta_{12}^2 + \beta_{21}^2 + \beta_{22}^2 \right) \frac{H(n)\sigma^2}{m_1 + m_2} \]

leading to the first-order conditions

\[
b_{1i}^* = \frac{q}{1 + \frac{rc_i H(n)\sigma^2}{m_1 + m_2}} \quad \text{and} \quad b_{2j}^* = \frac{q}{1 + \frac{rc_j H(n)\sigma^2}{m_1 + m_2}}.
\]

For signal \(y_l\) and director \(j\), we have

\[
\left[ \frac{r}{2} \frac{H(n)\sigma^2}{(m_1 + m_2)^2} \left( \sum_{i=1}^{n/2} b_{1i}^2 + \sum_{i=n/2+1}^{n} b_{2i}^2 \right) - k \right. \left. - \frac{r}{2} \frac{H(n)\sigma^2}{(m_1 + m_2)^2} \left( \beta_{11}^2 + \beta_{12}^2 + \beta_{21}^2 + \beta_{22}^2 \right) \right] \frac{\partial}{\partial \beta_{lj}} m_j^* (\beta_{lj})
\]

\[-r \beta_{lj} \frac{H(n)\sigma^2}{m_1 + m_2} = 0,
\]

where director \(j\)'s first-order condition implies

\[
\frac{\partial}{\partial \beta_{lj}} m_j^* (\beta_{lj}) = \sigma \frac{\sqrt{H(n)r}}{2k_j} \frac{\beta_{lj}}{\sqrt{\beta_{1j}^2 + \beta_{2j}^2}}.
\]

Simplifying these conditions with identical workers and directors (i.e., \(c_i \equiv c\) and \(k_i \equiv k\)) implies \(b_{1i}^* = b_{2i}^* \equiv b^*\), \(\beta_{lj}^* \equiv \beta^*\) and \(m_j^* \equiv m^*\). We use these identities to solve \(\beta^*\) and \(b^*\). In equilibrium, we then have

\[
2 m^* = \beta^* \sigma \sqrt{\frac{H(n)r}{k}};
\]

\[
b^* (\beta^*) = \frac{q}{1 + \frac{rc H(n)\sigma^2}{2m^*}} = \frac{q}{1 + \frac{c\sigma H(n)r}{\beta^*}};
\]

and

\[
\left[ \frac{r}{2} \frac{H(n)\sigma^2}{4m^*^2} nb^*^2 - k + \frac{r}{2} \frac{H(n)\sigma^2}{4m^*^2} 4\beta^*^2 \right] \sigma \sqrt{\frac{H(n)r}{4k}} - r \beta^* \frac{H(n)\sigma^2}{2m^*} = 0.
\]
Substituting \( m^*(\beta^*) = \frac{\beta^*}{2} \sigma \sqrt{\frac{H(n)r}{k}} \), we have

\[
0 = \left[ \frac{k}{2\beta^*2} nb^2 - k + \frac{k}{2\beta^*2} 4\beta^* \right] \sigma \sqrt{\frac{H(n)r}{4k}} - 2k\sigma \sqrt{\frac{H(n)r}{4k}}
\]

\[
0 = \left[ \frac{k}{2\beta^*2} nb^2 - k + 2k - 2k \right] \sigma \sqrt{\frac{H(n)r}{4k}}
\]

\[
0 = \frac{nb^2}{2\beta^*2} - 1
\]

\[
\beta^* = b^* \sqrt{\frac{n}{2}}
\]

We can incorporate this into the \( b^*(\beta^*) \) expression to obtain

\[
b^* = \frac{q}{1 + \frac{c\sigma \sqrt{H(n)kr}}{b^* \sqrt{n/2}}} = \frac{q}{1 + \frac{c\sigma 2H(n)kr/n}{b^*}}
\]

or

\[
b^* = q - c\sigma \sqrt{2rkH(n)/n}.
\]

Total Profits with exogenous \( n \), using \( 2m^* = \beta^* \sigma \sqrt{\frac{H(n)r}{k}} = b^* \sigma \sqrt{\frac{H(n)nr}{2k}} \), and \( \frac{1}{2} \text{Var}(y) = \frac{rH(n)s^2}{2m^*} = \frac{\sigma 2rkH(n)/n}{2b^*} \), is then given by:

\[
\Pi_{JS}(n) = n \left[ qb^* \frac{c}{c} - \frac{b^*2}{2c} - b^*2 \sigma \sqrt{2rkH(n)/n} \right] - k(2m^*) - 4\beta^*2 \sigma \frac{\sqrt{2rkH(n)/n}}{2b^*}
\]

\[
= \frac{nb^*}{2c} \left[ 2q - b^* - c\sigma \sqrt{2rkH(n)/n} \right] - k\beta^* \sigma \frac{\sqrt{H(n)nr}{2k}} - 4 \left( \frac{b^*}{2} \right)^2 \frac{\sigma \sqrt{2rkH(n)/n}}{2b^*}
\]

\[
= \frac{nb^*}{2c} \left[ 2q - b^* - c\sigma \sqrt{2rkH(n)/n} \right] - \frac{nb^*}{2} \sigma \sqrt{2rkH(n)/n} - nb^* \sigma \sqrt{2rkH(n)/n}
\]

\[
= \frac{nb^*}{2c} \left[ 2q - b^* - c\sigma \sqrt{2rkH(n)/n} - 3c\sigma \sqrt{2rkH(n)/n} \right]
\]

\[
= \frac{nb^*}{2c} \left[ q - 3c\sigma \sqrt{2rkH(n)/n} \right]
\]

To show claim (2), re-express \( \Pi_{IM}(n) \) as

\[
\Pi_{JS}(n) = \frac{nb^*}{2c} \left[ q - 3c\sigma \sqrt{2rkH(n)/n} \right]
\]

\[
= \frac{n}{2c} \left[ q - c\sigma \sqrt{2rkH(n)/n} \right] \left[ q - 3c\sigma \sqrt{2rkH(n)/n} \right]
\]

Recall that with a single monitor, the profit function for exogenous \( n \) is given by \( \Pi(n) = \frac{r}{2c} \left( q - 2c\sigma \sqrt{rkG(n)/n} \right)^2 \). If \( G(n) = n^2 \) and \( H(n) = G\left(\frac{n}{2}\right) \), the value of a second director,
letting $\alpha = c\sigma \sqrt{r n k}$, is given by:

$$
\Pi_{JS}(n) - \Pi(n) = \frac{n}{2c} \left[ q - \sqrt{1/2} \alpha \right] \left[ q - \sqrt{9/2} \alpha \right] - \frac{n}{2c} (q - 2\alpha)^2
$$

$$
= \frac{n}{2c} \left[ (q - 2\alpha) + \alpha(2 - \sqrt{1/2}) \right] \left[ (q - 2\alpha) + \alpha(2 - \sqrt{9/2}) \right] - \frac{n}{2c} (q - 2\alpha)^2
$$

$$
= \frac{n\alpha}{2c} \left[ 4 - \sqrt{9/2} - \sqrt{1/2} \right] (q - 2\alpha) + (2 - \sqrt{1/2})(2 - \sqrt{9/2}) (\alpha)
$$

which is greater than zero if and only if: $q > \frac{5\alpha}{8 - 4\sqrt{2}} \approx 2.13\alpha$ ■

References

Daron Acemoglu, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti.


Praveen Kumar and K. Sivaramakrishnan. Who monitors the monitors? the effect of board independence on executive compensation and firm value.


