

# Optimal Team Size and Monitoring in Organizations

*Pierre Jinghong Liang*  
*Carnegie Mellon University*

*Madhav V. Rajan*  
*Stanford University*

*Korok Ray*  
*University of Chicago*

**ABSTRACT:** We formulate and analyze a model of team structure and monitoring within a Linear-Exponential-Normal (LEN) agency framework. We incorporate three key instruments in the internal design of an organization involving team production: team size, monitoring activities, and incentive contracts. We show that the complex trade-offs among these instruments lead to surprisingly simple implications. One such result is that the equilibrium level of pay-for-performance for workers is attenuated and is, at times, invariant to most environmental variables of interest. As such, our model helps explain the empirical puzzle of the lack of a trade-off for risk/incentives shown in standard agency models. Our work also demonstrates the presence of complementarities between team size and monitoring, and between worker talent and managerial monitoring ability. Finally, we derive predictions about the impact of environmental variables on the choice of optimal team size, incentives, and employee quality, even in the presence of an external marketplace for talent.

**Keywords:** *incentives; team size; monitoring.*

## I. INTRODUCTION

Teams play a pivotal role in modern organizations. The study of teams has occupied a similarly important place in modern economic thought, beginning with the work of Marschak and Radner (1972). The primary focus of inquiry has been the provision of incentives to motivate team members to overcome concerns regarding “free-riding.” Virtually all of this work, following the lead of Holmström (1982), takes team structure and size as given. Yet, the broader elements of organizational design also form a key component of modern firms (see Roberts 2004). Management consultants and the business

---

We are grateful to Jon Glover, Steve Huddart, Stefan Reichelstein, Amir Ziv, and three anonymous reviewers for many helpful suggestions. We have also benefited from comments by Jeremy Bertomeu, Shane Dikolli, Sunil Dutta, Qintao Fan, Nicole Johnson, Dhananjay Nanda, Richard Saouma, Haresh Sapra, Jason Schloetzer, and seminar participants at University of California, Berkeley, Carnegie Mellon, University of Chicago, City University of Hong Kong, Duke, and the Issues in Managerial Accounting Conference at The University of Texas at Austin.

Editor’s note: This paper was accepted by Steven Huddart.

*Submitted March 2006*  
*Accepted October 2007*

press increasingly exhort executives to view organizational design as yet another choice variable. The academic press has only recently begun to explore such notions (Simons 2005), and theoretical research on the topic is still nascent (see, e.g., Harris and Raviv 2002). Acemoglu et al. (2006) note that the growing interest in the determinants of firms' organizational choices is partly motivated by the belief that recent technological developments are favoring more decentralized firms.

In this paper, we provide a model of team formation and monitoring within the Linear-Exponential-Normal (LEN) agency framework that enables us to shed light on optimal incentives, team size, and the efficient matching of worker and managerial abilities. We consider three instruments in the internal design of an organization involving team production: team size, monitoring activities, and the incentive contracts offered to workers and managers. The complex interactions among the three instruments generate new insights into the optimal provision of incentives. One such insight is that when the error in the measurement of employee contribution to firm value satisfies mild regularity conditions, incentives are attenuated and in some cases optimally structured to be invariant to noise, risk preferences, and the cost of worker effort. Our findings offer a novel explanation for the lack of empirical link between risk and incentives (see, e.g., Bhattacharyya and Lafontaine 1995; Prendergast 2002).

A central feature of modern organizations is the careful design and implementation of internal and external monitoring activities. These activities influence (and are influenced by) other organizational design choices, such as team size and incentive schemes. In this study, we examine the optimal number of risk-averse workers and the nature of their incentives when the principal hires a risk-averse manager to watch over the workers. Much of the previous literature on monitoring has assumed that monitoring is either exogenous or is done by the principal herself. In this paper, we deviate from this (unrealistic) assumption and, by using a LEN formulation, generate closed-form solutions to the incentive problems among the principal, the self-interested manager, and the workers being monitored. We derive the determinants of optimal team size and characterize the impact of this choice on the strength of the incentives given to managers and workers. We also identify conditions under which it is optimal to employ a manager, and demonstrate the existence of endogenous complementarities across the skill levels of the manager and workers. Finally, we provide comparative statics results on the impact of changes in uncertainty and performance measurement technology on the efficient mix of worker and manager talent employed by the firm.

Given our choice to emphasize three main instruments—team size, monitoring, and incentives—in our analysis, our results are more likely to be relevant to firms with organizational features that emphasize these three dimensions. We believe modern knowledge-based industries fit our broad description well. They are more likely to be organized into teams of varying sizes; their incentives are more likely to be output based; and professional managers are engaged in monitoring efforts from individual workers. As an example, consider a software manager who supervises a team of software engineers. As the team grows, it becomes more difficult to coordinate these multiple engineers and this makes performance measurement harder. Our paper speaks to the efficient provision of incentives for the supervisor and employees, the limits placed on the size of teams by incentive and monitoring considerations, and the optimal mix of talent levels across workers and managers.

Some authors, notably Williamson (1967), have suggested that the inherent loss of control from hiring monitors ultimately limits the size of the firm. Calvo and Wellisz (1978) formalize this argument and demonstrate that Williamson's results require particular *ad hoc*

assumptions on the firm's monitoring technology. For example, if agents are unaware of when they are being monitored, then Calvo and Wellisz (1978) find that there is no hierarchical limit to firm size. In our model, the active monitoring reduces the variance of the group performance measure. As such, our managers can be interpreted as accountants or auditors whose job is to reduce measurement errors in performance metrics. The reduction in noise allows the principal to offer stronger incentives to the workers without the burden of paying additional risk premia. In our model, there is no loss of control with the delegation of monitoring to a third party, as our entire model is cast in a moral hazard setting.

Because our monitors are separate from the principal, our framework differs from the traditional role assigned to monitors in the accounting literature (see, e.g., Baiman and Demski 1980). More recent work in the economics literature has considered the possibility of having working agents monitor one another via either direct observation (Ma 1988; Varian 1990) or via indirect observation of correlated signals (Fischer and Hughes 1997). In his paper, Varian (1990) shows that it may be optimal to accept only groups of agents, because the agents will only select "good types" for their own group. In Section III, we allow for heterogeneous types of workers and managers and find that as risk increases, rather than selecting a better monitor, the principal will want to decrease the team size, and hire a lower quality monitor.

We utilize a multi-agent LEN framework to analyze issues related to optimal team incentives. Perhaps the earliest paper to employ this approach was McLaughlin (1994), who looked at the determinants of the links between individual compensation and team performance. In Adams (2006), the productive complementarities among team members in team production make the use of firm-wide incentives valuable, even in the presence of the induced free-rider problem. Recent work has extended the analysis of team incentives to incorporate elements related to optimal task assignments (Hughes et al. 2005) and the role of productive synergies in teams (Autrey 2005). Hughes et al. (2005) find sufficient conditions for an owner to prefer diverse task assignments as opposed to specific assignments. Autrey (2005) demonstrates that an owner favors the measurement of individual output when there is little synergy across workers' efforts, but prefers to compensate employees based on team output once the level of synergy is sufficiently high. None of the above papers incorporates managerial monitoring or allows for the endogenous determination of team size and composition.

Qian (1994) is one of the few studies that allows for an endogenous number of workers. Following Williamson (1967) and Calvo and Wellisz (1978), Qian examines the optimal size of a hierarchy and compensation throughout all tiers of the hierarchy. In Qian's (1994) paper, the principal benefits from having fewer tiers in the hierarchy because it is assumed that there are production losses at each tier and, in general, fewer tiers imply fewer managers to compensate. On the other hand, if the firm has fewer tiers, then each manager's span of control is enlarged and hence monitoring becomes less effective; the principal is thus forced to pay a higher efficiency wage to ensure that the managers put forth costly monitoring efforts. In our paper, the manager serves to reduce the noise in measuring workers' effort, and hence we only consider a two-tier hierarchy: principal at the top, and workers accompanied by the manager at the bottom. In Qian's (1994) model, the agents are assumed to be risk-neutral with limited liability and are compensated with two-tier wages. By contrast, we assume that all workers and the manager are risk-averse, and are compensated via the optimal linear contract.

From the standpoint of the accounting literature, our work is most closely related to Ziv (1993, 2000). Ziv (1993) considers the interaction between a firm's internal information

structure and its optimal size. The study focuses on how the induced effort level and firm size respond to alternative, exogenous, information structures (such as whether signal variance is constant or increasing in firm size). Ziv (2000) introduces self-interested managers and studies the optimal number of layers of a firm's hierarchy and the number of agents in each layer. In his model, the principal can monitor the agents herself, which eliminates the incentive problem with monitoring. Hiring supervisors to monitor the lower-layer agents, on the other hand, introduces such an incentive problem. Thus, Ziv (2000) concludes that "flatter" organizations are optimal in most cases.<sup>1</sup>

In our paper, we assume that the absentee principal is unable to engage in monitoring directly, and must hire one or more self-interested managers and motivate them to engage in monitoring. Comparing this setting with a benchmark (called mechanical monitoring) allows us to clearly establish the trade-offs inherent in hiring a manager (i.e., the gains from improving the worker's incentive versus the incentive problem associated with the manager's monitoring). In addition, we are able to characterize incentive contracts for the workers and manager, and can determine, in closed-form, the optimal number of workers. In turn, this allows us to demonstrate the presence of organizational synergies when manager and worker abilities are exogenously given, as well as in cases where the firm must hire employees of varying abilities (and associated reservation wages) in a marketplace for talent.

Huddart and Liang (2005) study related issues in the context of a partnership. In a partnership, the sharing of profits invites all partners to shirk their productive efforts. While monitoring activities may improve the incentive situation, they may also lead to additional shirking because the partners also act as monitors. A key difference in our paper is the presence of a risk-neutral principal who eliminates the risk-sharing considerations inherent in a partnership setting. In addition, we model managers as agents specialized in monitoring, while not directly involved in actual production, thereby capturing real-world settings beyond those in professional service firms organized as partnerships. Our setting allows us to analyze the nature of the interaction among design instruments (incentive contracts, worker size, and monitoring assignments), how such interaction leads to the attenuation of worker incentives, and the value of hiring a manager.

The remainder of the paper is organized as follows. Section II lays out the basic model and characterizes its solution under both the benchmark (mechanical) and managerial monitoring settings. Section III derives the existence of complementarities in the selection of managerial and worker talent, in particular when external labor markets are present, and provides additional comparative statics results. The last section offers concluding comments and suggests ways in which the analysis in this paper could be extended or generalized.

## II. THE MODEL

Consider a principal (the firm) contracting with  $n$  workers. Each worker exerts effort  $e_i$  at cost of effort  $C_i(e_i) = \frac{1}{2}c_i e_i^2$ . The cost of effort parameter  $c_i$  measures the quality/ability of the worker: higher quality workers have a lower  $c_i$ . Let  $N = \{1, 2, \dots, n\}$  denote the set of workers. Joint output is given by:

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

---

<sup>1</sup> Also see Baldenius et al. (2002) for related work along these lines.

We assume that  $x$  is either unverifiable or is realized too late to be used in contracting. Instead, the principal can install a mechanical monitoring system at a price (normalized to zero), which produces a signal  $y$ .<sup>2</sup> This is the only variable on which contracts can be written. We specify that:

$$y = \sum_{i \in N} e_i + \varepsilon, \text{ where } \varepsilon \sim N(0, G(n)\sigma^2).$$

Note that  $\sigma$  measures the uncertainty in performance measurement. 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 team size grows. This reflects a basic assumption on the monitoring technology, namely that the performance of larger teams is harder to measure because of problems with coordination, communication difficulties, etc. In addition, we assume that  $G(\cdot)$  satisfies the following technical condition:

$$\frac{d}{dn} \left( \frac{n \cdot G'(n)}{G(n)} \right) \geq 0. \quad (1)$$

This condition is not too restrictive, and is satisfied by a wide range of functional forms. 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 functions 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.

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:

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

yielding the incentive constraint,  $e_i = b_i/c_i$ .<sup>3</sup> Suppose each worker has an outside option, denoted by  $\bar{u}$ , which is normalized to zero.<sup>4</sup> 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 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. So the firm selects bonus coefficients (incentives) to maximize total surplus:

<sup>2</sup> We present this benchmark case in order to highlight the costs and benefits that emerge when a self-interested manager is hired to provide the informative signal about workers' efforts.

<sup>3</sup> Notice that for each worker, the effort choice is also a dominant strategy response in the workers' subgame.

<sup>4</sup> We relax this assumption in Section III.

$$\max_{b_i} \sum_{i \in N} \left[ qe_i - C_i(e_i) - \frac{r}{2} b_i^2 G(n) \sigma^2 \right].$$

This objective function reveals the two costs of production that the principal must pay the workers. The first is the personal cost of providing effort,  $C_i(e_i)$ . The second is the risk-premium,  $\frac{r}{2} b_i^2 G(n) \sigma^2$ . This is the incentive cost, which is the main focus of our study. After inserting the worker's incentive constraint, the first-order condition gives the optimal incentives for an exogenous  $n$ :

$$\hat{b}_i(n) = \frac{q}{1 + rc_i G(n) \sigma^2}, \quad (2)$$

yielding the standard risk/incentive trade-offs. Observe that a larger team leads to a higher variance of the team performance measure, which leads to the firm cutting back on incentives. Also note that the equilibrium effort from each worker is  $\hat{b}_i/c_i$ , which is lower than the first-best effort level of  $q/c_i$ .

### Incentives and Optimal Team Size

In the standard LEN model (e.g., a single-agent, single-task, single-period model), the optimal incentive weight is  $\hat{b} = 1/(1 + rc\sigma^2)$  and the pay-for-performance sensitivity  $\partial \hat{b} / \partial \sigma$  captures how incentives vary with the risk or uncertainty in the model. In the standard model this derivative is negative, suggesting a trade-off between risk and incentives. Yet the empirical support for this trade-off is weak, as noted in the introduction. We show now that adjusting for team size attenuates the effect of risk/uncertainty on incentives, which may help explain why a strong negative trade-off is not observed empirically.

In our model, assuming identical workers ( $c_i = c$ ) leads to  $\hat{b}_i(n) = \hat{b}(n)$ , which in turn allows us to solve for the optimal incentives and total profit for a given  $n$ :

$$\hat{b}(n) = \frac{q}{1 + rcG(n)\sigma^2} \quad \text{and} \quad \hat{\Pi}(n) = \frac{nq\hat{b}(n)}{2c}. \quad (3)$$

The firm chooses an optimal team size, denoted  $\hat{n}$ , which maximizes  $\hat{\Pi}(n)$ . As  $\hat{\Pi}$  and its derivative are continuous on  $[0, \infty)$  and  $\hat{\Pi}'(0) > 0$ , a necessary condition for  $\hat{n}$  to be an optimal team size is that  $\hat{\Pi}'(\hat{n}) = 0$ . This yields the following first-order condition:

$$\hat{n}G'(\hat{n}) - G(\hat{n}) = \frac{1}{rc\sigma^2}. \quad (4)$$

Since the expression on the left is 0 at  $n = 0$  and is strictly increasing in  $n$  (from the convexity of  $G(\cdot)$ ), a solution,  $\hat{n}$ , exists and is unique. Furthermore, it is easy to verify that  $\hat{\Pi}''(\hat{n}) < 0$ , so that  $\hat{n}$  is indeed the optimal team size.

Implicit differentiation in Equation (4) yields our first key result:

**Proposition 1:** The optimal team size decreases as either worker risk aversion  $r$ , worker cost of effort  $c$ , or uncertainty in performance measurement  $\sigma^2$  rises.

The ability to adjust team size optimally in turn alters the optimal adjustments in incentives. We now examine how incentive strength  $\hat{b}$  reacts to changes in uncertainty  $\sigma^2$ , worker risk aversion  $r$ , and worker cost of effort  $c$  when team size is adjusted optimally, compared to the case when team size is exogenous. Let us denote any of the symbols  $r$ ,  $c$ , or  $\sigma^2$  by  $v$ . Suppose the current team size is  $n_0$ . If the team size were determined exogenously, then the effect of changes in  $v$  would be given by:

$$\delta_{\text{ex}}^v(n_0) \equiv \left. \frac{\partial \hat{b}}{\partial v} \right|_{n=n_0}, \tag{5}$$

where  $\hat{b}$  is given by Equation (3). If team size is chosen optimally, the effect of changes in  $v$  is given by:

$$\delta_{\text{end}}^v(n_0) \equiv \left. \frac{\partial \hat{b}}{\partial v} \right|_{n=n_0} + \left. \frac{\partial \hat{b}}{\partial n} \right|_{n=n_0} \cdot \left. \frac{\partial \hat{n}}{\partial v} \right|_{\hat{n}=n_0}, \tag{6}$$

where  $\partial \hat{n} / \partial v$  is obtained by implicit differentiation in Equation (4). We show next that when team size is adjusted efficiently in reaction to changes in external factors, the directional impact on the incentive strength offered to workers is unaltered, but the magnitude of the impact is dampened.

**Proposition 2:** For any set of parameter values  $(\sigma^2, r, c)$ ,  $\delta_{\text{ex}}^v(\hat{n}) < \delta_{\text{end}}^v(\hat{n}) \leq 0$ .

Stated differently,  $|\delta_{\text{end}}^v(\hat{n})| < |\delta_{\text{ex}}^v(\hat{n})|$ , i.e., pay-for-performance sensitivity under optimally chosen team size is *always* lower than it would be if the team size were given exogenously. In contrast to the standard agency model, the firm here has two instruments: team size and incentives. An increase in  $r$ ,  $c$ , or  $\sigma^2$  results in a direct decrease in  $\hat{b}$  and an indirect increase in  $\hat{b}$  through the decrease in  $\hat{n}$ . Therefore, the direct effect of a decrease in incentives in response to a change in environmental parameters is partially offset by an increase in incentives due to the indirect effect of a decrease in team size (incentives being decreasing in team size.)

An important special case of the attenuation of incentives occurs when  $G(\cdot)$  is a power function,  $G(n) = An^\gamma$  (with  $A > 0$  and  $\gamma > 1$ ). In this case the pay-for-performance sensitivity is not just attenuated, but eliminated entirely.

**Proposition 3:** The optimal incentive weight  $b$  does not depend on parameter values  $(\sigma^2, r, c)$  if and only if  $G(\cdot)$  is a power function.

That is, incentives exhibit an *invariance property* when  $G(\cdot)$  is a power function: the optimal incentives do not vary at all with  $\sigma^2$ ,  $r$ , or  $c$ . The firm responds to changes in these parameters by adjusting the team size and keeping the incentives fixed.

To understand the intuition behind this result, suppose first that  $n$  is exogenous, and  $\sigma^2$  increases. From Equation (3),  $\hat{b}$  decreases, since the firm cuts incentives as the risk premium rises. As the incentives fall, so do the profits. Now if the firm decreases the team size, a smaller  $n$  reduces the risk premium and hence lifts the incentives. This rise in incentives offsets the fall in profits from the smaller team size. When  $G(\cdot)$  is a power function, the firm decreases the team size *exactly* to offset the increase in uncertainty, leaving the incentives constant. Essentially, team size is a more flexible instrument for the firm than are incentives.

If  $G(\cdot)$  is a power function, then the optimal team size can be derived in closed-form as:

$$\hat{n} = \left( \frac{1}{(\gamma - 1) \text{Arc}\sigma^2} \right)^{1/\gamma}. \quad (7)$$

Substituting this expression into the equation for the workers' incentives yields the equilibrium strength of incentives and total profit at the optimal team size:

$$\hat{b}(\hat{n}) = q \cdot \left( \frac{\gamma - 1}{\gamma} \right) \text{ and } \hat{\Pi}(\hat{n}) = \frac{\hat{n}q^2(\gamma - 1)}{2c\gamma}. \quad (8)$$

From Equation (8), the invariance property is readily observable: the optimal incentive weight,  $\hat{b}(\hat{n})$ , is independent of  $\sigma$ ,  $r$ , and  $c$ .

Before ending our discussion of the benchmark model, we comment on the robustness of the above findings. Two aspects of our formulation invite special examination—the use of a single group measure to compensate each worker, and the specification of an additively separable production function. Similar results (e.g., the attenuation and invariance findings, and the impact of external factors on team size) continue to hold when each of these assumptions is altered. In particular, the results generalize immediately to a setting where each worker is compensated using his own *individual performance signal*, as opposed to the common group measure  $y$ . They also extend to a scenario where the production technology exhibits complementarities among workers, as opposed to the linear technology with uniform marginal productivity ( $q$ ) in the benchmark model.<sup>5</sup>

### Adding a Manager

Now we introduce a manager to the benchmark setting. The manager serves the role of actively and endogenously supplying the monitoring activities (as opposed to a mechanical monitoring system). Endogenous monitoring alters the economic trade-off on team size and worker incentives (and thus welfare), but we show that the attenuation result continues to hold when team size is endogenously chosen.

We model the manager as an effort- and risk-averse economic agent who can produce a signal about the collective efforts of the team under his supervision. In particular, the manager exerts monitoring effort  $m$  at personal cost  $C(m) = k \cdot m$ , where  $k$  is the manager's cost of effort. The parameter  $k$  measures the manager's quality/ability: higher quality (or lower cost) managers have a lower  $k$ . The manager's effort reduces the variance on the team performance measure. So, while we use the term manager, we could equivalently refer to a monitor or an internal auditor, as the function of such an individual is to improve the quality of the measurement of performance of other workers within the firm. Specifically, we assume:

<sup>5</sup> Details of these results are available from the authors on request. More generally, our particular modeling choice (i.e., placing the team size effect on the variance of the performance measure) is, in fact, fairly general. One can show that placing the team size effect on worker-effort costs, performance measure sensitivity, or the marginal benefit of effort are all equivalent ways of representing the model. We are grateful to one of our reviewers for alerting us to this point on scaling, and for demonstrating it mathematically.

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

Task-specialization is such that the manager's effort reduces the variance of  $y$  while the worker's effort increases its mean.<sup>6</sup> Assume  $m$  is unobservable so the manager must be motivated to supply monitoring efforts. The firm offers a linear contract to the manager consisting of a salary and a bonus on the team performance measure,<sup>7</sup> i.e.,  $\omega = \alpha + \beta y$ . The manager thus solves:

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

Even though all workers and the manager are under linear contracts, the provision of incentives to exert effort comes through in different ways. The manager's effort does not affect the expectation of his wage, but such effort does decrease its variance. Given the manager is risk-averse, he has an incentive to reduce the variance. The first-order condition yields the incentive constraint for the manager:

$$m^* = \beta \sigma \left( \frac{G(n)r}{2k} \right)^{1/2}. \quad (9)$$

The manager'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  $m$ .<sup>8</sup>

From Equation (9), we learn that, for a given incentive coefficient  $\beta$ , if the measurement process worsens ( $\sigma^2$  increases) or the manager becomes more risk-averse ( $r$  increases), the risk premium grows and hence the manager works harder at monitoring. Moreover, as the manager's span of control grows ( $n$  increases) the manager also works harder. Adding workers boosts the variance of  $y$  and hence increases the risk premium. Because the manager dislikes variance, he spends more time monitoring to compensate for the additional risk cost. However, the manager's incentive coefficient  $\beta$  is endogenous, and depends on the parameters  $\sigma$  and  $r$ . We return to address the overall effect on equilibrium monitoring efforts later (in Proposition 5).

As before, the firm sets the salary levels ( $a_i, \alpha$ ) such that the individual rationality constraint binds for every worker and the manager. Assume the opportunity wage for the

<sup>6</sup> By this modeling choice, we choose to emphasize the monitoring role of managers, leaving their other functions (such as fulfilling direct productive tasks) as second-order effects. Within this monitoring role, the manager can, for example, exert effort to coordinate multiple agents within the team or facilitate communication across agents. As a reduced form, we model this as decreasing the variance on the team performance measure, as opposed to explicitly framing the exact consequence of the managerial action, in order to keep the level of complexity manageable.

<sup>7</sup> We believe this is a simple and natural way to model the compensation arrangement between the principal and the manager. In practice, managers' compensation is, in large part, determined by the performance of the group under their supervision; see, for example, Bushman et al. (1995). Providing managerial incentives via the signal produced by the manager represents a sharp contrast to the model in Ziv (2000), where the firm provides incentives to managers via a signal about managerial effort that is generated by yet another manager or, ultimately, the principal. While Ziv's (2000) model potentially allows for the investigation of multi-layer hierarchies, the imposed complexity precludes development of a tractable framework.

<sup>8</sup> 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.

manager is  $\bar{u}_M$ , which is also normalized to zero. The firm therefore maximizes total surplus, as before:

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

where:

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

Substituting in each worker's incentive constraint gives:

$$\max_{\beta, b_i} \sum_{i \in N} \left[ \frac{qb_i}{c_i} - \frac{b_i^2}{2c_i} - \frac{r}{2} b_i^2 \frac{G(n)\sigma^2}{m^*} \right] - km^* - \frac{r}{2} \beta^2 \frac{G(n)\sigma^2}{m^*}, \tag{10}$$

where  $m^*$  is given by Equation (9). The first-order conditions yield the optimal incentives for the worker:

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

Similar to the earlier setting with a mechanical monitor, 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 manager's effort ( $m$ ) matters in choosing optimal incentives for workers. Of course, the firm affects  $m$  by its choice of the manager's incentive  $\beta$ .

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

$$\left[ \frac{r}{2} \frac{G(n)\sigma^2}{m^{*2}} \sum_{i \in N} b_i^2 - k + \frac{r}{2} \frac{G(n)\sigma^2}{m^{*2}} \beta^2 \right] \frac{\partial}{\partial \beta} m^* - r\beta \frac{G(n)\sigma^2}{m^*} = 0. \tag{11}$$

Providing more managerial incentive (i.e., increasing  $\beta$ ) induces more managerial effort (see Equation (9)). The benefit is two-fold. First, it reduces the variance of the performance measure  $y$  and thus the risk-premium needed for the manager because the manager's contract is written on  $y$ . Second, increasing  $\beta$  reduces the 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 managerial monitoring effort. The manager is self-interested and wants to exert effort to reduce the variation in his own compensation. But in doing so, he reduces the wage variances of all workers. This is beneficial to the firm because it reduces salary costs ( $\alpha_i^*$  and  $\alpha^*$ ). These two (marginal) benefits are shown as the two positive terms in Equation (11). The cost of providing managerial incentives is also two-fold. First, a higher incentive leads to more personally costly managerial effort, for which the manager requires compensation. Second, a higher incentive leads to managerial pay that is more sensitive to

the signal variations, which increases the risk-premium. These two (marginal) costs are shown as the two negative terms in Equation (11).<sup>9</sup>

Substituting the manager’s incentive constraint (9) into Equation (11), and rearranging terms yields:

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

The manager’s incentives thus increase with each worker’s incentives. Intuitively, as the firm increases  $b_i$ , it induces more effort from the workers but also increases the risk premium. To compensate, the firm increases  $\beta$ , which induces the manager to spend more effort monitoring the workers; in turn, this reduces the risk premium for everyone.

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

$$b_i^*(n) = \frac{q}{1 + 2c_i \sqrt{\frac{rG(n)\sigma^2 k}{\sum_{j \in N} b_j^{*2}}}} \tag{12}$$

Despite the fact that Equation (12) is an implicit function of  $b_i$ , it can be shown, as in the setting with a mechanical monitor, that the strength of incentives ( $b_i$ ) falls as the cost of effort ( $c_i$ ) increases.

**Optimal Incentives and Team Size with a Manager**

Assuming identical workers ( $c_i = c$ ) leads to  $b_i^* = b^*$ , which in turn allows us to solve Equation (12) and derive a closed-form solution for the equilibrium worker incentive for a given  $n$ :

$$b^*(n) = q - 2c\sigma\sqrt{rkG(n)/n}. \tag{13}$$

Recall that the equilibrium manager incentive ( $\beta^*$ ), monitoring effort ( $m^*$ ), and worker effort ( $e^*$ ) are all functions of  $b^*$ .<sup>10</sup> From Equation (10), we can now write the profit function with a manager for a given  $n$ , denoted by  $\Pi(n)$ , as follows:

<sup>9</sup> Consider a slight alteration of the model where the manager exerts a productive effort, say  $e_m$ , which, similar to the workers’ efforts  $e_i$ , increases the output and the aggregate performance measure  $y$ . 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 (11). 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 trade-off in the managerial monitoring activities, which is the focus of our paper, persists and leads to the same qualitative results.

<sup>10</sup> 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.

$$\begin{aligned} \Pi(n) &= n \left[ \frac{qb^*}{c} - \frac{b^{*2}}{2c} - \frac{r}{2} b^{*2} \frac{G(n)\sigma^2}{\frac{b^*\sigma}{2} \sqrt{\frac{rnG(n)}{k}}} \right] - 2k \frac{b^*\sigma}{2} \sqrt{\frac{rnG(n)}{k}} \\ &= \frac{n(b^*(n))^2}{2c}. \end{aligned} \tag{14}$$

Now we characterize the optimal  $n$  in this setting.

**Lemma 1:** Assume  $c_i = c$  for all  $i \in N$ , and the principal hires a manager with ability  $k$ . There exists a unique optimal team size,  $n^*$ , which is characterized by the following implicit function:

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

where  $h(n) = \sqrt{G(n)/n}$ .

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

**Proposition 4:** The optimal team size is decreasing in measurement uncertainty  $\sigma$ , worker risk aversion  $r$ , and worker and manager costs of effort,  $c$  and  $k$ , respectively.

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

We next study whether the ability to adjust team size affects the strength of incentives, as in the mechanical monitoring setting. Let  $v$  denote any of the parameters  $\sigma$ ,  $c$ ,  $\sqrt{r}$ , or  $\sqrt{k}$ . The impact on the workers' incentive parameter,  $b$ , of factor  $v$  is then given by:

$$\varphi_{\text{ex}}^v(n_0) \equiv \left. \frac{\partial b^*}{\partial v} \right|_{n=n_0}$$

and:

$$\varphi_{\text{end}}^v(n_0) \equiv \left. \frac{\partial b^*}{\partial v} \right|_{n=n_0} + \left. \frac{\partial b^*}{\partial n} \right|_{n=n_0} \cdot \left. \frac{\partial n^*}{\partial v} \right|_{n^*=n_0},$$

for the exogenous and the endogenous team-size cases, respectively. Since the monitoring is now performed by a self-interested manager, an immediate question is whether a similar comparison can be performed on the strength of the incentive,  $\beta$ , offered to the manager.

We refer to the corresponding constructs for the manager as  $\psi_{\text{ex}}^v(n_0)$  and  $\psi_{\text{end}}^v(n_0)$ , respectively. Finally, it is also of interest to identify the relative impact of parameter changes on the manager's and workers' equilibrium levels of pay-for-performance. The following result provides a full characterization of these issues.

**Proposition 5:** For any set of parameter values  $(\sigma^2, c, r, k)$ , and for a change in any factor  $v = \sigma, c, \sqrt{r}$ , or  $\sqrt{k}$ :

$$\psi_{\text{ex}}^v(n^*) = \psi_{\text{end}}^v(n^*) \leq \varphi_{\text{ex}}^v(n^*) < \varphi_{\text{end}}^v(n^*) \leq 0,$$

with a strict inequality in the middle for any  $n^* > 2$ .

There are several implications embedded in the above string of inequalities. First, as in the mechanical monitor setting,  $|\varphi_{\text{end}}^v(n^*)| < |\varphi_{\text{ex}}^v(n^*)|$ , i.e., the introduction of team size as a choice variable unambiguously attenuates worker incentives. An increase in  $\sigma, c, r$ , or  $k$  results in a decrease in the strength of incentives offered to the workers, regardless of whether team size is exogenously given or optimally adjusted. In the latter case, however, changes in any of these exogenous factors leads to a smaller change in the  $b$  offered to workers, i.e., the impact of variations in the external environment on the optimal contract offered to workers is dampened.

Second, we find, rather surprisingly, that no such attenuation occurs for the manager's incentives. Starting from the efficient team size, the variation in the manager's incentive rate when any external factor is changed is identical regardless of whether team size is adjusted to the new optimal level. Moreover, the decrease in the manager's incentive coefficient for an increase in any factor,  $v$ , is greater than the corresponding drop in the workers' incentive coefficient. We thus find that, even though both workers and managers are risk-averse, the manager's incentives are unambiguously more sensitive to changes in exogenous parameters (uncertainty, risk aversion, or costs of effort) than are the incentives of the workers. Our model thus yields the testable prediction that one should empirically observe greater variation in the incentives for supervisors than those for production workers.

The intuition for the result is as follows. We know from the previous section that  $\beta^* = b^*\sqrt{n}/2$ . Thus, parameters affect the manager's incentives via two channels: through their effects on  $b^*$  and through their effects on the team size. The effect of the parameters on  $\beta^*$  is stronger than that on  $b^*$  whenever the effects on  $b^*$  and  $n$  work in the same direction. Proposition 4 shows that an increase in any of the parameters  $\sigma, c, r$ , or  $k$  causes the firm to shrink the size of the team. Thus, whenever the firm reduces  $b^*$  in response to these parameter changes, the effects of the parameters on  $b^*$  and  $n$  amplify each other; as a result,  $\beta^*$  responds more than does  $b^*$ .

We next show that the extreme version of attenuation, i.e., the invariance property for workers' incentives, continues to hold for power functions even in the presence of a manager.

**Proposition 6:** The optimal  $b$  does not depend on parameter values  $(\sigma, c, r, \text{ or } k)$  if and only if  $G$  is a power function,  $G(n) = An^\gamma$ .

A different way of stating this result is as follows: the weak inequality,  $\varphi_{\text{end}}^v(n^*) \leq 0$ , in the statement of Proposition 5 is strict unless  $G(n)$  belongs to the class of power functions. At an intuitive level, we know from Proposition 3 that the invariance property holds in this new setting if and only if  $G(n)/m^*$  is a power function. But, by Equation (9),  $m^*(n)$

is a power function if and only if  $G(n)$  is a power function. Consequently,  $G(n)/m^*$  is a power function if and only if  $G(n)$  is.

Propositions 5 and 6 together imply that if  $G(\cdot)$  is a power function, then the manager's incentives respond to parameter changes, while the incentives of the workers do not. To see this explicitly, consider the following closed-form solutions for the equilibrium levels of the workers' and manager's incentives when team size is chosen optimally:

$$b^*(n^*) = q \cdot \left( \frac{\gamma - 1}{\gamma} \right) \quad (15)$$

and:

$$\beta^*(n^*) = \frac{q}{\sqrt{2}} \left( \frac{\gamma - 1}{\gamma} \right) \left( \frac{q}{2\gamma c \sigma \sqrt{A r k}} \right)^{1/(\gamma-1)}. \quad (16)$$

Recall that, in general,  $\beta^* = b^* \sqrt{n/2}$ , and  $n$  is decreasing in  $\sigma$ ,  $c$ ,  $r$ , and  $k$ . Since  $b^*$  is unaffected by exogenous parameters when  $G(\cdot)$  is a power function, the combined effect is that  $\beta^*$  decreases in all of the parameters. The result seems counterintuitive, as one might expect the firm to raise incentives for the manager to monitor the workers more precisely if, for example, the measurement process worsens ( $\sigma$  rises). Instead, the firm shrinks the size of the team to respond to the increase in risk premium. Smaller teams require less monitoring and, hence, the firm can afford to reduce the incentives for the manager. The firm is once again using team size as the main instrument to respond to exogenous changes in the environment.

The findings in Propositions 2, 3, 5, and 6 provide a possible explanation for the lack of strong empirical support for the negative relation between risk and incentives suggested by traditional agency models (see Prendergast 2002), who suggests that firms may contract directly on employee effort in some settings, and contract on output in less certain settings. Under this construct, the strength of incentives given to employees need not vary monotonically with the level of uncertainty. Our model differs in that we assume that employee effort is unobservable and that contracts with employees can only be structured on observed joint output. In effect, while the Prendergast (2002) result points to the omission of delegation as a choice variable in traditional agency models, our finding is based on the ability of the principal to control team size in addition to employee incentives. To this extent, our model does in fact embrace a quite large class of organizations, since we consider any organizations where teams are key to the production process. In contrast, the existing empirical literature on the risk/incentive trade-off has looked only within a handful of narrow environments (e.g., executive compensation, sharecropping, sports tournaments).

### Value of the Manager

Following our analyses of the mechanical monitoring and self-interested manager settings, we conclude this section by identifying when it is profitable for the firm to hire a manager. Suppose that firm size is exogenously given, and that  $c_i = c$  for all  $i \in N$ . Assume the principal has the choice of either using a fixed monitor/information system or hiring a manager with known ability  $k$ . Then for any given set of parameters  $(q, r, c, \sigma^2)$ , the firm compares the expected profit under the two alternatives. The value of hiring a manager is given by:

$$\begin{aligned} \Pi(n) - \hat{\Pi}(n) &= \frac{n}{2c} \left[ \left( q - 2c\sigma\sqrt{rkG(n)/n} \right)^2 - \frac{q^2}{1 + rcG(n)\sigma^2} \right] \\ &= \frac{\sigma}{2} \sqrt{rG(n)} \left[ q\sqrt{n} \left[ \frac{q\sigma\sqrt{rmG(n)}}{1 + rcG(n)\sigma^2} - 4\sqrt{k} \right] + 4c\sigma\sqrt{rkG(n)} \right]. \end{aligned}$$

It follows immediately that a sufficient condition for the owner to find it in his interest to hire a manager with ability  $k$  is:

$$\frac{q\sigma\sqrt{rmG(n)}}{1 + rcG(n)\sigma^2} \geq 4\sqrt{k}. \tag{17}$$

As one might expect, the manager is valuable if his ability is sufficiently high: lower values of  $k$  imply fewer difficulties in motivating an effort- and risk-averse manager to work. In addition, the expression in Equation (17) shows that employing a manager is unambiguously worthwhile when the market value of output,  $q$ , is high. More interestingly, an analysis of the profit difference reveals that it is more worthwhile for the owner to hire a manager when the workers are of higher quality (i.e.,  $c$  is lower). We explore more aspects of this complementarity between workers and managers in the next section.

The results with regard to the level of uncertainty in monitoring technology are less straightforward. In particular, it can be shown that the manager is not valuable when  $\sigma^2$  is either too low (since a mechanical monitor can then do just as well) or too high (because of the high cost of motivating the manager and paying him a risk premium in that case). However, for all intermediate values of the uncertainty parameter,  $\sigma^2$ , the owner finds it beneficial to employ a manager. This suggests that if the precision of the performance measurement increases (i.e.,  $\sigma$  decreases) over the life cycle of a firm, then the firm would operate without a monitoring manager initially, employ a monitoring manager when  $\sigma$  has decreased enough, and substitute in a mechanical monitor at a later, more mature stage of firm's life cycle when the performance measurement system is sufficiently precise.

When firm size is endogenous, then for a given set of parameters, the firm compares the expected profit under the mechanical monitor at the optimal team size and the expected profit under a hired manager at the (possibly different) optimal team size. While the economic intuition stated in the previous paragraph still holds, no general expression for the profit differential can be derived because of the implicit functions governing the optimal team sizes. We can, however, obtain additional insights by looking at the case where  $G(n) = An^\gamma$ . In this setting, the value of hiring a manager is:

$$\Pi(n^*) - \hat{\Pi}(\hat{n}) = \frac{q^2(\gamma - 1)}{2c\gamma} \cdot \left( \frac{n^*(\gamma - 1)}{\gamma} - \hat{n} \right).$$

As  $\gamma > 1$ , it follows that it is profitable to hire a manager if and only if the optimal team size with a manager exceeds the optimal team size without the manager. Managers assist in team performance measurement, and hence allow the formation of larger teams. The link between valuable managers and team sizes points to a synergy between the use of managers and the number of people assigned to them for supervision. What is interesting here is that the synergy (or complementarity) is not assumed of the production technology, but is due to the inherent moral hazard problem associated with team production.

To summarize, in this model, the firm has three design instruments to maximize its expected profit. They are: (1) the number of workers ( $n$ ), (2) the choice of whether to hire a manager, and (3) the choice of incentive coefficients for the workers and manager ( $b$  and  $\beta$ ). When the environment changes, as reflected by a variation in ( $q$ ,  $\sigma$ ,  $c$ ,  $r$ , or  $k$ ), all three design variables may change in response. The special property of the responses is a certain robustness in the worker-incentives; in particular,  $b^*$  responds less, and sometimes not at all, to changes other than in the market price of output,  $q$ , relative to the case when team size is fixed. For changes in all other environmental variables, the firm reacts more by changing team size and monitoring levels, and less by changing workers' incentives.

### III. COMPLEMENTARITY BETWEEN MANAGERS AND WORKERS

In the previous section, we demonstrated the presence of endogenous synergies between team size and managerial supervision. In this section, we explore more such aspects of optimal incentives and monitoring in team settings. Of specific interest to us is the link between worker and manager quality. Recall that the parameters  $c$  and  $k$  in our model reflect the costs of effort for the workers and the manager; better workers and managers have lower  $c$ 's and  $k$ 's. This section explores complementarity, in the sense of optimal matching, between managers and workers. That is, should better managers be asked to supervise better workers or worse ones? We show that it is always optimal to do the former, i.e., manager and worker abilities are complements. We then explore how the selection of the optimal manager and worker combination in a labor market varies with uncertainty and other parameters of our model. Throughout the section, we will restrict our attention to the case when  $G(\cdot)$  is a power function,  $G(n) = An^\gamma$ , with  $A > 0$ ,  $\gamma > 1$ .

#### Positive Sorting

When  $G(\cdot)$  is a power function, the closed-form solutions for the workers' and manager's incentive contracts are given in Equations (15) and (16). The solutions for the other variables of interest are provided below:

$$n^* = \left( \frac{q}{2\gamma c \sigma \sqrt{Ark}} \right)^{2/(\gamma-1)}; \tag{18}$$

$$m^* = \frac{q\sigma}{2} \left( \frac{\gamma-1}{\gamma} \right) \sqrt{\frac{Ar}{k}} \left( \frac{q}{2\gamma c \sigma \sqrt{Ark}} \right)^{(\gamma+1)/(\gamma-1)}; \tag{19}$$

$$\Pi^* = \frac{n^*(b^*)^2}{2c} = \frac{q^2}{2c} \left( \frac{\gamma-1}{\gamma} \right)^2 \left( \frac{q}{2\gamma c \sigma \sqrt{Ark}} \right)^{2/(\gamma-1)}. \tag{20}$$

From Equation (20), note that the overall profit function is of the form

$$\Pi^* = K_0 k^{-1/(\gamma-1)} c^{-(\gamma+1)/(\gamma-1)}, \tag{21}$$

where  $K_0 > 0$  is independent of  $k$  and  $c$ . Taking derivatives of Equation (21) shows that:

$$\frac{\partial \Pi^*}{\partial c} < 0 \quad \text{and} \quad \frac{\partial \Pi^*}{\partial k} < 0. \tag{22}$$

That is, better workers and managers raise profits. What drives this result? It is tempting

to argue that it is easier for better workers to deliver a given level of effort, so the firm reduces their incentives, thus reducing the risk premium payments and, hence, raising profits. Similarly, it seems plausible that because better managers are more effective at reducing variance, this will also lead to lower risk premiums and higher profits. However, this logic neglects the main point of this paper: that the firm is simultaneously adjusting the team size, and this can reverse the usual comparative statics. To see this, note that in equilibrium  $m^*$  is increasing in  $n^*$ , which in turn is decreasing in  $k$ . A better manager allows the firm to increase team size, which causes the manager to work harder in equilibrium. Recall that, in general:

$$\text{Var}(w_i) = b_i^2 \frac{G(n)\sigma^2}{m} \quad \text{and} \quad \text{Var}(\omega) = \beta^2 \frac{G(n)\sigma^2}{m}, \tag{23}$$

so more managerial effort will pull wage variance downward. But the larger team size counteracts this effect. In fact, it turns out that for any power function the variance-increasing team-size effects of  $c$  and  $k$  dominate their direct variance-decreasing managerial effort effects. Substituting the optimal  $b^*$  and  $\beta^*$  from Equations (15)–(16) and the optimal  $n^*$  and  $m^*$  from Equations (18)–(19) into Equation (23) shows that, for  $G(n) = An^\gamma$ :

$$\text{Var}(w_i) = K_1 c^{-1} \quad \text{and} \quad \text{Var}(\omega) = K_2 c^{-(\gamma+1)/(\gamma-1)} k^{-1/(\gamma-1)}, \tag{24}$$

where  $K_1, K_2 > 0$  are constants independent of  $c$  and  $k$ . We see that better workers and managers (lower  $c$ 's and  $k$ 's) in fact increase wage variance. In particular, observe that the risk premium:

$$RP = \frac{nr}{2} \text{Var}(w) + \frac{r}{2} \text{Var}(\omega),$$

increases with better workers and managers as well. We know that  $n$  decreases in  $c$  and  $k$  by (18), while  $\text{Var}(w)$  decreases in  $c$  and  $\text{Var}(\omega)$  decreases in  $c$  and  $k$  by Equation (24).

But if the risk payments are increasing, how do profits rise? Remember that larger teams not only inflate the risk premium, but also earn more production revenue for the firm. The firm maximizes total surplus, which (for identical workers) is:

$$TS = n \cdot \{qe - C(e)\} - [C(m) + RP].$$

Suppose the firm hires a better manager. This does not change the worker's effort choice, so the term in braces is unchanged. As shown above, the manager works harder and the risk premium rises; as a result, the term in brackets rises, pulling down profits. Finally, the firm hires a larger team, so the revenue from production lifts profits upward. The net effect is that the benefits from higher production exceed the costs of more managerial and more risk payments, so profits overall rise. A similar argument holds for better workers.

Taking cross partial derivatives of Equation (21), we see that:

$$\Pi_{ck}^* \equiv \frac{\partial^2 \Pi^*}{\partial c \partial k} = \frac{\gamma + 1}{(\gamma - 1)^2} K_0 k^{-\gamma/(\gamma-1)} c^{-2\gamma/(\gamma-1)} > 0, \tag{25}$$

which gives positive sorting: it is optimal to assign better managers to better workers.

Precisely, the condition above shows that  $\partial\Pi^*/\partial c$  increases in  $k$  and  $\partial\Pi^*/\partial k$  increases in  $c$ . Now  $-\partial\Pi^*/\partial c$  is the marginal return from hiring a better worker. So, as the quality of the manager improves (i.e.,  $k$  falls),  $-\partial\Pi^*/\partial c$  increases, so the marginal benefit from hiring a better worker rises. Said differently, better workers raise profits, and in fact raise profits more under better managers.

To see this visually, refer to Figure 1. The two curves represent profit as a function of  $k$  for different costs of worker effort. Observe that both curves are decreasing: as the quality of the manager improves ( $k$  decreases), the firm earns more profit. Equation (25) says that  $\Pi_k^*(\cdot)$  increases in  $c$ . So as  $c_2$  rises to  $c_1$ ,  $\Pi_k^*(c_2, k)$  rises to  $\Pi_k^*(c_1, k)$ . Since  $\Pi_k^*(\cdot) < 0$ , this means  $\Pi_k^*(c_1, k)$  is less negative than  $\Pi_k^*(c_2, k)$ , i.e., the profit function is flatter in  $k$  when  $c = c_1$ . In other words, as  $k_1$  falls to  $k_2$ , profits will rise. But the increase in profit under the better worker ( $\Delta_2$ ) exceeds the increase in profit under the worse worker ( $\Delta_1$ ). So as the firm hires a better manager, the firm earns more profit by hiring a better worker as well.

This shows that good managers and good workers are complements. The intuition once again lies in the optimal team size. From Equation (18), it follows that the cross partial of  $n^*$  with respect to  $c$  and  $k$  is positive. With better workers and better managers, the firm can afford to increase team size; team size increases the most with good workers working under good managers. A good manager can more easily reduce the variance for a larger team, and even more so if the team consists of good workers. This complementarity between manager and worker quality at the team level translates to a complementarity at the profit level.

### Optimal Manager and Agent Quality

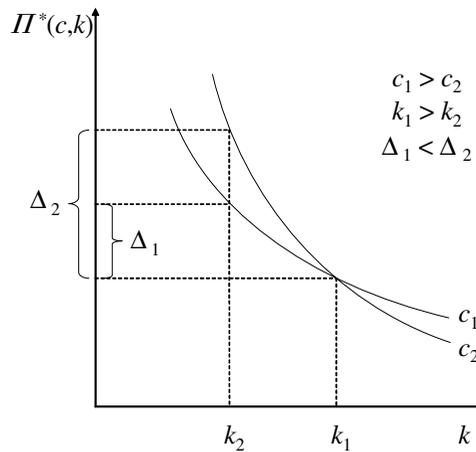
The previous subsection shows that the firm earns more profits by hiring better workers in conjunction with better managers. This begs the question: Why not hire the best workers and monitor possible? The reason, of course, is that employing high-quality talent is expensive. We have, as does virtually all of the compensation/agency literature, thus far assumed a partial equilibrium setting where the labor market is captured only via the assumed (exogenous) reservation wage. In this section, we introduce the idea of a labor market by explicitly specifying the reservation wage as a function of labor talent.<sup>11</sup> This enables us to draw inferences about the impact of firm-specific factors, as well as worker and manager characteristics, on the optimal team size and labor mix employed by organizations.

To model this in a parsimonious manner, we specify that the firm hires the manager and workers in distinct external labor markets. In particular, suppose that the quality  $k > 0$  of the manager and the quality  $c > 0$  of the worker are observable by both the firm and the market. A manager of quality  $k$  receives market wage  $\bar{u}_M(k)$ . Better managers (lower  $k$ ) earn higher market wages, so  $\bar{u}_M(k)$  is a positive and strictly decreasing function of  $k$ . Similarly, each identical worker receives a market wage of  $\bar{u}_W(c)$ , a positive and decreasing function of  $c$ .

It is clear that the firm must, and will, pay exactly these market wages to acquire its labor. As such, the firm sets the salaries ( $a, \alpha$ ) such that the workers' and manager's expected utilities precisely equal their market wages. Given binding (IR) constraints, the firm's profit function for exogenous  $n$  is then given by:

<sup>11</sup> Other papers that have modeled an outside labor market explicitly include Sappington (1983) and Ricart i Costa (1988). Huddart and Liang (2005) implicitly consider an outside labor market by characterizing partnerships that optimize per-partner utility, thereby making the reservation utility for any partner endogenous.

**FIGURE 1**  
Positive Sorting across Managers and Workers



The curves represent profit as a function of manager quality ( $k$ ) for different levels of worker quality ( $c$ ). Higher levels of managerial or worker quality are indicated by lower values of  $k$  or  $c$ , respectively. The curves denote positive sorting as the marginal benefit ( $\Delta$ ) from having a better manager ( $k_2 < k_1$ ) is greater when the worker is also of higher quality ( $c_2 < c_1$ ).

$$\Pi(n, c, k|\sigma, r) = \frac{nb^2}{2c} - n \cdot \bar{u}_W(c) - \bar{u}_M(k),$$

where  $b = b^* = q - 2c\sigma\sqrt{rkG(n)/n}$  is the optimal incentive payment for each worker (see Equation (13)).

Hiring a better quality manager (decreasing  $k$ ) allows the firm to increase the size of the team, and hence generate more revenue from production. At the same time, better managers are more expensive, since  $\bar{u}_M(k)$  increases as  $k$  decreases. The optimal  $k$  trades off the increased revenue from larger teams against the increased cost of a better quality manager, where this cost reflects the manager’s market wage. A similar trade-off holds for worker quality. The firm chooses the optimal worker and manager quality in addition to optimal team size by solving the following program:

$$\max_{n,c,k} \Pi(n, c, k|\sigma, r) \text{ such that } n, c, k > 0, \text{ and } b \geq 0. \tag{26}$$

We have the following result.

**Proposition 7:** When  $G(\cdot)$  is a power function, the optimal choices of  $b$  and  $c$  do not depend on  $\sigma$  or  $r$ .

The invariance property thus holds to an even greater extent in this setting. For exogenous changes in uncertainty or worker risk aversion, the firm optimally adjusts the team size and the manager quality, but keeps worker incentives and quality fixed. So invariance applies not just to incentives but to the quality of the worker as well. The inclusion of

additional choice variables in the firm's problem allows the firm to use these instruments to adjust to changes in its environment. Essentially, the team size and manager quality are more flexible instruments in making the adjustments, compared to strength of incentives and worker quality.

If the firm does not alter  $b$  or  $c$  after changes in  $\sigma$  or  $r$ , how does it change  $n$  and  $k$ ? As uncertainty or risk aversion increases, one might expect the firm to hire a better manager, since better managers by definition are more effective at reducing output variance. Instead, we have the following result:

**Proposition 8:** Suppose  $G(\cdot)$  is a power function. As uncertainty ( $\sigma$ ) or risk aversion ( $r$ ) increases, it is optimal to decrease the team size, hire a lower quality manager, and keep the quality of the workers fixed.

Once again, it is important to consider that in equilibrium the firm adjusts the team size optimally. As shown earlier, the firm will shrink the team in response to higher uncertainty. Because managers are costly, it is a waste of resources to assign an expensive, high-quality manager to a small team when a lower quality manager will suffice. Even though the expensive manager by himself can more effectively reduce output variance than a cheaper manager, the firm uses team size as the primary instrument to deal with the increase in uncertainty, and then selects the quality of the manager to fit the team.

When  $G(\cdot)$  is a power function, the parameter  $\gamma$  measures the quality of the performance technology. To see this, observe that when  $G(n) = An^\gamma$ , the output variance is:

$$\text{Var}(y) = \frac{G(n)}{m} \sigma^2 = \frac{A\sigma^2}{m/n^\gamma}.$$

As  $\gamma$  increases, the manager's effectiveness of monitoring ( $m/n^\gamma$ ) decreases, and hence the output variance increases. In other words,  $\gamma$  captures how quickly signal quality degrades as team size increases. The firm must of course bear the cost of this via higher risk payments. The next result characterizes how changes in the  $\gamma$  parameter affect the firm's choice of manager and worker quality.

**Proposition 9:** Suppose  $G(\cdot)$  is a power function. As the performance measurement technology improves ( $\gamma$  decreases), it is optimal to enlarge the team size, hire better quality managers, and hire lower quality workers.

Note that any given manager is more effective at reducing variance if  $\gamma$  is low. Naturally, more effective managers can handle a larger number of workers, so the firm expands the team size. But surprisingly, a change in  $\gamma$  affects the optimal quality of the workers and the manager in opposite directions: the firm will hire lower quality workers but better quality managers. Again, focus on the effects coming from the optimal team size. As the team size increases, the firm hires a better manager to handle the larger team. In this sense, the quality of the performance measurement system and the quality of the manager are complementary. Better managers are more effective at utilizing a high-quality performance measurement system than managers of lower ability.

We also find that the firm hires lower quality workers. This arises from the combination of two forces: first, high-quality workers are now costly, and second, the firm is already earning more revenue from larger teams. Therefore, the firm expands the size of the team

but lowers the quality of each team member. So team size and worker quality are substitutes in the firm's profit function, whereas team size and manager quality are complements. This arises directly from the nature of the monitoring production technology.

A noteworthy feature of the preceding comparative statics is that they require virtually no restrictions on the outside option functions. It is necessary only to impose that the functions are positive and strictly decreasing, but not to specify their functional form. The mere imposition of a cost of worker and manager quality suffices to generate rich comparative statics on the parameters of the model.

#### IV. CONCLUSION

In this paper, we study three instruments in the internal design of an organization involving team production: team size, monitoring activities, and the incentive contracts offered to workers and managers. When both types of agents (workers and managers) are self-interested and have the ability to take hidden actions, there exist complex interactions in the trade-offs among these instruments. We show, however, that such complex interactions produce rather simple and stark implications.

One such result is the robustness of worker contracts. We characterize the properties of attenuation and invariance whereby, compared to settings where team size is fixed, the equilibrium pay-for-performance scheme offered to workers varies less, if at all, with respect to most of the environmental variables of interest, including the extent of uncertainty, costs of effort, and risk preferences. By contrast, the incentive schemes offered to managers exhibit substantially higher levels of sensitivity to these parameters. Our work thus yields a clear prediction about the relative variation in the pay-for-performance sensitivities of employees in production versus supervisory roles in organizations.

We also characterize the optimal size of worker teams in various settings and provide testable predictions about the impact of environmental variables on the choice of optimal team size and structure, even in the presence of an external marketplace for talent. Of particular note, our work suggests the following implication: in settings with relatively lower levels of intrinsic uncertainty and/or better technology for measuring performance, team sizes will be higher and the quality of the workers in the teams will be either unchanged or lower, while the managers hired to oversee them will have greater levels of skill. The first of these implications is consistent with the finding by Rajan and Wulf (2006) of a significant flattening in the corporate hierarchy and a move to greater spans of control over the 1986–1998 time period, a shift they ascribe in part to improvements in information technology. More generally, Guadalupe and Wulf (2007) note that there is relatively little empirical evidence regarding the relationship between information technology and the structure of the internal hierarchy. The results in our paper provide several hypotheses that may be tested by future empirical research in this area.

Our theoretical exploration may be extended in at least three ways. First, we restrict attention to settings in which a single self-interested manager is responsible for monitoring a team of workers. It would be of great interest to extend the model framework to study the interaction between multiple managers who are asked to monitor a given worker team. Given the findings in Christensen and Sabac (2002), it is likely that the correlation in the managers' signals will prove pivotal in determining the value of having supervisory teams, and their impact on the optimal team size. This analysis could be expanded further, under alternative assumptions regarding other types of monitoring synergies, to explore ideas related to shirking in monitoring tasks or the assignment of managers to monitor different subgroups within a team. Another promising avenue for future study is the extension of our model to a multi-period setting. This would allow for analysis of the trade-offs inherent

in team incentives and monitoring in a dynamic setting, and under various scenarios involving the firm’s ability to commit to long-term incentive structures. Finally, all of our work has been carried out in a multi-agent LEN framework. We use this specification because it would be virtually impossible to replicate our analysis in a general agency model. However, it would be of interest to examine (perhaps via numerical simulations) whether our comparative static results, for example, hold in a broader range of settings.

**APPENDIX**

**Proof of Proposition 1:** Let  $v$  denote any of the symbols  $r, c, \sigma^2$ . Let  $\bar{v}$  denote the product of the other two (i.e., if  $v = r$ , then  $\bar{v} = c\sigma^2$ ). Implicitly differentiating (4) in  $v$  yields:

$$\frac{\partial \hat{n}}{\partial v} = -\frac{1}{\bar{v}\hat{n}G''(\hat{n})v^2} < 0. \tag{27}$$

■

**Proof of Proposition 2:** As before, let  $v$  denote any of the symbols  $r, c, \sigma^2$  and let  $\bar{v}$  denote the product of the other two. Differentiating (3) and applying (27), we obtain:

$$\delta_{\text{ex}}^v(\hat{n}) = \left. \frac{\partial \hat{b}}{\partial v} \right|_{n=\hat{n}} = -\frac{q\bar{v}G(\hat{n})}{(1 + rc\sigma^2G(\hat{n}))^2} < 0; \tag{28}$$

$$\delta_{\text{end}}^v(\hat{n}) = \left. \frac{\partial \hat{b}}{\partial v} \right|_{n=\hat{n}} + \left. \frac{\partial \hat{b}}{\partial n} \right|_{n=\hat{n}} \cdot \frac{\partial \hat{n}}{\partial v} = -\frac{q\bar{v}}{(1 + rc\sigma^2G(\hat{n}))^2} \left[ G(\hat{n}) - \frac{G'(\hat{n})}{rc\sigma^2\hat{n}G''(\hat{n})} \right]. \tag{29}$$

From (4), we can substitute for  $\frac{1}{rc\sigma^2}$  into the latter expression to obtain:

$$\begin{aligned} \delta_{\text{end}}^v(\hat{n}) &= -\frac{q\bar{v}}{(1 + rc\sigma^2G(\hat{n}))^2} \left[ G(\hat{n}) - \frac{G'(\hat{n})}{\hat{n}G''(\hat{n})} \cdot [\hat{n}G'(\hat{n}) - G(\hat{n})] \right] \\ &= -\frac{q\bar{v}G(\hat{n})}{(1 + rc\sigma^2G(\hat{n}))^2} \left[ \frac{G'(\hat{n})}{G''(\hat{n})} \cdot \left( \frac{G''(\hat{n})}{G'(\hat{n})} - \frac{G'(\hat{n})}{G(\hat{n})} + \frac{1}{\hat{n}} \right) \right] \leq 0, \end{aligned} \tag{30}$$

where the last weak inequality follows from the convexity of  $G(\cdot)$  and Condition (1). From (28) and (30), it follows that  $\delta_{\text{end}}^v(\hat{n}) > \delta_{\text{ex}}^v(\hat{n})$  if and only if:

$$\frac{G'(\hat{n})}{G''(\hat{n})} \cdot \left( \frac{G''(\hat{n})}{G'(\hat{n})} - \frac{G'(\hat{n})}{G(\hat{n})} + \frac{1}{\hat{n}} \right) < 1,$$

which reduces to:  $\hat{n}G'(\hat{n}) - G(\hat{n}) > 0$ , which holds for all  $\hat{n} > 0$  since  $G(0) = 0$  and  $G(\cdot)$  is convex. ■

**Proof of Proposition 3:** For the “if” part of the result, suppose that  $G(n) = An^\gamma$ ,  $A > 0$ ,  $\gamma > 1$ . Then, for arbitrary  $n$ , the expression in square brackets in (30) simplifies to:

$$\begin{aligned}
 1 - \frac{G'(n)^2}{G(n)G''(n)} + \frac{G'(n)}{nG''(n)} &= 1 - \frac{A^2\gamma^2n^{2\gamma-2}}{(An^\gamma) \cdot (A\gamma(\gamma - 1)n^{\gamma-2})} + \frac{A\gamma n^{\gamma-1}}{nA\gamma(\gamma - 1)n^{\gamma-2}} \\
 &= 1 - \frac{\gamma}{\gamma - 1} + \frac{1}{\gamma - 1} = 0.
 \end{aligned}
 \tag{31}$$

Therefore,  $\delta_{\text{end}}^v(n) = 0$  for all values of  $n$ , i.e., pay-for-performance sensitivity is unaffected by changes to the parameter values ( $\sigma^2$ ,  $r$ ,  $c$ ) once team size is adjusted optimally.

For necessity, note from Proposition 1 that changes to  $\sigma^2$ ,  $r$ , or  $c$  lead to variations in the optimal  $n$  via Equation (4). Therefore, if  $b$  is invariant to the parameters of the problem, then it must be the case that  $\delta_{\text{end}}^v(n) = 0$  for an open set of values of  $n$ . From (30), this implies that for all  $n$  in some interval:

$$\begin{aligned}
 \frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} &= 0 \\
 \Leftrightarrow \frac{d}{dn} \left( \ln \left[ \frac{nG'(n)}{G(n)} \right] \right) &= 0 \\
 \Rightarrow \frac{nG'(n)}{G(n)} &= \theta, \\
 \Rightarrow G'(n) - \frac{\theta}{n} \cdot G(n) &= 0
 \end{aligned}
 \tag{32}$$

for some constant  $\theta > 0$ . Solving the first-order linear differential equation in (32) reveals that  $G(\cdot)$  must belong to the class of power functions, thus completing the proof. ■

**Proof of Lemma 1:** From Equations (13) and (14), the owner’s profit function is:

$$\Pi(n) = \frac{n}{2c} (q - 2c\sigma\sqrt{rkG(n)/n})^2,
 \tag{33}$$

and this is defined over the relevant range where  $b^*(n) = q - 2c\sigma\sqrt{rkG(n)/n} \geq 0$ . Note

that  $h(n) = \sqrt{\frac{G(n)}{n}}$  is an increasing function of  $n$ . This implies  $n \in [0, \bar{n}]$ , where  $\bar{n}$  satisfies:  $q - 2c\sigma\sqrt{rkh(\bar{n})} = 0$ . Evaluating  $\Pi(n)$  at the two bounds, it is clear that  $\Pi(\bar{n})$ . Moreover:

$$\Pi(0) = \frac{1}{2c} \cdot \lim_{n \rightarrow 0} (nq^2 + 4c^2\sigma^2rkG(n) - 4qc\sigma\sqrt{rknG(n)}) = 0.$$

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 team size is given by:

$$\Pi'(n) = \frac{b^*(n)}{2c} (q - 2c\sigma\sqrt{rkh(n)} - 4nc\sigma\sqrt{rkh'(n)}).$$

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

$$h(n^*) + 2n^*h'(n^*) = \frac{q}{2c\sigma\sqrt{rk}}. \tag{34}$$

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

$$h'(n) = \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \left( \frac{1}{n} - \frac{G(n)}{G'(n)n^2} \right) \tag{35}$$

and:

$$h''(n) = \frac{\sqrt{n}G'(n)}{4n\sqrt{G(n)}} \left( \frac{2G''(n)}{G'(n)} + \frac{3G(n)}{n^2G'(n)} - \frac{G'(n)}{G(n)} - \frac{2}{n} \right). \tag{36}$$

Together, (35) and (36) imply that:

$$\begin{aligned} 3h'(n) + 2nh''(n) &= \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left( 2 \frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} \right) \\ &> \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left( \frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} \right) \geq 0, \end{aligned} \tag{37}$$

where the strict and weak inequalities follow from  $G''(\cdot) > 0$  and Condition (1), respectively. Therefore, a unique  $n^*$  exists satisfying (34). To verify the second-order conditions at this point, we differentiate  $\Pi'(n)$  with respect to  $n$  to obtain:

$$\begin{aligned} \Pi''(n) &= \frac{b^{*'}(n)}{2c} (q - 2c\sigma\sqrt{rk}h(n) - 4nc\sigma\sqrt{rk}h'(n)) \\ &\quad + \frac{b^*(n)}{2c} (-2c\sigma\sqrt{rk}h'(n) - 4c\sigma\sqrt{rk}h''(n) - 4nc\sigma\sqrt{rk}h''(n)). \end{aligned}$$

Evaluating this expression at  $n^*$ , we have:

$$\Pi''(n^*) = 0 + \frac{b^*(n^*)}{2c} (-2c\sigma\sqrt{rk})(3h'(n^*) + 2n^*h''(n^*)) < 0,$$

where the last inequality follows from (37) and  $b^*(n^*) > 0$ . We can therefore conclude that  $n^*$  is the globally optimal team size. ■

**Proof of Proposition 4:** Let  $v$  denote any of  $c$ ,  $\sigma$ ,  $\sqrt{r}$ , and  $\sqrt{k}$ , and let  $\bar{v}$  denote the product of the other three. Implicitly differentiating the condition for  $n^*$  in (34) shows that:

$$\frac{\partial n^*}{\partial v} = (3h'(n^*) + 2n^*h''(n^*))^{-1} \left( -\frac{q}{2\bar{v}v^2} \right) < 0,$$

as the first term is strictly positive from (37). ■

**Proof of Proposition 5:** As before, let  $v$  denote any of  $c$ ,  $\sigma$ ,  $\sqrt{r}$ , and  $\sqrt{k}$ , and let  $\bar{v}$  denote the product of the other three. First, differentiating (13), we immediately obtain:

$$\varphi_{\text{ex}}^v(n^*) = -2\bar{v}h(n^*) < 0. \tag{38}$$

Next, implicitly differentiating the first order-Condition (34) in Lemma 1, we have:

$$\frac{\partial n^*}{\partial v} = -\frac{q\bar{v}}{2(3h'(n^*) + 2nh''(n^*))(v\bar{v})^2} < 0 \text{ (from (37)).}$$

We can then compute  $\varphi_{\text{end}}^v(n^*)$  as:

$$\begin{aligned} \varphi_{\text{end}}^v(n^*) &= \frac{\partial b^*}{\partial v} \Big|_{n=n^*} + \frac{\partial b^*}{\partial n} \Big|_{n=n^*} \cdot \frac{\partial n^*}{\partial v} \Big|_{n=n^*} = -2\bar{v}h(n^*) \\ &\quad + \frac{2q\bar{v}h'(n^*)}{(3h'(n^*) + 2n^*h''(n^*))(2v\bar{v})}. \end{aligned}$$

Substituting for  $q/(2v\bar{v})$  using Condition (34), we obtain:

$$\varphi_{\text{end}}^v(n^*) = -2\bar{v} \left[ h(n^*) - \frac{h'(n^*)(h(n^*) + 2n^*h'(n^*))}{3h'(n^*) + 2n^*h''(n^*)} \right] \tag{39}$$

$$= -4\bar{v} \left[ \frac{h'(n^*)h(n^*) + n^*h(n^*)h''(n^*) - n^*h'(n^*)^2}{3h'(n^*) + 2n^*h''(n^*)} \right]. \tag{40}$$

The denominator of (40) is strictly positive (from (37)). Using Equations (35) and (36), the numerator can be simplified to:

$$\frac{G'(n)}{2} \cdot \left( \frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} \right) \geq 0, \tag{41}$$

where the inequality follows from Condition (1).

We have therefore established that  $\varphi_{\text{end}}^v(n^*) \leq 0$ . Moreover, using (37) and the fact that  $h'(\cdot) > 0$ , a direct comparison of (38) and (39) reveals that  $\varphi_{\text{ex}}^v(n^*) < \varphi_{\text{end}}^v(n^*) \leq 0$ .

Turning to the manager's incentives, we know that with identical workers:

$$\beta^* = b^* \sqrt{\frac{n}{2}}.$$

Therefore:

$$\psi_{\text{ex}}^v(n^*) = \sqrt{\frac{n^*}{2}} \cdot \varphi_{\text{ex}}^v(n^*) = -2\bar{v}h(n^*) \sqrt{\frac{n^*}{2}} < 0. \tag{42}$$

A comparison of (38) and (42) immediately reveals that  $\psi_{\text{ex}}^v(n^*) \leq \varphi_{\text{ex}}^v(n^*)$ , with strict inequality for all  $n^* > 2$ .

To complete the proof, we next demonstrate that  $\psi_{\text{ex}}^v(n^*) = \psi_{\text{end}}^v(n^*)$ . Since:

$$\psi_{\text{ex}}^v(n^*) = \psi_{\text{end}}^v(n^*) + \left. \frac{\partial \beta^*}{\partial n} \right|_{n=n^*} \cdot \left. \frac{\partial n^*}{\partial v} \right|_{n=n^*},$$

it is sufficient to show that  $\left. \frac{\partial \beta^*}{\partial n} \right|_{n=n^*} = 0$ . Using (13), we can expand  $\beta^*$  as follows:

$$\beta^* = b^* \sqrt{\frac{n}{2}} = \sqrt{\frac{n}{2}} \cdot (q - 2c\sigma\sqrt{rkG(n)/n}) = \sqrt{\frac{n}{2}} \cdot (q - 2v\bar{v}\sqrt{G(n)/n}).$$

Differentiating this expression, we obtain:

$$\begin{aligned} \left. \frac{\partial \beta^*}{\partial n} \right|_{n=n^*} &= \frac{q}{\sqrt{2}} \frac{1}{2\sqrt{n^*}} - \frac{v\bar{v}}{\sqrt{2}} \frac{G'(n^*)}{\sqrt{G(n^*)}} = \frac{v\bar{v}}{\sqrt{2n^*}} \cdot \left( \frac{q}{2v\bar{v}} - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}} \right) \\ &= \frac{v\bar{v}}{\sqrt{2n^*}} \cdot \left( h(n^*) + 2nh'(n^*) - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}} \right) \text{ (from (34))} \\ &= \frac{v\bar{v}}{\sqrt{2n^*}} \cdot \left( \sqrt{\frac{G(n^*)}{n^*}} + \frac{\sqrt{n^*}G'(n^*)}{\sqrt{G(n^*)}} - \frac{\sqrt{G(n^*)}}{\sqrt{n^*}} - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}} \right) = 0. \end{aligned}$$

■

**Proof of Proposition 6:** Suppose  $G(n) = An^\gamma$ ,  $A > 0$ ,  $\gamma > 1$ . Then, for arbitrary  $n^*$ , the numerator of  $\varphi_{\text{end}}^v(n^*)$ , given in reduced form in (41), simplifies to:

$$\begin{aligned} \frac{G'(n^*)}{2} \cdot \left( \frac{G''(n^*)}{G'(n^*)} - \frac{G'(n^*)}{G(n^*)} + \frac{1}{n^*} \right) &= \frac{G'(n^*)}{2} \\ \cdot ((\gamma - 1)(n^*)^{-1} - \gamma(n^*)^{-1} + (n^*)^{-1}) &= 0. \end{aligned}$$

Thus,  $\varphi_{\text{end}}^v(n^*) = 0$  everywhere, i.e., pay-for-performance is totally insensitive to changes in any of the four parameters of interest, ( $\sigma^2$ ,  $r$ ,  $c$ ,  $k$ ), provided team size is adjusted optimally.

For the “only if” argument, we have shown in Proposition 5 that changes to  $\sigma^2$ ,  $r$ ,  $c$ , or  $k$  lead to strict variations in the optimal  $n$  via Equation (34). Therefore, if  $b$  is invariant to any of these parameters, it implies that  $\varphi_{\text{end}}^v(n) = 0$  for an open set of values of  $n$ . From Equations (39)–(41), this implies that for all  $n$  in some open interval:

$$\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} = 0.$$

As in the proof of Proposition 3, this implies that  $G(\cdot)$  belongs to the class of power functions. ■

**Proof of Proposition 7:** When  $G(n)$  is a power function, we can write  $h(n) = \sqrt{G(n)/n} = Bn^\tau$ , where  $B$ ,  $\tau$  are arbitrary, positive real numbers. The objective function (see (26)) is then:

$$\Pi(n, c, k; \sigma, r) = \frac{n(q - 2c\sigma\sqrt{rk}Bn^\tau)^2}{2c} - n \cdot \bar{u}_w(c) - \bar{u}_M(k). \tag{43}$$

If we use  $b = q - 2c\sigma\sqrt{rk}Bn^\tau$  instead of  $n$  (change of variables), then:

$$n = \left(\frac{1}{B\sigma\sqrt{rk}}\right)^{1/\tau} \left(\frac{q - b}{2c}\right)^{1/\tau}.$$

Substituting into (43) and getting rid of  $n$ , we obtain the objective function:

$$\hat{\Pi}(b, c, k; \sigma, r) = \left(\frac{1}{B\sigma\sqrt{rk}}\right)^{1/\tau} \left(\frac{q - b}{2c}\right)^{1/\tau} \left(\frac{b^2}{2c} - \bar{u}_w(c)\right) - \bar{u}_M(k).$$

The optimal  $b$  and  $c$  are then determined from the following optimization:

$$\max_{b,c} \hat{\Pi}(b, c, k; \sigma, r) \text{ such that } 0 \leq b < q \text{ and } c > 0.$$

Obviously, this problem is equivalent to:

$$\max_{b,c} \left(\frac{q - b}{2c}\right)^{1/\tau} \left(\frac{b^2}{2c} - \bar{u}_w(c)\right)$$

such that  $0 \leq b < q$  and  $c > 0$ , which does not depend on  $\sigma$  or  $r$ . ■

**Proof of Proposition 8:** Expanding the objective function in (43), we obtain:

$$\begin{aligned} \Pi(n, c, k; \sigma, r) &= \frac{q^2}{2c} n - 2q\sigma\sqrt{rk}Bn^{\tau+1} + 2c\sigma^2rkB^2n^{2\tau+1} \\ &\quad - n \cdot \bar{u}_w(c) - \bar{u}_M(k). \end{aligned} \tag{44}$$

The maximization problem now becomes:

$$\max_{n,c,k} \Pi(n, c, k|\sigma, r) \text{ such that } n, c, k > 0 \text{ and } q - 2c\sigma\sqrt{rk}h(n) \geq 0.$$

First, observe that the comparative statics with respect to  $\sigma$  and  $r$  share the same sign. The optimal choice functions are  $Z^*(\sigma, r)$  where  $Z = n, c$  or  $k$ . Note that parameters  $\sigma$  and  $r$  enter the problem only through the expression  $\sigma\sqrt{r}$ . This implies that  $Z^*(\sigma, r) = Z^*(\sigma\sqrt{r}, 1)$ . Therefore,  $\text{sign} \frac{\partial Z^*(\sigma, r)}{\partial \sigma} = \text{sign} \frac{\partial Z^*(\sigma, r)}{\partial r} = \text{sign} \frac{\partial Z^*(\sigma\sqrt{r}, 1)}{\partial \sigma}$ .

**Assumption 1:** *The optimum is reached in the interior.*

First Order Conditions:

$$\Pi_n = \frac{q^2}{2c} - 2q(1 + \tau)\sigma\sqrt{rk}Bn^\tau + 2(1 + 2\tau)c\sigma^2rkB^2n^{2\tau} - \bar{u}_w(c) = 0; \tag{45a}$$

$$\Pi_c = n \left( -\bar{u}'_w(c) - \frac{q^2}{2c^2} + 2\sigma^2rkB^2n^{2\tau} \right) = 0; \tag{45b}$$

$$\Pi_k = -\bar{u}'_M(k) - q\sigma\sqrt{\frac{r}{k}}Bn^{\tau+1} + 2c\sigma^2rB^2n^{2\tau+1} = 0. \tag{45c}$$

The optimal values  $n^*(\sigma, r)$ ,  $c^*(\sigma, r)$ , and  $k^*(\sigma, r)$  solve system (45). The implicit function proposition ensures the existence of differentiable implicit functions  $n^*(\sigma, r)$ ,  $c^*(\sigma, r)$ , and  $k^*(\sigma, r)$ , provided the second derivative matrix  $D^2\Pi$  is nondegenerate.

**Assumption 2:** Matrix  $D^2\Pi$  is negative definite at the optimal solution.

A matrix is negative definite if and only if the leading principal minors switch signs:

$$\Pi_{nn} < 0 \tag{46}$$

$$\Pi_{nn}\Pi_{cc} - \Pi_{nc}^2 > 0 \tag{47}$$

$$\det D^2\Pi < 0 \tag{48}$$

Equation (48) shows that we can apply the implicit function theorem to system (45). Differentiating (45) w.r.t.  $\sigma$ , we get:

$$\begin{pmatrix} \partial n^*/\partial\sigma \\ \partial c^*/\partial\sigma \\ \partial k^*/\partial\sigma \end{pmatrix} = -(D^2\Pi)^{-1} \begin{pmatrix} \Pi_{n\sigma} \\ \Pi_{c\sigma} \\ \Pi_{k\sigma} \end{pmatrix}. \tag{49}$$

**Second derivatives.** Differentiating (45a), we get:

$$\Pi_{nn} = 4\tau(1 + 2\tau)c\sigma^2rkB^2n^{2\tau-1} - 2q\tau(\tau + 1)\sigma\sqrt{rk}Bn^{\tau-1}. \tag{50}$$

Differentiating (45b) with respect to  $n$  and using the FOC for  $c$ , we get:

$$\Pi_{nc} = -\bar{u}'_w(c) - \frac{q^2}{2c^2} + 2\sigma^2rkB^2n^{2\tau} + 4\tau\sigma^2rkB^2n^{2\tau} = 4\tau\sigma^2rkB^2n^{2\tau}, \tag{51}$$

where  $-\bar{u}'_w(c) - \frac{q^2}{2c^2} + 2\sigma^2rkB^2n^{2\tau} = 0$  by (45b). Differentiating (45c) with respect to  $n$ , we get:

$$\Pi_{nk} = -q(\tau + 1)\sigma\sqrt{\frac{r}{k}}Bn^\tau + 2(2\tau + 1)\sigma^2crB^2n^{2\tau}.$$

In a similar manner:

$$\begin{aligned}
 \Pi_{cc} &= n \left( -\bar{u}_W''(c) + \frac{q^2}{c^3} \right), \\
 \Pi_{ck} &= 2\sigma^2 r B^2 n^{2\tau+1}, \\
 \Pi_{kk} &= -\bar{u}_M''(k) + \frac{1}{2} q \sigma r^{1/2} k^{-3/2} B n^{\tau+1}, \\
 \Pi_{n\sigma} &= -2q(1 + \tau) \sqrt{rk} B n^\tau + 4(1 + 2\tau) c \sigma r k B^2 n^{2\tau}, \\
 \Pi_{c\sigma} &= 4\sigma r k B^2 n^{2\tau+1}, \\
 \Pi_{k\sigma} &= -q \sqrt{\frac{r}{k}} B n^{\tau+1} + 4c \sigma r B^2 n^{2\tau+1}, \\
 \Pi_{nr} &= -q(1 + \tau) \sigma \sqrt{\frac{k}{r}} B n^\tau + 2(1 + 2\tau) c \sigma^2 k B^2 n^{2\tau}, \\
 \Pi_{cr} &= 2\sigma^2 k B^2 n^{2\tau+1}, \\
 \Pi_{kr} &= -q \sigma \frac{1}{2\sqrt{rk}} B n^{\tau+1} + 2c \sigma^2 B^2 n^{2\tau+1}.
 \end{aligned}
 \tag{52}$$

Finally, by direct calculations, using the formulas above, we can see that:

$$\Pi_{n\sigma} = \frac{n}{\tau\sigma} \Pi_{nn}, \quad \Pi_{nk} = \frac{n}{2\tau k} \Pi_{nn}.
 \tag{53}$$

Similarly:

$$\Pi_{c\sigma} = \frac{n}{\tau\sigma} \Pi_{nc}, \quad \Pi_{ck} = \frac{n}{2\tau k} \Pi_{nc}.
 \tag{54}$$

**Behavior of  $c^*(\sigma, r)$ .** Applying Cramer’s rule to the linear system (49), we get:

$$\begin{aligned}
 -\det D^2\Pi \frac{\partial c^*}{\partial \sigma} &= -\Pi_{n\sigma} (\Pi_{nc}\Pi_{kk} - \Pi_{ck}\Pi_{nk}) \\
 &+ \Pi_{c\sigma} (\Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2) - \Pi_{k\sigma} (\Pi_{nn}\Pi_{ck} - \Pi_{nk}\Pi_{nc}).
 \end{aligned}$$

After eliminating  $\Pi_{nk}$ ,  $\Pi_{n\sigma}$ ,  $\Pi_{c\sigma}$  and  $\Pi_{ck}$  by substituting (53) and (54), we obtain:

$$\begin{aligned}
 -\det D^2\Pi \frac{\partial c^*}{\partial \sigma} &= -\frac{n}{\tau\sigma} \Pi_{nn} \left( \Pi_{nc}\Pi_{kk} - \frac{n^2}{4\tau^2 k^2} \Pi_{nc}\Pi_{nn} \right) \\
 &+ \frac{n}{\tau\sigma} \Pi_{nc} \left( \Pi_{nn}\Pi_{kk} - \frac{n^2}{4\tau^2 k^2} (\Pi_{nn})^2 \right) \\
 &- \Pi_{k\sigma} \left( \frac{n}{2\tau k} \Pi_{nn}\Pi_{nc} - \frac{n}{2\tau k} \Pi_{nn}\Pi_{nc} \right) = 0.
 \end{aligned}$$

Now  $\det D^2\Pi \neq 0$ , from (48), so  $\frac{\partial c^*}{\partial \sigma} = 0$ .

**Behavior of  $k^*(\sigma, r)$ .** Again using Cramer’s rule for the linear system (49):

$$\begin{aligned}
 -\det D^2\Pi \frac{\partial k^*}{\partial \sigma} &= \Pi_{n\sigma}(\Pi_{nc}\Pi_{ck} - \Pi_{cc}\Pi_{nk}) \\
 &- \Pi_{c\sigma}(\Pi_{nn}\Pi_{ck} - \Pi_{nc}\Pi_{nk}) + \Pi_{k\sigma}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2).
 \end{aligned}$$

We eliminate  $\Pi_{nk}$ ,  $\Pi_{n\sigma}$ ,  $\Pi_{c\sigma}$ , and  $\Pi_{ck}$  by substituting for (53) and (54):

$$\begin{aligned}
 -\det D^2\Pi \frac{\partial k^*}{\partial \sigma} &= \frac{n}{\tau\sigma} \Pi_{nn} \left( \frac{n}{2\tau k} (\Pi_{nc})^2 - \frac{n}{2\tau k} \Pi_{cc}\Pi_{nn} \right) \\
 &- \frac{n}{\tau\sigma} \Pi_{nc} \left( \frac{n}{2\tau k} \Pi_{nn}\Pi_{nc} - \frac{n}{2\tau k} \Pi_{nc}\Pi_{nn} \right) + \Pi_{k\sigma}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) \\
 &= (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) \left( \Pi_{k\sigma} - \frac{n^2}{2\tau^2 k\sigma} \Pi_{nn} \right).
 \end{aligned} \tag{55}$$

The last factor in the expression above is, after simplifying:

$$\Pi_{k\sigma} - \frac{n^2}{2\tau^2 k\sigma} \Pi_{nn} = \frac{Bn^{\tau+1}}{\tau} \sqrt{\frac{r}{k}} b. \tag{56}$$

Substituting back into (55), we get:

$$\frac{\partial k^*}{\partial \sigma} = -\frac{1}{\det D^2\Pi} (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) \frac{Bn^{\tau+1}}{\tau} \sqrt{\frac{r}{k}} b > 0,$$

because  $\det D^2\Pi < 0$  and  $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$ , by (47) and (48).

**Behavior of  $n^*(\sigma, r)$ .** We use Cramer’s rule for the linear system (49) to obtain:

$$\begin{aligned}
 -\det D^2\Pi \frac{\partial n^*}{\partial \sigma} &= \Pi_{n\sigma} (\Pi_{cc}\Pi_{kk} - (\Pi_{ck})^2) \\
 &- \Pi_{c\sigma}(\Pi_{nc}\Pi_{kk} - \Pi_{nk}\Pi_{ck}) + \Pi_{k\sigma}(\Pi_{nc}\Pi_{ck} - \Pi_{cc}\Pi_{nk}).
 \end{aligned}$$

We eliminate  $\Pi_{nk}$ ,  $\Pi_{n\sigma}$ ,  $\Pi_{c\sigma}$ , and  $\Pi_{ck}$  by substituting (53) and (54). This yields:

$$-\det D^2\Pi \frac{\partial n^*}{\partial \sigma} = \left( \frac{n}{\tau\sigma} \Pi_{kk} - \frac{n}{2\tau k} \Pi_{k\sigma} \right) (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2). \tag{57}$$

We first show that the first factor in the last expression above is negative. Since the ordering of the rows and columns of the Hessian  $D^2\Pi$  is arbitrary, we can swap the second and third columns and rows of the matrix. As before, the matrix is negative definite if and only if the leading principal minors alternate in sign. This implies that  $\Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2 > 0$ . Using (53):

$$0 < \Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2 = \Pi_{nn} \left( \Pi_{kk} - \frac{n^2}{4\tau^2k^2} \Pi_{nn} \right). \tag{58}$$

As  $\Pi_{nn} < 0$  from (46), inequality (58) implies:

$$\Pi_{kk} < \frac{n^2}{4\tau^2k^2} \Pi_{nn}. \tag{59}$$

Using this inequality to eliminate  $\Pi_{kk}$  from (57), and using (56) to simplify, we get

$$\begin{aligned} \frac{n}{\tau\sigma} \Pi_{kk} - \frac{n}{2\tau k} \Pi_{k\sigma} &< \frac{n}{\tau\sigma} \times \frac{n^2}{4\tau^2k^2} \Pi_{nn} - \frac{n}{2\tau k} \Pi_{k\sigma} \\ &= -\frac{n}{2\tau k} \frac{Bn^{\tau+1}}{\tau} \sqrt{\frac{r}{k}} b < 0. \end{aligned} \tag{60}$$

Thus, from (57) we get:

$$\frac{\partial n^*}{\partial \sigma} = -\frac{1}{\det D^2\Pi} \left( \frac{n}{\tau\sigma} \Pi_{kk} - \frac{n}{2\tau k} \Pi_{k\sigma} \right) (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) < 0,$$

where  $\det D^2\Pi < 0$  by (48),  $\frac{n}{\tau\sigma} \Pi_{kk} - \frac{n}{2\tau k} \Pi_{k\sigma} < 0$  by (60), and  $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$  by (47). ■

**Proof of Proposition 9:** When  $G(n) = An^\gamma$ ,  $h(n) = \sqrt{G(n)/n} = Bn^\tau$ , where  $\tau = \frac{\gamma - 1}{2} > 0$ . As such, comparative statics with respect to  $\gamma$  mirror those with respect to  $\tau$ . We compute the latter, and start by differentiating (45) with respect to  $\tau$ :

$$\begin{aligned} \Pi_{n\tau} &= (4(1 + 2\tau)c\sigma^2rkB^2n^{2\tau} - 2q(1 + \tau)\sigma\sqrt{rk}Bn^\tau) \left( \frac{1}{1 + 2\tau} + \ln n \right) \\ &\quad - \frac{2\tau}{1 + 2\tau} q\sigma\sqrt{rk}Bn^\tau. \end{aligned}$$

Since  $4(1 + 2\tau)c\sigma^2rkB^2n^{2\tau} - 2q(1 + \tau)\sigma\sqrt{rk}Bn^\tau = \frac{n}{\tau} \Pi_{nn}$  by (50):

$$\Pi_{n\tau} = \Pi_{nn} \frac{n}{\tau} \left( \frac{1}{1 + 2\tau} + \ln n \right) - \frac{2\tau}{1 + 2\tau} q\sigma\sqrt{rk}Bn^\tau. \tag{61}$$

The two remaining equations are straightforward:

$$\Pi_{c\tau} = 4\sigma^2rkB^2n^{2\tau+1} \ln n = \Pi_{nc} \frac{n}{\tau} \ln n \text{ (see Equation (51))} \tag{62}$$

$$\Pi_{k\tau} = \left( -q\sigma \sqrt{\frac{r}{k}} Bn^{\tau+1} + 4c\sigma^2rB^2n^{2\tau+1} \right) \ln n = \Pi_{k\sigma} \ln n \text{ (see Equation (52))} \tag{63}$$

**Behavior of  $c^*(\tau)$ .** By applying Cramer’s rule to a system analogous to (49) (with  $\tau$  in place of  $\sigma$ ), we get:

$$-\det D^2\Pi \frac{\partial c^*}{\partial \tau} = -\Pi_{n\tau}(\Pi_{nc}\Pi_{kk} - \Pi_{nk}\Pi_{ck}) + \Pi_{c\tau}(\Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2) - \Pi_{k\tau}(\Pi_{nn}\Pi_{ck} - \Pi_{nc}\Pi_{nk}).$$

We use (53) and (54) to eliminate  $\Pi_{ck}$  and  $\Pi_{nk}$  and obtain:

$$-\det D^2\Pi \frac{\partial c^*}{\partial \tau} = (\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau}) \left( \Pi_{kk} - \frac{n^2}{4\tau^2k^2} \Pi_{nn} \right). \tag{64}$$

We transform the first factor in Equation (64), using (61) and (62) to eliminate  $\Pi_{n\tau}$  and  $\Pi_{c\tau}$ :

$$\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau} = \Pi_{nc} \left( -\Pi_{nn} \frac{n}{\tau} \frac{1}{1+2\tau} + \frac{2\tau}{1+2\tau} q\sigma\sqrt{rk}Bn^\tau \right) > 0, \tag{65}$$

since  $\Pi_{nc} > 0$  by (51), and  $-\Pi_{nn} > 0$  by (46). Finally, we get:

$$\frac{\partial c^*}{\partial \tau} = -\frac{1}{\det D^2\Pi} (\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau}) \left( \Pi_{kk} - \frac{n^2}{4\tau^2k^2} \Pi_{nn} \right) < 0,$$

since  $\det D^2\Pi < 0$  by (48),  $\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau} > 0$  by (65), and  $\Pi_{kk} - \frac{n^2}{4\tau^2k^2} \Pi_{nn} < 0$  by (59).

**Behavior of  $n^*(\tau)$ .** By applying Cramer’s rule to a system analogous to (49) (with  $\tau$  in place of  $\sigma$ ), we get:

$$-\det D^2\Pi \frac{\partial n^*}{\partial \tau} = \Pi_{n\tau}(\Pi_{cc}\Pi_{kk} - (\Pi_{ck})^2) - \Pi_{c\tau}(\Pi_{nc}\Pi_{kk} - \Pi_{ck}\Pi_{nk}) + \Pi_{k\tau}(\Pi_{nc}\Pi_{ck} - \Pi_{cc}\Pi_{nk}).$$

We use (53) and (54) to eliminate  $\Pi_{ck}$  and  $\Pi_{nk}$  from this equation to obtain:

$$-\det D^2\Pi \frac{\partial n^*}{\partial \tau} = (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) \frac{n}{2\tau k} \left( \frac{n}{2\tau k} \Pi_{n\tau} - \Pi_{k\tau} \right) + (\Pi_{n\tau}\Pi_{cc} - \Pi_{c\tau}\Pi_{nc}) \left( \Pi_{kk} - \frac{n^2}{4\tau^2k^2} \Pi_{nn} \right), \tag{66}$$

where  $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$  by (47), and  $\Pi_{kk} - \frac{n^2}{4\tau^2k^2} \Pi_{nn} < 0$  by (59). Define  $E_1 \equiv \frac{n}{2\tau k}$

$\Pi_{n\tau} - \Pi_{k\tau}$  and  $E_2 \equiv \Pi_{n\tau}\Pi_{cc} - \Pi_{c\tau}\Pi_{nc}$ . Before we can calculate their signs, we must make an additional assumption in order to fix the sign of  $n^*$ .

**Assumption 3:** *At the optimum, the value of  $n$  is greater than one, i.e.,  $n^* > 1$ .*

Using (61) and (63) with this assumption, we get:

$$E_1 = -\sigma \ln n \left( \Pi_{k\sigma} - \frac{n^2}{2\tau^2 k \sigma} \Pi_{nn} \right) + \frac{n^2}{2\tau^2 k} \Pi_{nn} \frac{1}{1 + 2\tau} - \frac{q\sigma B n^{\tau+1}}{(1 + 2\tau)} \sqrt{\frac{r}{k}} < 0, \tag{67}$$

where  $\sigma \ln n > 0$  by Assumption 3,  $\Pi_{k\sigma} - \frac{n^2}{2\tau^2 k \sigma} \Pi_{nn} > 0$  by (56), and  $\Pi_{nn} < 0$  by (46).

Using (61) and (62) similarly, we get:

$$E_2 = (\Pi_{cc}\Pi_{nn} - (\Pi_{nc})^2) \frac{n}{\tau} \ln n + \Pi_{cc} \left( \Pi_{nn} \frac{n}{\tau} \frac{1}{1 + 2\tau} - \frac{2\tau}{1 + 2\tau} q\sigma \sqrt{rk} B n^\tau \right) > 0, \tag{68}$$

because  $\Pi_{cc}\Pi_{nn} - (\Pi_{nc})^2 > 0$  by (47),  $\frac{n}{\tau} \ln n > 0$  by Assumption 3,  $\Pi_{cc} < 0$ , and  $\Pi_{nn} < 0$

by Assumption 2. Combining (66), (67), and (68), we see that  $\frac{\partial n^*}{\partial \tau} < 0$ .

**Behavior of  $k^*(\tau)$ .** By applying Cramer’s rule to a system analogous to (49) (with  $\tau$  in place of  $\sigma$ ), we get:

$$-\det D^2\Pi \frac{\partial k^*}{\partial \tau} = \Pi_{n\tau}(\Pi_{nc}\Pi_{ck} - \Pi_{nk}\Pi_{cc}) - \Pi_{c\tau}(\Pi_{nn}\Pi_{ck} - \Pi_{nc}\Pi_{nk}) + \Pi_{k\tau}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2).$$

To get rid of  $\Pi_{ck}$  and  $\Pi_{nk}$ , we use (53) and (54). This yields:

$$-\det D^2\Pi \frac{\partial k^*}{\partial \tau} = (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) \left( \Pi_{k\tau} - \frac{n}{2\tau k} \Pi_{n\tau} \right) > 0,$$

since  $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$  by (47), and  $\Pi_{k\tau} - \frac{n}{2\tau k} \Pi_{n\tau} = -E_1 > 0$  by (69). Therefore,

$$\frac{\partial k^*}{\partial \tau} > 0. \quad \blacksquare$$

**REFERENCES**

Acemoglu, D., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti. 2006. Technology information and the decentralization of the firm. Working paper, London School of Economics.  
 Adams, C. 2006. Optimal team incentives with CES production. *Economics Letters* 92 (July): 143–148.

- Autrey, R. 2005. Team synergy, team composition and performance measures. Working paper, Harvard Business School.
- Baiman, S., and J. Demski. 1980. Economically optimal performance evaluation and control systems. *Journal of Accounting Research* 18 (Supplement): 184–220.
- Baldenius, T., N. Melumad, and A. Ziv. 2002. Monitoring in multiagent organizations. *Contemporary Accounting Research* 19 (Winter): 483–511.
- Bhattacharyya, S., and F. Lafontaine. 1995. Double-sided moral hazard and the nature of share contracts. *RAND Journal of Economics* 26 (Winter): 761–781.
- Bushman, R., R. Indjejikian, and A. Smith. 1995. Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. *Journal of Accounting Research* 33 (Supplement): 101–128.
- Calvo, G., and S. Wellisz. 1978. Supervision, loss of control, and the optimum size of the firm. *Journal of Political Economy* 86 (October): 943–952.
- Christensen, P., and F. Sabac. 2002. Dynamic agency and accounting informativeness. Working paper, University of Alberta.
- Fischer, P., and J. Hughes. 1997. Mutual monitoring and best agency contracts. *Journal of Institutional and Theoretical Economics* 153 (June): 334–355.
- Guadalupe, M., and J. Wulf. 2007. The flattening firm and product market competition: The effect of trade costs and liberalization. Working paper, Columbia University.
- Harris, M., and A. Raviv. 2002. Organizational design. *Management Science* 48 (July): 852–865.
- Holmström, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 13 (Autumn): 324–340.
- Huddart, S., and P. Liang. 2005. Profit sharing and monitoring in partnerships. *Journal of Accounting and Economics* 40 (December): 153–187.
- Hughes, J., L. Zhang, and J. Xie. 2005. Production externalities, congruity of aggregate signals, and optimal task assignments. *Contemporary Accounting Research* 61 (Summer): 398–408.
- Ma, C. 1988. Unique implementation of incentive contracts with many agents. *Review of Economic Studies* 55 (October): 555–571.
- Marschak, J., and R. Radner. 1972. *Economic Theory of Teams*. Cowles Foundation. New Haven, CT: Yale University Press.
- McLaughlin, K. 1994. Individual compensation and firm performance: The economics of team incentives. Working paper, University of Chicago.
- Prendergast, C. 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy* 110 (October): 1071–1102.
- Qian, Y. 1994. Incentives and loss of control in an optimal hierarchy. *The Review of Economic Studies* 61 (July): 527–544.
- Rajan, R., and J. Wulf. 2006. The flattening firm: Evidence on the changing nature of firm hierarchies from panel data. *Review of Economics and Statistics* 88 (November): 759–773.
- Ricart i Costa, J. 1988. Managerial task assignment and promotions. *Econometrica* 56 (March): 449–466.
- Roberts, J. 2004. *The Modern Firm: Organizational Design for Performance and Growth*. Oxford, U.K.: Oxford University Press.
- Sappington, D. 1983. Limited liability contracts between principal and agent. *Journal of Economic Theory* 29 (February): 1–21.
- Simons, R. 2005. *Levers of Organization Design*. Boston, MA: Harvard Business School Press.
- Varian, H. 1990. Monitoring agents with other agents. *Journal of Institutional and Theoretical Economics* 146 (March): 153–174.
- Williamson, O. 1967. Hierarchical control and optimum firm size. *Journal of Political Economy* 75 (April): 123–138.
- Ziv, A. 1993. Performance measures and optimal organization. *Journal of Law, Economics, and Organization* 9 (April): 30–50.
- . 2000. Information technology and optimal firm structure. *Journal of Accounting Research* 38 (Autumn): 297–328.