MANAGEMENT SCIENCE
Vol. 61, No. 2, February 2015, pp. 335–353
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2013.1858

Sorting Effects of Performance Pay

Maris Goldmanis
Department of Economics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom, maris.goldmanis@rhul.ac.uk

Korok Ray
School of Business, George Washington University, Washington, District of Columbia 20052, korok@gwu.edu

Compensation not only provides incentives to an existing manager but also affects the type of manager attracted to the firm. This paper examines the dual incentive and sorting effects of performance pay in a simple contracting model of endogenous participation. Unless the manager is highly risk averse, sorting dampens optimal pay-performance sensitivity (PPS) because PPS beyond a nominal amount transfers unnecessary (information) rent to the manager. This helps explain why empirical estimates of PPS are much lower than predictions from models of moral hazard alone. The model also predicts that sorting under asymmetric information causes the firm to turn away more candidates than would be efficient; PPS increases in the cost of hiring the manager and in the manager's outside option, but decreases in output risk, information risk, and managerial risk aversion; and the firm becomes more selective in hiring as either the manager's outside option, the cost of hiring, risk aversion, output risk, or information risk increases.

Keywords: accounting; compensation; performance measurement; management; agency theory; contract theory
History: Received June 28, 2011; accepted September 24, 2013, by Mary Barth, accounting. Published online in *Articles in Advance* March 31, 2014.

1. Introduction

Managerial compensation serves many functions. It provides incentives, attracts talent, ensures retention, offers feedback, and communicates the goals and objectives of the firm. Yet, the vast majority of theoretical and empirical work on executive pay considers only incentive effects. Executive contracts not only provide incentives to the existing manager, but they also attract new types of managers to the firm, from either internal or external labor markets. Thus, performance pay has a sorting effect, in that it sorts the potential pool of managers tomorrow in addition to providing incentives to the incumbents today. The objective is to understand the sorting effects of performance pay, namely, how the firm will solve the dual problem of providing incentives ex post, and sorting new types of managers ex ante.

Unless the manager is highly risk averse, we find that sorting dampens optimal pay-performance sensitivity (PPS). Some PPS is necessary to attract the high types (of managers) and repel the low types (of managers). High types prefer performance pay because their high ability drives productivity and, therefore, their income. One the other hand, low types prefer to collect their outside option rather than receive low compensation from the firm. However, PPS allows high-type managers to collect information rents. The highest type of manager captures the most information rents. These information rents are costly for the firm, and thus the firm seeks to limit this rent transfer by lowering PPS. Therefore, sorting exerts a downward pressure on PPS. Since the firm must also provide incentives to a manager once he is hired, the incentive effect exerts an upward pressure on PPS. The optimal PPS balances these twin competing effects, bringing the theory on PPS closer to its empirics.[1]

We adopt a simple contracting framework that permits a solution to the dual sorting and incentive problems. A risk-neutral manager has private information on his ability, and the firm's contracts are incomplete since they cannot easily extract this private information through a complex menu of contracts.[2] The firm proposes a contract, which consists of a salary and bonus, representing the fixed and variable components of compensation. Based on this contract, the candidate manager decides whether to join the firm. If so, he exerts productive effort. After nature resolves production uncertainty, the output is realized and the

[1] Existing estimates of PPS (0.325%, according to Jensen and Murphy 1990) are much lower than predictions from the canonical model of moral hazard alone, even when factoring in risk aversion.

[2] This amounts to a restriction of communication between the firm and the agent through the contracts. The firm cannot tailor its contracts to the full complexity of the manager's private information. Baker and Jorgensen (2003), Melumad et al. (1997), and Ray (2007b) make a similar assumption. Bushman et al. (2000) also work in a world of private predecision information.

firm pays the manager based upon the negotiated contract.

The crux of the analysis rests on the firm's joint choice of salary and bonus, where the bonus measures the PPS of the manager's compensation. We start with a benchmark model of sorting alone (without incentives) and find that a low PPS minimizes the manager's information rent. The downward pressure that sorting exerts on the PPS becomes even stronger in the more general model, which combines sorting and incentives. There, compensation induces participation ex ante as well as determines effort (and therefore profits) ex post. While a higher level of PPS induces higher effort ex post (which the firm desires), this also transfers higher information rents to the manager (which the firm would like to avoid). In equilibrium, the positive ex post effect must balance the negative ex ante effect on the margin.

Our paper is closest in spirit to Dutta (2008) and Baker and Jorgensen (2003). Both operate in a LEN (linear contract, exponential utility, normal errors) framework and consider an agent whose ability affects output. Our model shares more assumptions with Baker and Jorgensen (2003), but it follows the approach of Dutta (2008). Like Baker and Jorgensen (2003), we work in a world of predecision information, assume ability and effort are complements, and disallow communication between the principal and agent. Like Dutta (2008), we aim to characterize the optimal contract (although implicitly) and make comparisons with the benchmark moral hazard model without information. Our primary difference is that we make participation endogenous. Both papers assume the principal will hire the worst type of manager; however, we demonstrate how some managers are not necessarily profitable for the firm. As such, the contract in our model must solve the dual problem of participation and incentives.[3]

Endogenous participation provides a new analytical lens that extends prior work. Dutta (2008) shows that when a risk-averse manager's ability is firm-specific (his outside options are invariant to his ability), adverse selection considerations mute PPS. In our model, sorting is guaranteed to dampen PPS only when risk aversion and output uncertainty are small; in such cases, the need to reduce information rents under sorting uniformly forces PPS down. If, however, risk aversion and output uncertainty become large, the cost of risk begins to dominate the information rents effect. For lower-ability managers (for whom the relative cost of risk is highest), this causes PPS in the canonical model (perfect information) to fall lower than the PPS with sorting (i.e., sorting inflates the optimal PPS). Dutta (2008) also finds that if a manager has firm-specific human capital, optimal PPS falls in the variance in managerial ability. We observe the same phenomenon under sorting.

Our model delivers a number of comparative statics that fortify intuition and can potentially be empirically tested. First, when the manager's cost of employment or outside options rise, the surplus generated by each manager shrinks, forcing the firm to hire more selectively; by raising hiring standards, the firm reduces the total information rents accruing to the manager, and therefore can afford to raise PPS. Second, when the variation between managers increases, the information asymmetry problem worsens, and the information rents accruing to the manager rise, forcing the firm to decrease PPS to limit these information rents. Third, when managerial risk aversion or output variability is high, this higher cost of risk decreases the surplus produced by a given manager, causing the firm to hire better managers. But although hiring fewer types of managers will reduce total information rents that could allow the firm to increase PPS, the firm will not do so (since that would only load more risk onto the risk-averse manager); hence, PPS will decline and the classical trade-off between risk and incentives still stands. All our predictions on PPS should be straightforward to test, given the large empirical executive compensation literature that calculates PPS from proxy statements, annual reports, and other financial statements of the firm. Measuring hiring standards is harder but not impossible, especially if one observes data on applications and decisions.

Models of adverse selection and moral hazard each enjoy voluminous theoretical literatures (for surveys, see Baiman 1991, Hart and Holmstrom 1987). But there have been only limited attempts to combine both in a single model. The fusion has proven notoriously difficult and researchers have made simplifying assumptions to make the analysis tractable.[4] In general, the literature remains largely separate, even

[3] Communication, effort, and ability are perfect substitutes that allow Dutta (2008) to characterize the optimal contract, whereas Baker and Jorgensen (2003) derive comparative statics without solving for the optimal contract in closed form. Dutta (2008) assumes full participation, ability-contingent outside options, that ability and effort are substitutes ($e + \theta$), and considers risk aversion throughout. We assume no communication, fixed outside options, that effort and ability are complements ($e\theta$), and examine risk neutrality as well as risk aversion.

[4] Jullien et al. (1999) derive some preliminary results on a joint moral hazard and adverse selection model, although it is difficult to draw general conclusions from their analysis. Sung (2005) makes progress in a continuous time framework, Darrough and Stoughton (1986) examine the joint problem in a particular financial context, and Bernardo et al. (2001) operate in the world of capital budgeting. Hagerty and Siegal (1988) show that contracts under moral hazard and adverse selection are observationally equivalent.

though real-life contracts must solve both problems simultaneously.[5] Our paper contributes to the emergent literature on the sorting/matching approach to understanding compensation. This literature was pioneered by Rosen (1981), who proposed a neoclassical model of compensation set through a labor market, where wages match heterogeneous employees to heterogeneous firms. This literature has grown recently through Himmelberg and Hubbard (2000), Arya and Mittendorf (2005), Ray (2007a), Liang et al. (2008), Gabaix and Landier (2008), and Edmans et al. (2009). Our model blends this neoclassical approach to compensation with a principal agent model, and, as such, provides a bridge between the sorting/matching approach and the agency approach of compensation.

2. The Basic Model

To fix ideas, consider the basic model with sorting, but no incentives. A firm (the principal) considers employing a single manager (the agent). The firm has only one chance to hire a manager; if it does not hire the current candidate, it will not receive further applicants; hence, the problem is one of sorting, not one of search. This is justified, for example, when the firm can expand freely, creating a division for each manager, or when the firm needs to hire a manager for a specific project on short notice. The manager is risk neutral. The manager has a type θ, which he knows but the firm does not. Hereafter, "type" refers to the type θ of the manager; although there is a single manager, there is a continuum of types. The firm's uncertainty on θ is represented by the density f with cumulative distribution function F, over support $\Theta = [0, 1)$, with mean μ_θ and variance σ_θ^2. The firm's output with (a manager of type) θ is

$$x = \gamma\theta + \epsilon, \tag{1}$$

where ϵ is distributed symmetrically with mean 0 and variance σ^2. The parameter $\gamma > 0$ represents the complementarity between the firm and the manager. High γ firms produce more output with high θ types than with low θ types. The manager enjoys an outside option $\bar{u} > 0$, which represents his outside opportunities.[6] The firm bears a fixed cost $k - m\theta$ to employ a manager of type θ, where $k > m > 0$.[7] This shows why a better manager is less costly to employ. A better manager is less likely to make mistakes or bad decisions. The parameter k can also track, for example, the level of general versus firm-specific human capital: firms that require more specialized skills (finance or technology) may bear a larger cost of installing and training the manager, compared to firms that require more general skills (retail, commodities).[8] The parameter m captures the return to a better manager; when m is large, the cost saving of a better manager is large.

2.1. First Best

A social planner maximizes total surplus, which is the value that the manager produces, net of all costs and over and above the total outside options of the parties. The expected surplus for each θ is $E[TS \mid \theta] = \gamma\theta - (k - m\theta) - \bar{u}$. Ex post efficiency will require that total surplus be positive for each θ. This occurs when

$$\theta > \frac{k + \bar{u}}{m + \gamma} \equiv \theta^{FB}. \tag{2}$$

Thus, ex post efficiency establishes a marginal type θ^{FB}, above which a manager of type θ generates positive surplus. Observe that this first-best cutoff θ^{FB} rises in k and $\bar{u}$ and falls in m and γ; therefore, it is efficient for the firm to require a better manager when employment is expensive (to compensate for the high fixed cost of hiring), when managers have higher outside options (to compensate the managers for their higher opportunity costs), and when the quality of the match between the firm and the manager is low (to compensate for the lower productivity of a poor match).

We will assume throughout the paper that $\theta^{FB} < 1$ (i.e., $k + \bar{u} < m + \gamma$). This simply ensures that it is possible for a manager to produce positive surplus. If this condition did not hold, the problem would be uninteresting because it would always be optimal for the firm to exit the market.

The formula for the first-best cutoff provides insight into when sorting matters. The cost of hiring the manager has both a fixed component (k) and a variable component ($m\theta$), where the variable component varies not per unit produced but rather for an incremental change in the manager's ability. Firms with high fixed components (high k) are those where it is costly to install a manager. For example, these can be firms in technical industries that require a high level of industry-specific or firm-specific human capital (biotechnology, financial services). For such firms, it is important to obtain a high-quality manager to compensate for these high fixed costs; as such, the efficient cutoff θ^{FB} will be high. Firms that require more

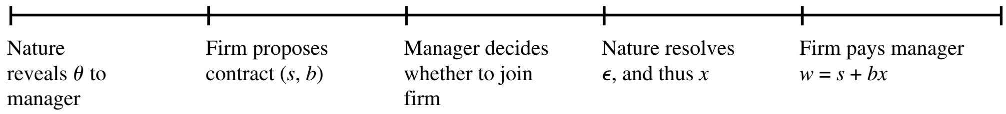
[5] Armstrong et al. (2010) solve the joint problem numerically, simulating the optimal CEO contract under realistic assumptions on the agent's risk aversion and actual executive contracts.

[6] The manager's outside option $\bar{u}$ is fixed and does not vary with θ. However, the crux of the results of the basic model still holds under outside options that are increasing in the manager's type, provided the outside options don't increase too steeply.

[7] The assumption that $k > m > 0$ ensures that the hiring cost is always positive.

[8] See Corollary 2 in §3 for a discussion of empirical proxies for k and implications for cross-sectional variation.

Figure 1 **Timeline of the Basic Model**



general human capital (consumer products, retailing) may have lower fixed costs and, therefore, lower needs for able managers.

The coefficient on the variable component of the cost function, m, tracks how much an incremental increase in quality decreases the cost to the firm. Firms with high variable components are those that markedly benefit from managerial ability. In such companies, the need for sorting is lower because it is built into the cost function. Such companies are very sensitive to ability, since they markedly decrease the firm's cost function. High-ability managers are productive at such firms, but so are low-ability managers because of the sensitivity of the cost function (its steep slope). In contrast, firms with a low variable component (low m) need sorting the most, as only highly able managers will be able to produce value for the firm. Under a low m, lower-ability managers are worth less to the firm; hence, they must be screened out through a high θ^{FB} hurdle.[9]

The firm cannot observe the type of the manager and thus must induce his employment through a compensation contract. A contract consists of a salary s and a bonus b. For tractability, we restrict attention to linear contracts of the form

$$w = s + bx. \tag{3}$$

This reflects the main feature of most compensation schemes, which have a fixed salary and a bonus that depends on some performance measure. Contracts are incomplete in that the firm cannot offer the manager a menu of contracts that depend on an announcement of the manager's type.

When would such incomplete contracts occur? Eggleston et al. (2000) provide a number of reasons why contracts are simpler and less complete than in orthodox economic theory: negotiation costs, differential monitoring dynamics, social conventions, reliance on trust and reputation, bounded rationality, and enforcement costs. For example, Joskow (1987) finds that incomplete contracts predominate when relationship-specific investment is low. For our setting, the most relevant constraint is the cost of enforcement of contracts. When enforcement costs are high, contracts tend to become simple. This is documented in the literature on law and finance by La Porta et al. (1997, 1998). They find that when rule of law and enforcement of property rights are weak, contracts remain incomplete. Thus, our setting is more relevant in these environments, which (as La Porta et al. document) occur in emerging markets and civil law countries.

The timing of the game is displayed in Figure 1.

2.2. Manager's Problem

A manager of type θ will join the firm if his expected wage exceeds his outside option, i.e., if $E[w \mid \theta] \geq \bar{u}$. Since the manager's expected wage $E[w \mid \theta] = s + b\gamma\theta$ is linear in θ, this results in two possibilities: If $b = 0$, the manager joins whenever $s \geq \bar{u}$, regardless of the manager's type (there is no sorting). If $b > 0$, there exists a threshold θ^*, such that $E[w \mid \theta^*] = \bar{u}$ and the manager joins if and only if $\theta \geq \theta^*$ (positive sorting). Note that $b < 0$ is never optimal. Now, since $E[w \mid \theta^*] = \bar{u}$, the threshold θ^* is

$$\theta^* = \frac{\bar{u} - s}{b\gamma}. \tag{4}$$

Thus, when $b > 0$, positive sorting takes place if and only if $1 > \theta^* > 0$. If $\theta^* \geq 1$, no manager is hired, and there is no production (we know this is not optimal because we have assumed that it is efficient to hire at least some type of manager). If $\theta^* < 0$, a manager of any type is hired. Note that when $b > 0$, $\theta^* > 0$ if and only if $s < \bar{u}$.

Finally, observe that when actual positive sorting occurs, the expected wage of the marginal type θ^* exactly equals his outside option ($E[w \mid \theta^*] = \bar{u}$), whereas every $\theta > \theta^*$ enjoys an information rent ($E[w \mid \theta] - \bar{u} > 0$). This information rent accrues because the firm, not knowing the true type of the manager, must offer a bonus that is high enough to attract even the lowest type of manager. Since the wage is increasing in a manager's type, this bonus will result in any higher type receiving a wage that is strictly higher than the outside option. If the firm knew the true type of any such manager, it would pay him just enough to induce him to join the firm.

[9] Although this may seem counterintuitive, think formally that total surplus rises in k and falls in m. In particular, m governs the slope of the surplus function, and θ^{FB} is the cutoff, where total surplus breaks even. Firms with high m have steep surplus functions and therefore can afford to hire less able managers; hence, they have lower thresholds. Firms with low m have surplus functions that rise slowly in θ, so a high θ is necessary to create value.

2.3. Firm's Problem

The firm earns profit from output, pays out wages, and bears the costs of employing the manager. Thus, the ex post expected profit for each θ is

$$E[\pi \mid \theta] = \gamma\theta(1-b) - s - (k - m\theta). \tag{5}$$

Thus, the firm's profits will rise in the manager's productivity θ and the quality of his match with the firm γ but will fall in the compensation parameters s and b. If the firm chooses to induce positive sorting, it will choose $b > 0$ and s to solve

$$\max_{b,\,s} \int_{\theta^*(s,\,b)}^{1} E[\pi \mid \theta] f(\theta)\, d\theta. \tag{6}$$

Write the marginal manager as $\theta^*(s, b)$ to illustrate this threshold's dependence on the contract parameters. Even without a moral hazard problem of the manager, the compensation contract has a role to play as a sorting instrument for the firm. The contract parameters s and b will affect the firm's payoff in two ways. First, they will determine the mix of types attracted to the firm; second, they will determine the firm's expected wage payments made to every manager who then joins the firm. This dual effect of the contract (determining participation ex ante and expected wage payments ex post) is apparent from (6), and will be a constant theme throughout, even under moral hazard and risk aversion.

Since there is no incentive problem in the basic model, the contract serves only to sort types. Ex post expected profit $E[\pi \mid \theta]$ decreases in salary and bonus, so sorting is costly for the firm. But it is necessary because the marginal type $\theta^*(s, b)$ decreases in s and b. Thus, as the firm raises either salary or bonus (decreasing ex post profits), it can raise ex ante profits because it attracts more types to the firm (thereby expanding the area of integration in (6)). The manager's participation depends only on whether his expected total wages exceed his outside options. He is effectively indifferent to receiving salary or bonus as long as he earns more at the firm than in the outside market. The firm, however, prefers to sort using salary rather than bonus. Intuitively, the manager's information rent $E[w \mid \theta] - \bar{u}$ increases in his bonus, since a high bonus boosts a high-type manager's compensation relative to his outside option. The firm seeks to minimize these information rents since, like all rents, they come at the cost of the firm's surplus through higher wages. In fact, the firm will pay as small a bonus as possible and a salary as close as possible to the manager's outside option. This is just enough to induce the desired set of managers to accept the job, leading to the first proposition. (All proofs are in the appendix.)

Proposition 1. *In the pure sorting model, the optimal contract consists of $b = \epsilon_b$ and $s = \bar{u} - \epsilon_s$ and induces $\theta^* = \theta^{FB} + \epsilon_\theta$, where ϵ_b, ϵ_s, and ϵ_θ are infinitesimal positive quantities.*

In the optimal contract, the firm sets a very small positive bonus. The salary is set so as to achieve the desired amount of sorting. Since the bonus is approximately zero, one should not be surprised that the desired sorting threshold is approximately equal to the efficient threshold, or that the salary approximately equals the manager's outside option. Note, however, that the firm does not set $s = \bar{u}$ and $b = 0$ *exactly*, because then every θ would weakly prefer to work at the firm, and there would be no sorting. By introducing a small positive bonus and reducing the salary appropriately, the firm can achieve positive sorting, since this modified contract is unattractive to low-productivity managers. This strictly improves the firm's profit.

In equilibrium, the firm achieves almost efficient sorting. It should be noted, however, that the actual sorting threshold is slightly *higher* than the efficient one, i.e., the firm screens out some managers whom it would be efficient to hire. The outcome is, therefore, not actually first-best but only *approximately* first-best. This deviation from the efficient sorting threshold, which happens through the firm's effort to avoid the additional information rents accruing to higher-type managers when the threshold is lowered to the efficient level, will recur more prominently in the full model with incentives, where even approximate first-best sorting will not be possible.

Why is a small bonus better at sorting than a large bonus? Everyone is happy with a large bonus, both the high types and the low types. But only the high types will accept a small bonus, because their high ability can outweigh the low per-unit pay from the small bonus. Low-types, on the other hand, receive a very low payoff when the bonus is small and, therefore, would not choose to join the firm. The small bonus thus achieves the separation of types more effectively than a large bonus does. In addition, because a high bonus benefits high types more than low types, raising the bonus raises the information rents accruing to the high types, due to the firm's need to ensure the participation of lower types; thus, a high bonus is more wasteful to the firm than a low bonus.

Proposition 1 shows that performance pay does, indeed, have a sorting effect but that this effect operates at very low PPS. To induce efficient sorting, one needs only a nonzero slope of the wage profile; more than this is unnecessary because it transfers unnecessary information rent to the agent. This is consistent with the vast empirical literature of PPS of executive contracts.[10] The sorting effect exerts downward pressure on performance pay.

3. Combining Sorting and Incentives

Now suppose that the manager exerts costly and unobservable effort at the firm. This induces a moral hazard problem on the part of the manager, since the firm cannot observe effort perfectly but must induce it through its compensation contract. At the same time, this same compensation contract is used to attract managers to the firm. Thus, contracts will now serve the dual purpose of attracting workers and providing incentives. Output is now given by

$$x = \gamma\theta e + \epsilon, \tag{7}$$

where ϵ still has mean 0 and variance σ^2. The manager exerts effort at a quadratic cost of e at $C(e) = 0.5ce^2$ with $c > 0$, and he continues to enjoy an outside option $\bar{u}$. Observe that the manager's type θ and effort choice e are complements, so more able types have higher marginal productivities of labor and are more productive to the firm. In fact, there are two levels of complementarity: between the ability θ and effort e, as well as between ability and the quality of the match between the firm and manager γ.

The firm pays the manager $w = s + bx$. As before, the contract is linear and the firm cannot condition the contract on the manager's type. The expected output for each manager θ is $E[x \mid \theta] = \gamma\theta e$, so more effort from the manager produces more revenue for the firm. For each θ, the average wage is $E[w \mid \theta] = s + b\gamma\theta e$. Assume the manager is risk neutral.[11] Figure 2 illustrates the timeline of the game.[12]

3.1. First Best

The manager has two decisions: whether to join the firm, and how hard to work. As such, the first best benchmark will also have two components: an efficient effort level, and an efficient cutoff for participation. Observe that the expected total surplus for each θ is

$$E[TS \mid \theta] = \gamma\theta e - C(e) - (k - m\theta) - \bar{u}. \tag{8}$$

Total surplus now not only includes expected output, the fixed cost of hiring a manager, and the manager's outside option, but also includes the manager's cost of effort. The first-best effort level that maximizes this is $e^{FB} = \gamma\theta/c$. Observe that the first-best effort level rises in both θ and γ. The total surplus obtained with first-best effort by TS^* is

$$E[TS^* \mid \theta] = \frac{(\gamma\theta)^2}{2c} - (k - m\theta) - \bar{u}. \tag{9}$$

Expected total surplus rises in both γ and θ and falls in c and k. This is positive if and only if

$$\theta > \frac{\sqrt{c[m^2c + 2\gamma^2(k+\bar{u})]} - mc}{\gamma^2} \equiv \theta^{FB}. \tag{10}$$

As before, θ^{FB} denotes an efficient cutoff, namely the minimal managerial type, such that it is efficient for the firm to employ any manager with $\theta > \theta^{FB}$. Observe that $\theta^{FB} > 0$, so it is always efficient to screen out some types of managers. In addition, just as before, we will assume that $\theta^{FB} < 1$ to avoid the trivial case where the firm cannot produce positive surplus with any manager.

The cutoff θ^{FB} rises in k, falls in γ, and rises in c. Thus, as technological or market factors lower the cost of supplying effort, it is efficient for the firm to become less selective regarding manager type and to have the manager work more.

3.2. Manager's Problem

A manager of type θ maximizes his expected wage less his cost of effort:

$$\max_e E[w \mid \theta] - C(e). \tag{11}$$

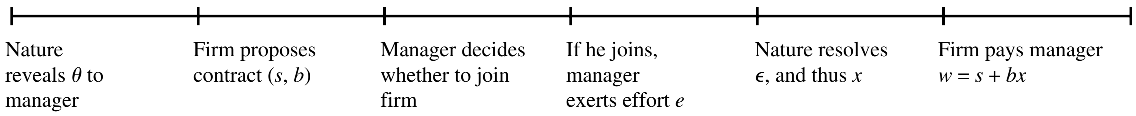
Given that $E[w \mid \theta] = s + b\gamma\theta e$ and $C(e) = 0.5ce^2$, the first-order condition yields the manager's incentive constraint (IC): $\hat{e} = b\gamma\theta/c$. Higher bonuses now have a clear incentive effect of inducing more effort. We can also safely ignore the possibility that $b \leq 0$ in equilibrium, since no surplus would be produced, and the firm would choose to exit the market. In addition, effort rises in both γ and θ, so more able types work more, as do types who are a better fit with the firm. This is exactly the sense in which there is complementarity in production: both γ and θ are complements with respect to effort. Finally, observe that $\hat{e} = e^{FB}$ if and only if $b = 1$, so that if the firm wanted to induce an efficient outcome, it would need to set $b = b^{FB} = 1$.

[10] Jensen and Murphy (1990) find that CEO wealth changes \$3.25 for every \$1,000 change in shareholder wealth. Morck et al. (1988) argue empirical estimates of PPS are "too low," since they deviate widely from theoretical predictions. Attempts to justify these low empirical estimates on risk aversion, such as Haubrich (1994), rely on estimating parameters of the model that are notoriously difficult to measure, such as the manager's cost of effort parameter.

[11] Section 4 considers a risk-averse manager.

[12] The timeline of the game opens the possibility of renegotiation of the contract after the manager accepts. But because the manager cannot communicate θ at any point, a manager who accepts the initial contract would also accept a renegotiated contract, since the information environment is unchanged. The firm's optimization problem is the same and, therefore, would offer the same contract to a manager after acceptance that it would before acceptance. Thus, the optimal contract would be renegotiation-proof. Details are available from the authors upon request.

Figure 2 Timeline of the Game with Effort



The participation decision now involves the manager's effort choice, which he makes conditional on facing a contract (s, b). A manager of type θ will join the firm if, in equilibrium, the manager earns more inside the firm than outside the firm, which occurs if $E[\hat{w} \mid \theta] - C(\hat{e}) \geq \bar{u}$, that is, $s + (b\gamma\theta)^2/(2c) \geq \bar{u}$. Since the left-hand side of this condition is strictly increasing in θ (due to $b > 0$), this simplifies to a threshold participation condition similar to the benchmark (no-effort) model. In particular, the manager joins if either $s > \bar{u}$[13] or $s \leq \bar{u}$ and $\theta > \theta^*$, where $\theta^* > 0$ satisfies $E[\hat{w} \mid \theta^*] - C(\hat{e}) = \bar{u}$, which yields

$$\theta^* = \frac{\sqrt{2c(\bar{u} - s)}}{b\gamma}. \tag{12}$$

Just as in the benchmark case without effort, the marginal type θ^* falls in both s and b, reinforcing the intuition that higher-wage payments attract more types to the firm. And just as θ^{FB}, θ^* rises in c.

3.3. Firm's Problem

The expected profit that the firm makes when hiring a manager of type θ is $E[\pi \mid \theta] = E[x \mid \theta] - E[w \mid \theta] - (k - m\theta)$. Using (IC), in equilibrium this becomes

$$E[\pi \mid \theta] = (\gamma\theta)^2 b(1-b)/c - s - (k - m\theta). \tag{13}$$

The firm will select a salary and bonus to maximize its expected profits. Thus, the firm will select s and b to maximize

$$\Pi(s, b) = \int_{\theta^*(s,\,b)}^{1} E[\pi \mid \theta] f(\theta)\, d\theta. \tag{14}$$

As before, writing $\theta^*(s, b)$ makes prominent the dependence of the marginal manager on the parameters of the compensation contract the firm sets. Because of the moral hazard problem, the compensation contract plays a dual role of both sorting types through θ^* and providing incentives through the expected wage $E[w \mid \theta]$. Solving this program explicitly is difficult because of the interaction between the sorting and incentive effects. In particular, b affects the manager's incentives to work, as well as the decision on whether to participate at all. Thus, the participation decision $\theta^*(s, b)$ is now endogenous. Nonetheless, the next proposition, proved in the appendix, presents the implicit solution that still has enough structure to provide insight into the dual sorting and incentive effects.

Proposition 2. *A firm contracting with a risk-neutral agent will select an optimal contract (s, b) that satisfies*

$$s = \left[\theta^*(1-b) - \frac{1 - F(\theta^*)}{f(\theta^*)} b\right] \frac{b\theta^*\gamma^2}{c} - (k - m\theta^*), \tag{15}$$

$$b = \left(2 - \frac{\theta^{*2}}{E[\theta^2 \mid \theta > \theta^*]}\right)^{-1}. \tag{16}$$

The sorting effect is immediately apparent in the optimal contract, since the contract now depends on the distribution of manager types. Recall that in the canonical agency problem, which includes principal contracting with a risk-neutral agent, the principal will make the agent the full residual claimant on the firm's output ($b^* = 1$) and will then take the rents back in the form of a negative "salary." This salary will be set just high enough to make the manager's expected payoff equal to his outside option; this is the "sell-the-firm" contract. The canonical model has empirical difficulties because empirical estimates of PPS are much lower than $b^* = 1$.

Adding sorting to the canonical model, however, moves the equilibrium away from the full "sell-the-firm" contract. For any $\theta^* \geq 0$, observe $E[\theta^2 \mid \theta > \theta^*] > \theta^{*2}$, and therefore $b < 1 = b^*$. Thus, the presence of sorting dampens the optimal bonus. This confirms the intuition from Proposition 1, which says that although a positive bonus is necessary to attract higher-quality managers to the firm, increasing the bonus too much is wasteful because it transfers excessive rents to the agent. In the full model considered here, the firm must provide incentives to the manager to induce him to work, which puts upward pressure on the bonus. The downward pressure from sorting, however, is still there: a high bonus transfers rents to any high-ability manager, which the firm tries to avoid by lowering the bonus and correspondingly increasing the salary (or, rather, reducing the constant transfer from the manager to the firm). The optimal bonus will trade off these twin effects, namely the downward pressure from sorting and the upward pressure from incentives.[14]

[13] In fact, it is obvious that $s > \bar{u}$ will never happen in equilibrium, since the firm could then increase its profits by lowering s to $\bar{u}$, which would decrease the rents of the manager without any adverse effects on participation or effort.

[14] If b_S is the bonus under sorting alone, b_I the bonus under incentives alone, and b_{IS} the bonus under both incentives and sorting, then $0 \approx b_S < b_{IS} < b_I = 1$. This gives a strict ordering. Dutta (2008) finds that $b_{IS} < b_I$, but does not offer a clear ordering between b_S and either b_{IS} or b_I.

How exactly will the firm trade off the optimal choice of salary and bonus? In the canonical model of a risk-neutral agent without sorting, there is a clean separation between salary and bonus; the bonus provides incentives, while the salary guarantees participation. Here, although the salary does not play a role in incentives, bonuses do affect sorting, since the marginal manager θ^* falls in b. Just like salaries, higher bonuses will attract more types to the firm. The proof of Proposition 2 details the first- order condition for the firm's optimization with respect to the optimal salary and bonus. This leads to the equilibrium condition

$$\frac{\partial\theta^*/\partial s}{\partial\theta^*/\partial b} = \frac{(\partial\theta^*/\partial s)\, E[\pi \mid \theta^*] f(\theta^*)}{(\partial\theta^*/\partial b)\, E[\pi \mid \theta^*] f(\theta^*)} = \frac{\int_{\theta^*}^{1} \partial E[\pi \mid \theta]/\partial s \, dF}{\int_{\theta^*}^{1} \partial E[\pi \mid \theta]/\partial b \, dF}. \tag{17}$$

Recall from §2 that sorting gives contracts a dual effect, namely, determining participation ex ante and the expected wage payments ex post. The same is true here, after including a moral hazard problem. Each piece of the compensation contract will have an ex post (right-hand term of (17)) and an ex ante effect (middle term of (17)). The equilibrium condition above states that the firm will optimally equalize the ratios of these ex ante and ex post effects. This is equivalent to equalizing the marginal rate of substitution between salary and bonus along the optimal participation threshold curve to each of those ratios. The firm chooses its salary and bonus such that the tradeoff between the costs of wages against the benefits of participation is equal across both contract parameters. At the equilibrium, the firm is indifferent between using salary and bonus.

Now we are firmly in a second-best world. Recall that in the effortless benchmark model the equilibrium was *approximately* efficient. With effort choice added, even approximate efficiency is too much to ask. We have already argued that to reduce information rents accruing to the manager, the firm will reduce the bonus from its efficient level (of one), which will result in inefficiently low effort. We will now see that the sorting threshold, which the firm chooses to induce by the equilibrium contract, is also inefficient.

Proposition 3. *When the agent is risk neutral, the firm's equilibrium sorting threshold exceeds its efficient level*: $\theta^* > \theta^{FB}$.

This result should not be too surprising. Even in the effortless benchmark model, in an attempt to reduce information rents, the firm turned away some types of managers whom it would have been efficient to hire, i.e., the equilibrium sorting threshold was slightly higher than the efficient threshold. In that model, the inefficiency was negligible since the sorting threshold converged to the efficient level as the bonus converged to zero. In the present model, however, this convergence does not happen, and the sorting threshold remains firmly above the efficiency benchmark.

To understand why the firm chooses to be more selective than an efficiency-driven planner, consider the marginal effects of raising the threshold slightly (by $\eta \approx 0$) above the efficient level. Since the total surplus and, hence, the profit generated by the marginal manager is zero, and the profit function is continuous, each of the types no longer hired will decrease the total expected profit by only a small amount, resulting in a total profit loss of order η^2. By not hiring these types, the firm can lower the salary by an amount of order η. This will result in a profit gain in the form of salary savings from *all* manager types, which will yield a total profit gain of order η. Clearly, for small η, the gain will exceed the loss ($\eta > \eta^2$), so that increasing the threshold will be profitable.

3.4. Comparative Statics

Our goal is not only to solve for the optimal contract, but also to conduct comparative statics, seeing how the optimal contract varies with the exogenous parameters in the model. We do this with an eye for predicting some of the empirical cross-sectional variation in managerial pay. Indeed, Murphy (1999) shows substantial cross-sectional variation in median CEO PPS across both industry and company size. We believe the variation along these two dimensions indicates the existence of variation across different variables. We hope our comparative statics will inspire empirical researchers to test whether the parameters in our model generate the cross-sectional variation we expect to see in real data.

The optimal salary and bonus both depend on the distribution of managerial talent, F. For tractability, we assume, throughout this section, that this distribution is uniform on $[\mu - a, \mu + a) \subseteq [0, 1)$.[15] However, numerical simulations suggest similar results for a much wider class of distributions.[16] We first investigate the effects of increases in what Dutta (2008) calls *information risk*, namely, the variance of the distribution of manager types.

Corollary 1. *When the variance of the type of manager increases, the PPS* (b) *decreases and the sorting threshold* (θ^*) *increases*.

[15] Strictly speaking, this violates the original assumption that the support of f is $[0, 1]$; however, this particular deviation is without loss of generality as long as $\mu - a < \theta^{FB} < \mu + a$.

[16] Numerical simulations for all of the corollaries in the paper are available in the online appendix, located at http://www.korokray.com, under the section "Academic Papers."

An increase in the variance of the type distribution has no effect on the expected surplus from a given manager type, given a fixed salary and bonus. However, since the increase in this variability increases the uncertainty that the firm has about the manager's type, the information asymmetry problem is exacerbated, and the information rents accruing to the manager increase. We saw in the discussion of Propositions 1 and 2 that the firm's desire to avoid information rents drives down the bonus. In Proposition 3, we saw the same considerations force the sorting threshold up. Thus, to counteract the upward pressure that the increase in manager type variability exerts on information rents, the firm can either lower the bonus or raise the sorting threshold. The proof of Corollary 1 shows that it will, in fact, optimally choose to do both.

This corollary is in contrast to the result of Baker and Jorgensen (2003), who found that the optimal bonus can increase in the variance of θ. In Baker and Jorgensen's model, the firm wants the manager to make use of his private information. Increasing the bonus induces the manager to make more use of this information, because its value rises due to a rise in the variance of θ. In our model, the focus is not on the efficient utilization of information by the manager but rather on the use of contracts to induce self-selection of productive managers into the firm. The fact that the value of information in our model does not directly affect the manager's choices removes the upward pressure of information risk on the bonus; meanwhile, the use of the contract for selection creates a downward pressure, because it means a rise in the type variance will increase information rents to the manager. Measuring information risk is difficult but far from impossible, given the wide variety of managerial performance measures available today. Empirical proxies include the manager's experience level, age, level of general versus specific human capital, or "thickness" of the managerial labor market.

Corollary 2. *As the manager's cost of employment* (k) *or outside option* ($\bar{u}$) *rises, the firm will increase both the PPS* (b) *and the sorting threshold* (θ^*).

First, note from the expression for the bonus in Proposition 2 that k and $\bar{u}$ affect the optimal bonus only indirectly, through their effects on the optimal sorting threshold θ^*. In particular, at least for uniform distributions, a higher θ^* is always accompanied by a higher bonus—this is intuitive because a higher sorting threshold reduces the information rents. Therefore, this induces the firm to increase the bonus to return to the earlier balance between incentives and information rents. The effect on the sorting threshold θ^*, in turn, is unambiguously positive: A rise in either the employment cost, or the manager's outside option, lowers the surplus generated by each manager type, forcing the firm to become more selective in hiring. As a result, an increase in k or $\bar{u}$ increases both the sorting threshold and the bonus. Finally, the effect on the salary is unambiguously negative in the case of k (because s decreases in both θ^* and b, and k has no direct effect on the manager's surplus), but ambiguous in the case of $\bar{u}$ (the direct upward effect of $\bar{u}$ on the salary through decreased manager surplus is counterbalanced by the indirect downward effect from an increase in θ^* and b).

Corollary 2 predicts that in industries with high fixed costs of hiring, firms will employ higher bonuses and lower salaries. Empirical proxies of k can be diverse. One example is a measure of firm or industry-specific human capital. For instance, a firm in the finance or technology sectors, which require deep, specialized knowledge, may have a high k; meanwhile, a more general consumer products or retail firm, which requires more general business knowledge, may correspond to a low k. One possible way to measure k may be the recruitment of managers across different industries or sectors. An industry that recruits future managers from within the same industry is likely to be one with a high fixed cost of training. That investment banks hire CEOs with long careers in the financial sector may indicate a high requirement for industry-specific knowledge and, therefore, a high k. Corollary 2 shows that such financial managers will receive high bonuses and low salaries, consistent with conventional wisdom and empirical surveys.

4. Risk Aversion

Now suppose that the manager is risk averse and dislikes volatility in his income. Without sorting, the canonical agency model with a risk-averse agent delivers the standard risk-incentives trade-off, where the principal seeks to provide incentives for the agent to work; however, such incentives load risk onto the agent, which he dislikes. How does sorting affect this analysis?

To fix ideas, suppose that production uncertainty ϵ is normally distributed and has a mean of 0 and a variance of σ^2. Assume the agent has constant absolute risk aversion (CARA) preferences with coefficient of absolute risk aversion $r > 0$. The timing of the game is the same as before, as are the first-best effort level e^{FB} and first-best participation cutoff θ^{FB}.

4.1. Manager's Problem

Given output $x = \gamma\theta e + \epsilon$ and a linear compensation contract $w = s + bx$, the certainty equivalent of the θ-manager's payoff is now

$$CE(\theta) = s + b\gamma\theta e - \frac{r}{2}b^2\sigma^2 - C(e). \tag{18}$$

Observe that a risk-averse manager now bears a disutility from uncertainty captured by the risk-premium term $(r/2)b^2\sigma^2$. This risk premium is the additional compensation necessary to induce a risk-averse manager to accept risk. Further observe that the risk premium is independent of the manager's effort choice. Therefore, a risk-averse manager's effort choice problem is identical to that of a risk-neutral manager, leading to identical incentive constraints. In particular, given a quadratic cost of effort $C(e) = 0.5ce^2$, the incentive constraint (IC) for a manager of type θ is still $\hat{e} = b\gamma\theta/c$. As with a risk-neutral manager, we can safely ignore the possibility that $b \le 0$ in equilibrium, since this would lead the firm to exit the market due to the inability to produce.

Given the contract (s, b), a manager of type θ will participate if and only if the certainty equivalent of the contract exceeds his outside option $\bar{u}$, or $CE(\theta) \ge \bar{u}$. Given equilibrium effort choice, this condition becomes $s + (b\gamma\theta)^2/(2c) - (rb^2\sigma^2)/2 \ge \bar{u}$. As before, the left-hand side of this condition is strictly increasing in θ for all positive θ (due to $b \ne 0$), making this a threshold participation condition. In particular, the manager joins if either $s > \bar{u} + rb^2\sigma^2/2$ (this is wasteful to the firm and will not happen in equilibrium) or $s \le \bar{u} + rb^2\sigma^2/2$ and $\theta > \theta^*$, where $\theta^* > 0$ satisfies $CE(\theta) \ge \bar{u}$, so that

$$\theta^* = \frac{\sqrt{2c(\bar{u} - s) + crb^2\sigma^2}}{b\gamma}. \qquad (19)$$

This threshold differs from the threshold for risk-neutral managers (Equation (12)) by the presence of the risk-premium term $crb^2\sigma^2$: the marginal type θ^* rises in the measure of the agent's risk aversion and in the volatility of the output. Indeed, because the expected wage increases in the manager's type, whereas (with CARA preferences) the cost of risk is independent of it, only a manager of higher θ could earn a high enough expected wage to offset the cost of risk. The higher that cost (i.e., the higher the level of risk aversion and the volatility of the output), the higher the manager's type needs to be, thus making the risk bearable. This implies that raising output volatility will not only cause fewer types to participate but will also ensure that the participating types are better. Insofar as it is possible to measure production uncertainty σ^2 and managerial type θ, this predicts that better types work in more uncertain environments.

4.2. Firm's Problem

The expected profit that the firm makes if it hires a manager of type θ is $E[\pi \mid \theta] = E[x \mid \theta] - E[w \mid \theta] - (k - m\theta)$. Using (IC), this becomes (13).

The form for expected profit is the same as before, namely, in the prior section with a risk-neutral agent. This occurs because risk aversion does not alter the agent's incentive constraint (although it does affect his participation decision). Since the incentive constraint is the same, conditional on hiring a manager θ, the firm's ex post profits from that manager are unchanged in the presence of risk aversion. For sure, risk aversion will affect the optimal bonus, which in turn will drive both incentives and participation. But taking this bonus as given, risk aversion does not change the expected profit function.

The firm will select s and b to maximize its expected profits. The equilibrium conditions take the same form as before, but there are a few key differences. Most importantly, $\partial\theta^*/\partial b$ has now changed to

$$\frac{\partial\theta^*}{\partial b} = \frac{rc\sigma^2}{\gamma^2 b\theta^*} - \frac{\theta^*}{b} = \frac{\theta^*}{b}\left(\frac{rc\sigma^2}{(\gamma\theta^*)^2} - 1\right). \qquad (20)$$

This expression makes the effect of risk aversion clear. Under risk neutrality, $r = 0$ and this derivative becomes unambiguously negative, just like $\partial\theta^*/\partial s$.[17] Increasing either the salary or the bonus raises the total compensation of the agent, making the contract more attractive to outsiders and thus drawing greater participation from the labor market, and reducing the marginal type θ^*. But now, $r > 0$ implies that risk aversion enters the participation decision. Raising incentives no longer unambiguously increases participation, since higher bonuses load more risk onto the compensation, which risk-averse agents dislike. If the levels of risk aversion and output volatility are sufficiently low, the positive effect that higher bonuses exert on the expected value of the compensation will dominate the unfavorable effect on the cost of risk. Thus, raising the bonus will attract more types, i.e., lower the sorting threshold, just as it did with risk-neutral managers. But, if risk aversion and output volatility are sufficiently high (namely, if $r\sigma^2 > \gamma^2\theta^{*2}/c$), then raising incentives will actually repel more types, and will cause the marginal type θ^* to rise in b. This shows that the relationship between performance pay and participation is subtle in the presence of risk aversion. Since risk-averse agents dislike uncertainty in their compensation, they may stay away from firms that offer large incentive packages.

Solving for the optimal contract in closed form is even more difficult than before, since risk aversion adds additional complexity to the participation decision. The participation threshold θ^* now depends on the bonus b in a nonmonotonic way, making the first-order condition for profit maximization more complex. Nonetheless, it is still possible to arrive at an implicit solution that can deliver intuition.

[17] Irrespective of risk preferences, $\partial\theta^*/\partial s = -c/(b^2\gamma^2\theta^*) < 0$.

Proposition 4. *A firm contracting with a risk-averse agent will select an optimal contract* (s, b) *that satisfies*

$$s = \left[\theta^*(1-b) - \frac{1-F(\theta^*)}{f(\theta^*)}b\right]\frac{b\theta^*\gamma^2}{c} - (k - m\theta),$$

$$b = \left(\frac{rc\sigma^2}{\gamma^2 E[\theta^2 \mid \theta > \theta^*]} + 2 - \frac{\theta^{*2}}{E[\theta^2 \mid \theta > \theta^*]}\right)^{-1}.$$

The formula for the optimal bonus in Proposition 4 is more complex than in Proposition 2. The tension between incentives and information rents is still there (as represented by the term $2-\theta^{*2}/E[\theta^2 \mid \theta > \theta^*]$, in the denominator, for the bonus of both risk-neutral and risk-averse managers), but now an additional force has entered the scene: risk aversion. The direct effect of this on the bonus is clearly negative, because the risk-aversion term in the denominator is positive and increasing in the levels of risk aversion and output variability when holding the sorting threshold fixed.

Introducing risk aversion fundamentally changes the nexus of trade-offs that the firm faces when choosing salary and bonus. Without risk aversion, salary and bonus are always substitutes and, therefore, are two equivalent ways of attracting the worker. But under sufficiently high risk aversion, they become complements. When risk aversion is high, a high bonus loads a large disutility onto the manager, and this, in fact, repels the manager from the firm. To rein him back in, the firm must counteract by raising his salary.

Recall that even without risk aversion, the interplay of the sorting and incentive effects of performance pay sent us into a second-best world, causing effort to be lower than first-best (due to the bonus being reduced from its efficient level of one), and making the firm turn away some managers whom it would have been efficient to hire. We suggested above, and will firmly establish in Corollary 4, that risk aversion further exacerbates the inefficiency of the incentives, pushing the bonus even further down and away from its efficient level (and thus similarly reducing the equilibrium effort). But what about the sorting? Because of the complex relationship between the bonus and the hiring threshold in the presence of risk aversion, one might wonder whether risk aversion might do away with the inefficiently high hiring threshold we saw with risk-neutral managers. The following result, however, shows that this is not the case. Even when agents are risk averse, the firm still chooses to turn away some managers who would have produced positive surplus.

Proposition 5. *When the agent is risk averse, the firm's equilibrium sorting threshold exceeds its efficient level*: $\theta^* > \theta^{FB}$.

The proof of the proposition follows along the same lines as the proof of Proposition 3. What causes the firm to set the threshold inefficiently high is its desire to avoid paying high information rents to the manager. When the firm increases the threshold slightly above the efficient level, the amount of money it saves, by paying lower information rents to all types of managers, is larger than the amount it loses by not hiring if the manager happens to be of the marginal type.

4.3. Comparative Statics

We begin this section by asking ourselves how the need to use the compensation scheme for sorting affects the optimal contract in the presence of risk aversion. Rather than repeating the same analysis from Corollaries 1 and 2, the new comparative statics will ask how PPS varies with the new parameter in the model, risk aversion. To this end, consider the canonical agency model, where the firm can perfectly observe the manager's type, so that there is no adverse selection. All other elements of the model remain the same as in the full model we have been investigating. In particular, output is still $x = \gamma\theta e + \epsilon$ (with $\epsilon \sim N(0, \sigma^2)$), and the firm's cost of employing the manager is $k - m\theta$. The manager still has cost of effort $C(e) = 0.5e^2$, constant absolute risk aversion r, and outside option $\bar{u}$. The firm is still constrained to linear contracts. As shown in the proof of Corollary 3, the optimal bonus for a manager of type θ in this canonical model (the "no-sorting bonus") is $b_{NS} = (1 + rc\sigma^2/(\gamma^2\theta^2))^{-1}$. Note that the no-sorting bonus increases in the manager's type, whereas the bonus in the full model (the "sorting bonus") does not depend on it (since the type is not observable).

In the canonical model, the optimal bonus falls in the cost-of-effort parameter c, the coefficient of absolute risk aversion r, and the variance of output σ^2. The direct effects (when holding the sorting threshold fixed) of all these parameters on the bonus in the full model are obviously identical to those in the canonical model. Nonetheless, unlike in the canonical model, we cannot immediately conclude that the full equilibrium effects will be the same: the equilibrium condition for the bonus also depends on the sorting threshold, which itself will vary as the parameters change. Much more careful analysis is needed to determine the overall equilibrium effects.

We saw before that, without risk aversion, sorting invariably depresses the bonus relative to the benchmark without sorting. The presence of risk aversion makes this relationship more subtle. When comparing the formula for b_{NS} to b in Proposition 4, we can decompose the difference of their inverses into two terms:[18]

$$b^{-1}-b_{NS}^{-1}=Rent+Risk,\qquad \text{where}$$

$$Rent=1-\frac{\theta^{*2}}{E[t^2\mid t>\theta^*]}\qquad\text{and}$$

$$Risk=-\frac{rc\sigma^2}{(\gamma\theta)^2}\left(1-\frac{\theta^2}{E[t^2\mid t>\theta^*]}\right).$$

The first term, $Rent>0$, represents the downward pressure on the sorting bonus due to information rents (and equals the entire difference between the sorting and no-sorting bonuses when there is no risk aversion). The second term, *Risk*, arises because the sorting bonus adjusts the bonus for risk based on average ability, whereas the no-sorting bonus tailors the risk adjustment to each specific type. Recall that the risk adjustment in both cases drags the bonus down (since a higher bonus loads undesirable risk onto the manager). Furthermore, the tailored risk adjustment decreases in manager ability, because higher-ability managers have higher output and the same risk as lower-ability managers, making risk relatively less important for the more able ones. Consequently, the average adjustment is lower than the tailored adjustment for managers of low ability, and the opposite ordering holds for managers of high ability. Thus, risk drags the average-based sorting bonus down less (respectively, more) than the tailored no-sorting bonus for low-ability (and, respectively, high-ability) managers. Mathematically, we see that *Risk* is increasing in θ, with $Risk<0$ for $\theta<E[t^2\mid t>\theta^*]$ and $Risk>0$ for $\theta>E[t^2\mid t>\theta^*]$.

We have thus determined that for high-ability managers, both the information rents effect (*Rent*) and the average-versus-tailored risk-adjustment effect (*Risk*) drag the sorting bonus down relative to the no-sorting bonus. For low-ability managers, however, the two effects move in opposite directions: while the rents effect still pushes the sorting bonus down relative to the no-sorting bonus, the risk-adjustment effect works against it. Since the risk-adjustment effect is increasing in the manager's risk aversion and output risk, it can overtake the rents effect when the values of these parameters are high. Summarizing these observations, we arrive at the following result.

Corollary 3. *For high-ability managers, sorting always dampens PPS, regardless of the level of risk aversion. Sorting increases PPS for low-ability managers if and only if $r\sigma^2$ is high.*

The proof of the corollary shows that there always exists thresholds $\underline{R}$ and $\bar{R}$, such that sorting dampens PPS for all managers if $r\sigma^2<\underline{R}$, whereas sorting increases PPS for some low-ability managers if $r\sigma^2>\bar{R}$.[19] The proof also shows that $\underline{R}$ and $\bar{R}$ are increasing in the productivity parameter γ (due to risk becoming relatively less important when output is high). Thus, sorting becomes more likely to increase the bonus relative to the perfect information benchmark when the manager is not very productive at his current firm (the complementarity between the firm and the manager is low). When the complementarity/productivity parameter γ is low, the risk-adjustment effect can come to dominate the information rents effect for low-ability managers, causing such managers' bonuses to be higher in the presence of sorting than they would be otherwise. This stands in contrast to prior work, such as Dutta (2008), who finds that optimal PPS, under symmetric information, always exceeds optimal PPS under asymmetric information (when managerial talent is firm-specific, the relevant case here).

We have demonstrated that risk aversion can cause sorting to increase the optimal PPS, but we should note that the level of risk aversion required for this to occur is quite extreme. In particular, the proof of Corollary 3 shows that sorting increases the PPS for some low-ability managers if and only if $r\sigma^2>\gamma^2(\theta^*)^2/c$. But from the definition of the threshold θ^*, we see that the salary is $s=\bar{u}-[\gamma^2(\theta^*)^2/c-r\sigma^2]b^2/2$, so that $r\sigma^2>\gamma^2(\theta^*)^2/c$ if and only if $s>\bar{u}$. Thus, sorting can increase a manager's PPS only if the manager is so risk averse that he views his equilibrium level of bonus as a punishment, i.e., he needs to be paid a fixed salary exceeding his outside option to accept the risk loaded onto him by the bonus.

We conclude this section by investigating how risk aversion and output uncertainty affect the optimal contract. For tractability, we assume that the manager type distribution is uniform.

Corollary 4. *When either the manager's level of risk aversion (r) or output variability (σ^2) increases, the firm will increase the sorting threshold (θ^*) and decrease the PPS (b).*

As we already observed in the discussion after Proposition 4, the direct effect of increases in risk aversion and output uncertainty on the bonus is unambiguously negative, as higher costs of risk reduce the returns from a higher bonus, since the manager values the risky part of its compensation less. However, the increased costs of risk also

[18] The expectations are taken with respect to the manager type distribution; the symbol t is used to avoid confusion between the variable of integration t and the fixed value θ.

[19] The thresholds $\underline{R}$ and $\bar{R}$ provide sufficient, but not necessary conditions for the two cases. Either of the two scenarios can occur when $\underline{R}<r\sigma^2<\bar{R}$.

decreases the overall surplus produced by a given type of manager, therefore compelling firms to stop hiring previously marginal types. In response, the firm raises the sorting threshold. This, in turn, has the side effect of reducing information rents, which allows the firm to raise the bonus. Nonetheless, the proof of the corollary shows that, at least with a uniform type distribution, this latter effect does not overcome the former. The combined effect is unambiguously negative on the bonus and unambiguously positive on the sorting threshold.

Corollary 4 also reinforces the classic trade-off between risk and incentives: as output variability rises, PPS decreases. The higher output variability leads to more selective hiring, which decreases total information rents because of a smaller pool manager types. Although the firm could afford to increase PPS because of these smaller information rents, doing so would load more risk onto the manager and make him less likely to accept the job (and could adversely affect participation and therefore profits for the firm). Just as a large empirical literature has tested the classic risk incentives trade-off (see Prendergast 2002), so too can our corollaries be compared to data. For example, PPS and output variation are fairly straightforward to measure, with the former estimated through compensation contracts, proxy statements, and disclosures of executive pay in financial statements. Measuring the sorting threshold is harder but not impossible. It would require observing applications and decisions for managerial jobs—counting who applied, who was offered, and who was accepted. These more granular ; on the labor market are becoming increasingly available and could then be used to test how the sorting threshold changes with the parameters of the model (r, σ^2, c, etc.).

5. Conclusion

The vast academic literature on performance pay has focused almost exclusively on incentive effects. This paper incorporates the dual incentive and sorting effects of performance pay and puts forth a simple and tractable model that provides basic insight into how a firm can solve the two problems jointly. The primary goal of this paper is to understand how sorting affects the firm's design of optimal compensation. Rather than focusing on inducing appropriate communication as in prior work, sorting concerns how the contract will affect the type of manager attracted to the firm. This new logic confirms some existing work (such as incentives raising PPS, while selection reduces PPS), but also provides new insights (sorting can inflate PPS if the manager's risk aversion is high or manager-firm complementarity is small; sorting under imperfect information causes the firm to turn away more candidates than would be efficient). Future work in this area will expand on the sorting and incentive effects of performance pay to include repeated interactions, retention effects, subjective bonuses, and multiple managers.

Acknowledgments

The authors thank Allan Timmermann, Cindy Alexander, Alex Lee, Richard Saouma, Yun Zhang, V. G. Narayanan, Jack Hughes, and Lee Pinkowitz, as well as seminar participants at George Washington University, Georgetown University, Harvard Business School, University of California, San Diego Rady School of Management, the U.S. Securities and Exchange Commission Office of Economic Analysis, and the American Accounting Association junior faculty theory conference. John Cannavo provided outstanding research assistance. George Washington University and PwC Faculty INQuires provided financial support.

Appendix. Proofs

Proof of Proposition 1. The firm's task is to choose b and s to maximize total expected profit. This reduces to the following problem:

$$\max\left\{\max_{b>0;\,s}\int_{\max(\theta^*(s,b),\,0)}^{1} E[\pi\mid\theta]f(\theta)\,d\theta,\right.$$
$$\max_{b<0;\,s}\int_{0}^{\min(\theta^*(s,b),\,1)} E[\pi\mid\theta]f(\theta)\,d\theta,$$
$$\left.\int_0^1 E[\pi(b=0;s=\bar{u})\mid\theta]f(\theta)\,d\theta,\,0\right\}.$$

Here, the first term corresponds to choosing the optimal positive bonus; the second term, to choosing the optimal negative bonus; the third term, to choosing $b=0$ and inducing participation (note that this means setting $s=\bar{u}$, since s above that level is wasteful to the firm), and the fourth term to setting $b=0$ and paying $s<\bar{u}$, so there is no participation.

It will be useful to let $\overline{TS}(\theta^*)$ be the expected total surplus with positive sorting with threshold θ^*. Note that we have already determined that $\overline{TS}(\theta^*)<\overline{TS}(\theta^{FB})$ for any $\theta^*\neq\theta^{FB}$. It will also be useful to denote by Π the total expected profit of the firm.

We begin by observing that when $\mathbf{b}=\mathbf{0}$, the expected profit is always strictly less than first-best total surplus $\overline{TS}(\theta^{FB})$. When $b=0$ and there is no participation, $\Pi=0<\overline{TS}(\theta^{FB})$. When $b=0$ and there is participation, $s=\bar{u}$, and $E[\pi\mid\theta]=E[TS\mid\theta]$ for all θ, so that $\Pi=\int_0^1 E[TS\mid\theta]f(\theta)\,d\theta=\overline{TS}(0)<\overline{TS}(\theta^{FB})$.

To explore the cases when $b>0$ or $b<0$, it will be convenient to recast the problem as one of choosing b and θ^*. Since the threshold is given by $b\gamma\theta^*=\bar{u}-s$, we can eliminate s from the problem by noting that $s=\bar{u}-b\gamma\theta^*$. Thus,

$$E[\pi\mid\theta]=(m+\gamma)\theta+b\gamma(\theta^*-\theta)-\bar{u}-k,$$

so that $\partial E[\pi]/\partial b=\gamma(\theta^*-\theta)$ and $\partial E[\pi]/\partial\theta^*=b\gamma$. Also note that $E[\pi\mid\theta=\theta^*]=(m+\gamma)\theta^*-\bar{u}-k=E[TS\mid\theta^*]$.

Now, when $b>0$,

$$\Pi=\int_{\theta^*}^{1} E[\pi\mid\theta]f(\theta)\,d\theta,$$

so that, for any $b > 0$,

$$\frac{\partial \Pi}{\partial b} = \int_{\theta^*}^{1} \gamma(\theta^* - \theta) f(\theta)\, d\theta < 0.$$

It follows that, in this range, it is optimal for the firm to set b as low as possible. In particular, if ϵ_b is the lowest available currency unit, the firm will set $b = \epsilon_b$.

Next, note that

$$\frac{\partial \Pi}{\partial \theta^*} = -[(m+\gamma)\theta^* - \bar{u} - k] + b\gamma[1 - F(\theta^*)],$$

and $\partial^2 \Pi / \partial \theta^{*2} = -(m+\gamma) - b\gamma f(\theta^*) < 0$, so that a necessary and sufficient condition for optimal θ^* is

$$[(m+\gamma)\theta^* - \bar{u} - k] = b\gamma[1 - F(\theta^*)].$$

Noting that the left-hand side is $E[TS \mid \theta^*]$, this reduces to

$$E[TS \mid \theta^*] = b\gamma[1 - F(\theta^*)].$$

Since the right-hand side is positive, it immediately follows that $\theta^* > \theta^{FB}$. We can thus write $\theta^* = \theta^{FB} + \epsilon_\theta$, where $\epsilon_\theta > 0$. Furthermore, as b approaches 0, the entire right-hand side of the first-order condition above also approaches 0. By continuity and monotonicity of $E[TS \mid \cdot]$, this implies that θ^* approaches $E[TS \mid \cdot]^{-1}(0) = \theta^{FB}$. Therefore, the fact that $b = \epsilon_b$ is infinitesimal implies that ϵ_θ is also infinitesimal. It now also follows that $s = \bar{u} - \epsilon_s$, where $\epsilon_s = b\gamma\theta^* > 0$ is also infinitesimal. Finally, $\Pi = \int_{\theta^*}^{1} E[\pi \mid \theta] f(\theta)\, d\theta \to \int_{\theta^{FB}}^{1} [(m+\gamma)\theta - \bar{u} - k] f(\theta)\, d\theta = \overline{TS}(\theta^{FB})$ as $b \to 0$.

We have now characterized the optimal contract when $b > 0$ and have shown that as long as ϵ_b is sufficiently small, Π from $b = \epsilon_b$ can get arbitrarily close to $\overline{TS}(\theta^{FB})$, which is more than what could be achieved with $b = 0$. Hence, setting $b > 0$ dominates setting $b = 0$. To complete the proof that $b > 0$ is the unique optimal solution, we now need only to show that $b < 0$ gives a lower profit.

Now, if $b < 0$,

$$\Pi = \int_{0}^{\theta^*} E[\pi \mid \theta] f(\theta)\, d\theta.$$

Noting that $E[\pi \mid \theta] < E[TS \mid \theta]$ for all $\theta < \theta^*$, we see that $\Pi < \int_0^{\theta^*} E[TS \mid \theta] f(\theta)\, d\theta$. Since $E[TS \mid \theta]$ is increasing in θ, we know that if this expression is positive, then $E[TS \mid \theta] > 0$ for all $\theta \geq \theta^*$, so that

$$\Pi < \int_0^{\theta^*} E[TS \mid \theta] f(\theta)\, d\theta + \int_{\theta^*}^{1} E[TS \mid \theta] f(\theta)\, d\theta$$
$$= \int_0^1 E[TS \mid \theta] f(\theta)\, d\theta = \overline{TS}(0) < \overline{TS}(\theta^{FB}).$$

Thus, $b < 0$ always gives profit that is lower than the maximum achievable with $b > 0$ (as long as ϵ_b is sufficiently small). This completes the proof. $\square$

Proof of Proposition 2. The firm chooses salary s and bonus b to solve

$$\max_{s,\,b} \int_{\theta^*(s,\,b)}^{1} E[\pi \mid \theta] f(\theta)\, d\theta.$$

Using Leibnitz's rule, the first-order conditions are

$$E[\pi \mid \theta^*] f(\theta^*) \frac{\partial \theta^*}{\partial s} = \int_{\theta^*}^{1} \frac{\partial E[\pi \mid \theta]}{\partial s} f(\theta)\, d\theta,$$
$$E[\pi \mid \theta^*] f(\theta^*) \frac{\partial \theta^*}{\partial b} = \int_{\theta^*}^{1} \frac{\partial E[\pi \mid \theta]}{\partial b} f(\theta)\, d\theta.$$

Combining these two equations leads to the equilibrium condition

$$\frac{\partial \theta^*}{\partial b} \int_{\theta^*}^{1} \frac{\partial E[\pi \mid \theta]}{\partial s} f(\theta)\, d\theta = \frac{\partial \theta^*}{\partial s} \int_{\theta^*}^{1} \frac{\partial E[\pi \mid \theta]}{\partial b} f(\theta)\, d\theta. \tag{21}$$

Now, the ex post expected profit is $E[\pi \mid \theta] = E[x \mid \theta] - E[w \mid \theta] - (k - m\theta) = \gamma\theta e(1-b) - s - (k - m\theta)$. Plugging in the incentive constraint $e = b\gamma\theta/c$,

$$E[\pi \mid \theta] = \frac{(\gamma\theta)^2}{c} b(1-b) - s - (k - m\theta). \tag{22}$$

The derivatives of expected profit with respect to salary and bonus are, respectively,

$$\frac{\partial E[\pi \mid \theta]}{\partial s} = -1 \quad \text{and} \quad \frac{\partial E[\pi \mid \theta]}{\partial b} = \frac{(\gamma\theta)^2}{c}(1 - 2b). \tag{23}$$

Recall that $\theta^* = \sqrt{2c(\bar{u} - s)}/(b\gamma)$. The partial derivatives are

$$\frac{\partial \theta^*}{\partial s} = \frac{-c}{(b\gamma)^2 \theta^*} \quad \text{and} \quad \frac{\partial \theta^*}{\partial b} = -\frac{\theta^*}{b}. \tag{24}$$

Combining these gives

$$\frac{\partial \theta^*}{\partial b} = \frac{\partial \theta^*}{\partial s} \cdot \frac{b\theta^{*2}\gamma^2}{c}. \tag{25}$$

Combining (23) with the equilibrium condition (21) gives

$$-\frac{\partial \theta^*}{\partial b}(1 - F(\theta^*)) = \frac{\partial \theta^*}{\partial s} \int_{\theta^*}^{1} \frac{(\gamma\theta)^2}{c}(1-2b) f(\theta)\, d\theta.$$

Combining with (25) gives

$$-\frac{b\theta^{*2}\gamma^2}{c} \cdot (1 - F(\theta^*)) = \int_{\theta^*}^{1} \frac{(\gamma\theta)^2}{c}(1-2b) f(\theta)\, d\theta.$$

Rearranging and simplifying gives

$$b = \left(2 - \frac{\theta^{*2}}{E[\theta^2 \mid \theta > \theta^*]}\right)^{-1}.$$

Inserting (23) into the first-order condition for s gives

$$E[\pi \mid \theta^*] f(\theta^*) \frac{-\partial \theta^*}{\partial s} = 1 - F(\theta^*).$$

Combining with (22), (24), and simplifying, yields

$$s = \left[\theta^*(1-b) - \frac{1 - F(\theta^*)}{f(\theta^*)} b\right] \frac{b\gamma^2\theta^*}{c} - (k - m\theta^*). \quad \square$$

Proof of Proposition 3. Noting that $\theta^* = \sqrt{2c(\bar{u} - s)}/(b\gamma)$, so that

$$s = \bar{u} - \frac{(\theta^*)^2 b^2 \gamma^2}{2c},$$

we can eliminate s from the problem and restate the problem as one of choosing b and θ^*. The first-order condition for θ^* is then

$$0 = -E[\pi \mid \theta^*] + \int_{\theta^*}^{1} \frac{\partial E[\pi \mid \theta]}{\partial \theta^*} f(\theta)\, d\theta. \tag{26}$$

But

$$E[\pi \mid \theta] = \frac{\gamma^2\theta^2 b(1-b)}{c} - \bar{u} + \frac{(\theta^*)^2 b^2\gamma^2}{2c} - k + m\theta,$$

so that

$$\frac{\partial E[\pi \mid \theta]}{\partial \theta^*} = \frac{b^2\gamma^2\theta^*}{c} > 0 \quad \forall\, \theta^* \in (0,1].$$

It follows that $\int_{\theta^*}^1 (\partial E[\pi \mid \theta]/\partial\theta^*) f(\theta)\, d\theta > 0$, so that (by (26)) $E[\pi \mid \theta^*] > 0$.

We will now see that this implies that $\theta^* > \theta^{FB}$. Denoting the manager's surplus by $E[MS \mid \theta]$, we note that by definition, $E[\pi \mid \theta] + E[MS \mid \theta] = E[TS \mid \theta]$: the sum of the firm's and the manager's surpluses equals total surplus for any θ. Evaluating this at $\theta = \theta^*$ and noting that by definition, $E[MS \mid \theta^*] = 0$, we obtain $E[TS \mid \theta^*] = E[\pi \mid \theta^*] > 0$. Furthermore, noting that the total surplus actually achieved cannot exceed its first-best level (obtained by setting $b = 1$), we must conclude that $E[TS^* \mid \theta^*] > 0 = E[TS^* \mid \theta^{FB}]$, where TS^* is the total surplus with first-best effort level. Since $E[TS^* \mid \theta] = \gamma^2\theta^2/(2c) - \bar{u} - k + m\theta$ is strictly increasing in θ for all $\theta > 0$, this proves that $\theta^* > \theta^{FB}$. □

General Notes for Comparative Statics. Throughout the paper, we examine comparative statics problems where we are interested in the dependence of two choice variables, x and y ($x = b$ and $y = s$ or $y = \theta^*$), on some parameter α. The choice variables are determined by the problem

$$\max_{x,y} \Pi(x, y; \alpha).$$

The first-order conditions are

$$0 = \Pi_x(x, y; \alpha), \tag{27}$$

$$0 = \Pi_y(x, y; \alpha). \tag{28}$$

We will drop the arguments from now on.

The second-order conditions are

$$\Pi_{xx} < 0, \tag{29}$$

$$\Pi_{yy} < 0, \tag{30}$$

$$\Pi_{xx}\Pi_{yy} - (\Pi_{xy})^2 > 0. \tag{31}$$

Fully differentiating the first-order conditions with respect to α gives

$$0 = \Pi_{x\alpha} + \Pi_{xx}\frac{\partial x}{\partial \alpha} + \Pi_{xy}\frac{\partial y}{\partial \alpha},$$

$$0 = \Pi_{y\alpha} + \Pi_{yx}\frac{\partial x}{\partial \alpha} + \Pi_{yy}\frac{\partial y}{\partial \alpha}.$$

This can be seen as a system of two linear equations in two unknowns, $\partial x/\partial\alpha$ and $\partial y/\partial\alpha$. Restating this as

$$\Pi_{xx}\frac{\partial x}{\partial \alpha} + \Pi_{xy}\frac{\partial y}{\partial \alpha} = -\Pi_{x\alpha}, \tag{32}$$

$$\Pi_{yx}\frac{\partial x}{\partial \alpha} + \Pi_{yy}\frac{\partial y}{\partial \alpha} = -\Pi_{y\alpha}, \tag{33}$$

we can write the system determinants as follows:

$$\Delta = \begin{vmatrix} \Pi_{xx} & \Pi_{xy} \\ \Pi_{yx} & \Pi_{yy} \end{vmatrix}, \tag{34}$$

$$\Delta_x = \begin{vmatrix} -\Pi_{x\alpha} & \Pi_{xy} \\ -\Pi_{y\alpha} & \Pi_{yy} \end{vmatrix}, \tag{35}$$

$$\Delta_y = \begin{vmatrix} \Pi_{xx} & -\Pi_{x\alpha} \\ \Pi_{yx} & -\Pi_{y\alpha} \end{vmatrix}. \tag{36}$$

Applying Cramer's rule and noting that $\Delta > 0$ (by the second-order condition (31)), we have the following results for the signs of the derivatives:

$$\begin{aligned} \operatorname{sgn}\left(\frac{\partial x}{\partial \alpha}\right) &= \operatorname{sgn}\left(\frac{\Delta_x}{\Delta}\right) = \operatorname{sgn}(\Delta_x) \\ &= \operatorname{sgn}(\Pi_{xy}\Pi_{y\alpha} - \Pi_{yy}\Pi_{x\alpha}), \end{aligned} \tag{37}$$

$$\begin{aligned} \operatorname{sgn}\left(\frac{\partial y}{\partial \alpha}\right) &= \operatorname{sgn}\left(\frac{\Delta_y}{\Delta}\right) = \operatorname{sgn}(\Delta_y) \\ &= \operatorname{sgn}(\Pi_{xy}\Pi_{x\alpha} - \Pi_{xx}\Pi_{y\alpha}). \quad \square \end{aligned} \tag{38}$$

Proof of Corollary 1. By assumption, $f(\theta) = 1/(2a)$ for all $\theta \in [\mu - a, \mu + a) \subseteq [0, 1)$ and zero otherwise. Simple calculation of uniform densities shows that $E[\theta] = \mu$ and $V[\theta] = a^2/3$. Also by assumption, the parameter values are such that the solution is nontrivial, i.e., $\mu - a < \theta^* < \mu + a$ in equilibrium. Now,

$$\Pr(\theta > \theta^*) = \frac{\mu + a - \theta^*}{2a},$$

and $f(\theta \mid \theta > \theta^*) = f(\theta)/\Pr(\theta > \theta^*) = (\mu + a - \theta^*)^{-1}$, so that

$$\begin{aligned} E \equiv E[\theta^2 \mid \theta > \theta^*] &= \int_{\theta^*}^{\mu+a} \theta^2 f(\theta \mid \theta > \theta^*)\, d\theta \\ &= \frac{(\mu+a)^3 - (\theta^*)^3}{3(\mu + a - \theta^*)}. \end{aligned} \tag{39}$$

From Proposition 2, $b = (2 - (\theta^*)^2/E)^{-1}$. Combining with (39),

$$\begin{aligned} b^{-1} &= 2 - \frac{3(\mu + a - \theta^*)(\theta^*)^2}{(\mu+a)^3 - (\theta^*)^3} \\ &= 2 - \frac{3}{A^2 + A + 1}, \quad \text{where } A \equiv \frac{\mu + a}{\theta^*}. \end{aligned} \tag{40}$$

Note that since $0 > A > 1$, we have $1/2 < b < 1$.

A rise in the variance of θ corresponds to a rise in a. We therefore need to determine the comparative statics with respect to a. To this end, we will use Equations (37) and (38) from the appendix section "General Notes for Comparative Statics," using $x = b$ and $y = \theta^*$ (i.e., casting the problem as one of choosing b and θ^*).

Plugging in the uniform distribution we are investigating and simplifying, the first-order conditions become

$$\begin{aligned} 0 &= \Pi_{\theta^*} \\ &= \frac{1}{2a}\left[\frac{b^2\gamma^2\theta^*}{c}(\mu+a) - \frac{\gamma^2(\theta^*)^2}{2c}(b^2+2b) + \bar{u} + k - m\theta^*\right], \end{aligned} \tag{41}$$

$$\begin{aligned} 0 = \Pi_b &= \frac{\gamma^2}{6ac}(\mu + a - \theta^*) \\ &\quad \cdot [3b(\theta^*)^2 + (1-2b)((\mu+a)^2 + \theta^*(\mu+a) + (\theta^*)^2)]. \end{aligned} \tag{42}$$

Taking derivatives of these expressions and noting that $\Pi_{\theta^*} = 0$ (for $\Pi_{\theta^* a}$) and $\Pi_b = 0$ (for Π_{ba}), we obtain the elements of Equations (37) and (38):

$$\Pi_{bb} = \frac{\gamma^2}{6ac}(\mu + a - \theta^*)[(\theta^*)^2 - 2(\mu+a)(\mu + a + \theta^*)], \tag{43}$$

$$\Pi_{\theta^*\theta^*} = \frac{1}{2ac}[b^2\gamma^2(\mu+a) - \gamma^2\theta^*(b^2+2b) - mc], \tag{44}$$

$$\Pi_{b\theta^*} = \frac{\gamma^2\theta^*}{2ac}[b(2(\mu+a)-\theta^*)-\theta^*], \tag{45}$$

$$\Pi_{ba} = \frac{\gamma^2}{6ac}(\mu+a-\theta^*)(1-2b)(2(\mu+a)+\theta^*), \tag{46}$$

$$\Pi_{\theta^* a} = \frac{b^2\gamma^2\theta^*}{2ac}. \tag{47}$$

We can also sign all these elements: $\Pi_{bb} < 0$ and $\Pi_{\theta^*\theta^*} < 0$ (by the second-order conditions for maximum); $\Pi_{b\theta^*} > 0$ (obtained by plugging in the value of b from (40) and analyzing the resulting expression); $\Pi_{ba} < 0$ (because $\theta^* < \mu + a$ by assumption and $b > 1/2$ as observed from (40)); $\Pi_{\theta^* a} > 0$ (obvious).

Now,

$$\begin{aligned}
&\Pi_{b\theta^*}\Pi_{ba} - \Pi_{bb}\Pi_{\theta^* a} \\
&\quad= \frac{\gamma^4\theta^*}{12a^2c^2}(\mu+a-\theta^*) \\
&\qquad\times\{[b(2(\mu+a)-\theta^*)-\theta^*](1-2b)(2(\mu+a)+\theta^*) \\
&\qquad- b^2[(\theta^*)^2 - 2(\mu+a)(\mu+a+\theta^*)]\},
\end{aligned}$$

so that

$$\begin{aligned}
&\operatorname{sgn}(\Pi_{b\theta^*}\Pi_{ba} - \Pi_{bb}\Pi_{\theta^* a}) \\
&\quad= \operatorname{sgn}([b(2(\mu+a)-\theta^*)-\theta^*](1-2b)(2(\mu+a)+\theta^*) \\
&\qquad- b^2[(\theta^*)^2 - 2(\mu+a)(\mu+a+\theta^*)]).
\end{aligned}$$

Substituting in $\mu + a = A\theta^*$ and $b = [2-(3/(A^2+A+1))]^{-1}$ from (40), the argument of the sgn operator above becomes $(\theta^*)^2(2A^6+6A^5-3A^4+8A^3+12A^2+3A-1)/(1-2A(1+A))^2$. Thus, by (38) with $x = b$ and $y = \theta^*$,

$$\begin{aligned}
\operatorname{sgn}\left(\frac{\partial\theta^*}{\partial a}\right) &= \operatorname{sgn}(\Pi_{b\theta^*}\Pi_{ba} - \Pi_{bb}\Pi_{\theta^* a}) \\
&= \operatorname{sgn}(2A^6+6A^5-3A^4+8A^3+12A^2+3A-1) \\
&= \text{“}+\text{”},
\end{aligned}$$

where the conclusion follows because $A = (\mu+a)/\theta^* > 1$ by definition, and the sixth-order polynomial above is positive for all $A > 1$.

Now, we have only to show that b is decreasing in a. By (37) with $x = b$ and $y = \theta^*$,

$$\operatorname{sgn}\left(\frac{\partial b}{\partial a}\right) = \operatorname{sgn}(\Pi_{b\theta^*}\Pi_{\theta^* a} - \Pi_{\theta^*\theta^*}\Pi_{ba}).$$

Since $\Pi_{ba} < 0$ and a, $m > 0$, we know that $\Pi_{b\theta^*}\Pi_{\theta^* a} - \Pi_{\theta^*\theta^*}\Pi_{ba} < \Pi_{b\theta^*}\Pi_{\theta^* a} - (\Pi_{\theta^*\theta^*} + m/(2a))\Pi_{ba}$. Therefore, to show that $\partial b/\partial a < 0$, it is sufficient to show that $\Pi_{b\theta^*}\Pi_{\theta^* a} - (\Pi_{\theta^*\theta^*} + m/(2a))\Pi_{ba} < 0$. Now,

$$\begin{aligned}
&\Pi_{b\theta^*}\Pi_{\theta^* a} - (\Pi_{\theta^*\theta^*} + m/(2a))\Pi_{ba} \\
&\quad= \frac{b\gamma^4}{4(ac)^2}\{b(\theta^*)^2[b(2(\mu+a)-\theta^*)-\theta^*] \\
&\qquad- \tfrac{1}{3}[b(\mu+a-\theta^*)-2\theta^*](\mu+a-\theta^*)(1-2b)(2(\mu+a)+\theta^*)\},
\end{aligned}$$

which equals $(4A^5 - 8A^4 - 5A^3 + 8A^2 + 2A - 1)(\theta^*)^3/(1-2A(1+A))^2$. This is negative whenever $A \in (1, \bar{A})$, where $\bar{A} \approx 2.1$ is the unique root of $2A^3 - 3A^2 - 3A - 1$ on $(1, 1)$.

To complete the proof, we must show that, in equilibrium, we always have $A \le \bar{A}$. In fact, we can show that $\bar{A}$ is precisely equal to the highest value that A can attain in equilibrium. First note that the equilibrium A is decreasing in $\bar{u}$ and k by Corollary 2. It also follows from the proof of the corollary that A is also decreasing in the full hiring cost $k - m\theta$. Thus, the highest possible equilibrium A is attained when $\bar{u} = k = m = 0$. Inserting these parameter values in the first-order conditions, we can easily verify that this highest equilibrium A is precisely $\bar{A}$. □

Proof of Corollary 2. We will use Equations (37) and (38) with $x = b$, $y = \theta^*$ and $\alpha = k$ or $\alpha = \bar{u}$. Observe that

$$\Pi_{b\alpha} = 0, \tag{48}$$

$$\Pi_{\theta^*\alpha} = f(\theta^*) > 0. \tag{49}$$

Now, by (38),

$$\operatorname{sgn}\left(\frac{\partial\theta^*}{\partial\alpha}\right) = \operatorname{sgn}(\Pi_{b\theta^*}\Pi_{b\alpha} - \Pi_{bb}\Pi_{\theta^*\alpha}) = -\Pi_{bb}f(\theta^*) > 0,$$

because $\Pi_{bb} < 0$ by the second-order condition for maximum.

Thus, θ^* is an increasing function of $\bar{u}$ and k, and hence, $A = (\mu + a)/\theta^*$ is a decreasing function of $\bar{u}$ and k. But by (40), $b = [2 - 3/(A^2+A+1)]^{-1}$, which for all $A > 1$ is a decreasing function of the single variable A, making b an *increasing* function of $\bar{u}$ and k. □

Proof of Proposition 4. As before, the firm chooses salary s and bonus b to solve

$$\max_{s,\,b}\int_{\theta^*(s,\,b)}^{1} E[\pi \mid \theta]f(\theta)\,d\theta.$$

Note that when holding b and s fixed, the presence of risk aversion does not change the manager's effort choice and the firm's expected profit for a given value of θ. Similarly, the firm's objective function is also the same as with a risk-neutral agent, except for the change to $\theta^*(s, b)$. Consequently, the equilibrium conditions given by Equations (21)–(23) in the proof of Proposition 2 still hold.

Next, recall from the manager's problem that in the presence of risk aversion,

$$\theta^* = \frac{\sqrt{2c(\bar{u}-s)+crb^2\sigma^2}}{b\gamma}.$$

The partial derivatives are therefore

$$\frac{\partial\theta^*}{\partial s} = \frac{-c}{(b\gamma)^2\theta^*} \quad\text{and}\quad \frac{\partial\theta^*}{\partial b} = \frac{rc\sigma^2}{b\gamma^2\theta^*} - \frac{\theta^*}{b}. \tag{50}$$

Substituting (50) and (23) into the equilibrium condition (21) and collecting the terms yields

$$b = \left(2 + \frac{rc\sigma^2}{\gamma^2 E} - \frac{\theta^*}{E}\right)^{-1},$$

where $E = E[\theta^2 \mid \theta > \theta^*]$.

Noting that the expression for the optimal s in the proof of Proposition 2 uses only the firm's first-order condition with respect to s and Equations (22) and (23), neither of which is affected by risk aversion, we obtain the same expression as we did with risk-neutral agents:

$$s=\left[\theta^*(1-b)-\frac{1-F(\theta^*)}{f(\theta^*)}b\right]\frac{b\theta^*\gamma^2}{c}-(k-m\theta).\quad\square$$

Proof of Proposition 5. The proof is almost identical to that of Proposition 3.

We first eliminate s from the problem and restate the problem as one of choosing b and θ^*. The first-order condition for θ^* is of the same form as with a risk-neutral agent, namely,

$$0=-E[\pi\mid\theta^*]+\int_{\theta^*}^1\frac{\partial E[\pi\mid\theta]}{\partial\theta^*}f(\theta)\,d\theta. \tag{51}$$

But

$$E[\pi\mid\theta]=\frac{\gamma^2\theta^2b(1-b)}{c}-\bar{u}+\frac{(\theta^*)^2b^2\gamma^2}{2c}-\frac{b^2r\sigma^2}{2}-k+m\theta,$$

so that

$$\frac{\partial E[\pi\mid\theta]}{\partial\theta^*}=\frac{b^2\gamma^2\theta^*}{c}>0\quad\forall\,\theta^*\in(0,1],$$

as with a risk-neutral agent. It follows that $\int_{\theta^*}^1(\partial E[\pi\mid\theta]/\partial\theta^*)\cdot f(\theta)\,d\theta>0$, so that (by (51)) $E[\pi\mid\theta^*]>0$.

We will now see that this implies that $\theta^*>\theta^{FB}$. Denoting the manager's surplus by $E[MS\mid\theta]$, we note that by definition, $E[\pi\mid\theta]+E[MS\mid\theta]=E[TS\mid\theta]$: the sum of the firm's and the manager's surpluses equals total surplus for any θ. Evaluating this at $\theta=\theta^*$ and noting that by definition, $E[MS\mid\theta^*]=rb^2\sigma^2/2>0$, we obtain $E[TS\mid\theta^*]>E[\pi\mid\theta^*]>0$. Furthermore, noting that the total surplus actually achieved cannot exceed its first-best level (obtained by setting $b=1$), we must conclude that $E[TS^*\mid\theta^*]>0=E[TS^*\mid\theta^{FB}]$, where TS^* is the total surplus with first-best effort level. Since $E[TS^*\mid\theta]=\gamma^2\theta^2/(2c)-\bar{u}-k+m\theta$ is strictly increasing in θ for all $\theta>0$, this proves that $\theta^*>\theta^{FB}$. $\square$

Proof of Corollary 3. First note that the manager's problem in the canonical model is the same as in the full model, so the manager's incentive compatibility (IC) condition is still $e=b\gamma\theta/c$, and the manager's individual rationality (IR) condition is

$$\bar{u}\le CE(\theta)=s+b\gamma\theta e-\frac{r}{2}b^2\sigma^2-\frac{1}{2}ce^2=s+\frac{b^2}{2}\left(\frac{\gamma^2\theta^2}{c}-r\sigma^2\right),$$

where the last equality follows after substituting in the IC condition. The firm's objective is to maximize $(1-b)(\gamma\theta)e-s-(k-m\theta)$ subject to the IR and and IC conditions. Substituting the IR and IC conditions into the firm's objective function yields the objective function

$$\max_b\left\{\frac{\gamma^2\theta^2}{c}\left(b-\frac{b^2}{2}\right)-\frac{r\sigma^2}{2}b^2-\bar{u}-k+m\theta\right\}.$$

Taking the derivative with respect to b and simplifying yields the optimal PPS:

$$b_{NS}=\left(1+\frac{cr\sigma^2}{\gamma^2\theta^2}\right)^{-1}.$$

Now, $b>b_{NS}$ if and only if $b^{-1}<b_{NS}^{-1}$. Substituting in b_{NS} from the line above and b from Proposition 4 turns the condition into

$$\frac{cr\sigma^2}{\gamma^2E}+2-\frac{(\theta^*)^2}{E}<1+\frac{cr\sigma^2}{\gamma^2\theta^2},$$

where $E=E[\tilde{\theta}^2\mid\tilde{\theta}>\theta^*]$. Note that $0<(\theta^*)^2<E<1$. Rearranging and simplifying gives

$$E-(\theta^*)^2<\frac{cr\sigma^2}{\gamma^2}\left(\frac{E}{\theta^2}-1\right).$$

This immediately proves the first part of the corollary: as θ approaches its upper limit of 1, the right-hand side becomes negative, while the left-hand side, which does not depend on θ, remains positive, violating the inequality above. Hence, $b>b_{NS}$ is not possible.

Further observe that the right-hand side is decreasing in θ, while the left-hand side is independent of it. Thus, the inequality holds for some θ if and only if it holds for the lowest type hired, namely, $\theta=\theta^*$. Substituting this value into the inequality and simplifying turns the condition into

$$1<\frac{cr\sigma^2}{\gamma^2(\theta^*)^2}.$$

Hence $b>b_{NS}$ for some manager types that are actually hired if and only if $r\sigma^2>\gamma^2(\theta^*)^2/c$.

Now, note that by Proposition 5 and the assumption of nontriviality of the problem we have $\theta^{FB}<\theta^*<1$. Hence, a sufficient condition for $r\sigma^2>\gamma^2(\theta^*)^2/c$ to hold is $r\sigma^2>\gamma^2/c\equiv\bar{R}$, whereas a sufficient condition for $r\sigma^2>\gamma^2(\theta^*)^2/c$ *not* to hold is $r\sigma^2<\theta^{FB}\gamma^2/c\equiv\underline{R}$. Note that both of these conditions are exogenous conditions on $r\sigma^2$, as neither $\bar{R}$ nor $\underline{R}$ depends on r or σ^2. This completes the proof of the second part of the corollary. (Finally, also note that both $\bar{R}=\gamma^2/c$ and $\underline{R}=\sqrt{m^2+2\gamma^2(k+\bar{u})/c^2}-m$ are increasing in γ and decreasing in c.) $\square$

Proof of Corollary 4. We need to determine the comparative statics with respect to r and σ^2. However, noting that these two variables enter the firm's profit function only as the product $r\sigma^2$, one can analyze only the effects of increasing r; the effects of increasing σ^2 will be identical. We will use Equations (37) and (38) from the appendix section "General Notes for Comparative Statics," using $x=b$, $y=\theta^*$, and $\alpha=r$.

Plugging in the uniform distribution we are investigating and simplifying, the first-order conditions become

$$0=\Pi_{\theta^*}=\frac{1}{2c}[b^2cr\sigma^2-b\gamma^2\theta^*(b(\theta^*-2)+2\theta^*)+2c(k+\bar{u}-m\theta^*)], \tag{52}$$

$$0=\Pi_b=\frac{\theta^*-1}{3c}[3bcr\sigma^2-\gamma^2(1+\theta^*+(\theta^*)^2-2b+b(\theta^*-2)\theta^*)]. \tag{53}$$

Taking derivatives of these expressions, we obtain the elements of Equations (37) and (38):

$$\Pi_{bb}=\frac{\theta^*-1}{3c}(3cr\sigma^2+\gamma^2(2+2\theta^*-(\theta^*)^2)), \tag{54}$$

$$\Pi_{\theta^*\theta^*}=\frac{1}{c}(b\gamma^2(b-(2+b)\theta^*)-cm), \tag{55}$$

$$\Pi_{b\theta^*} = br\sigma^2 - \frac{\gamma^2\theta^*}{c}(b(\theta^*-2)+\theta^*), \tag{56}$$

$$\Pi_{br} = b\sigma^2(\theta^*-1), \tag{57}$$

$$\Pi_{\theta^* r} = \frac{b^2\sigma^2}{2}. \tag{58}$$

We can also sign most of these elements: $\Pi_{bb} < 0$ and $\Pi_{\theta^*\theta^*} < 0$ (by the second-order conditions for maximum); $\Pi_{br} < 0$ (because $\theta^* < 1$ by assumption); $\Pi_{\theta^* r} > 0$ (obvious).

Because $2c(k + \bar{u} - m\theta^*) > 0$, (52) implies that $b^2cr\sigma^2 - b\gamma^2\theta^*(b(\theta^*-2)+2\theta^*) < 0$, which is equivalent to

$$\gamma^2[3\theta^*(2\theta^* + b(\theta^*-2))] - 3bcr\sigma^2 > 0. \tag{59}$$

Now,

$$\Pi_{b\theta^*}\Pi_{br} - \Pi_{bb}\Pi_{\theta^* r} = \frac{b\sigma^2(\theta^*-1)}{6c}\{3bcr\sigma^2 - \gamma^2[6(\theta^*)^2 + b(2-10\theta^*+5(\theta^*)^2)]\}, \tag{60}$$

so that

$$\operatorname{sgn}\left(\frac{\partial\theta^*}{\partial r}\right) \overset{(38)}{=} \operatorname{sgn}(\Pi_{b\theta^*}\Pi_{br} - \Pi_{bb}\Pi_{\theta^* r}) \overset{(60)}{=} \operatorname{sgn}(\gamma^2[6(\theta^*)^2 + b(2-10\theta^*+5(\theta^*)^2)] - 3bcr\sigma^2).$$

But

$$\{6(\theta^*)^2 + b(2-10\theta^*+5(\theta^*)^2)\} - \{3\theta^*(2\theta^* + b(\theta^*-2))\} = 2b(1-\theta^*)^2 > 0, \tag{61}$$

so that

$$\gamma^2[6(\theta^*)^2 + b(2-10\theta^*+5(\theta^*)^2)] - 3bcr\sigma^2 \overset{(61)}{>} \gamma^2[3\theta^*(2\theta^*+b(\theta^*-2))] - 3bcr\sigma^2 \overset{(59)}{>} 0.$$

Thus, $\operatorname{sgn}(\partial\theta^*/\partial r) > 0$, as claimed in the corollary.

Combining (53) and (59), we see that

$$1 + \theta^* + (\theta^*)^2 - 2b + b(\theta^*-2)\theta^* = \frac{3bcr\sigma^2}{\gamma^2} < 3\theta^*(2\theta^* + b(\theta^*-2)),$$

which simplifies to

$$(\theta^*)^2(5+2b) - \theta^*(1+4b) - 1 + 2b > 0. \tag{62}$$

Next, observe that by Proposition 4, $b^{-1} > 2 - (\theta^*)^2/E[\theta^2 \mid \theta > \theta^*]$. Given a uniform distribution on [0, 1], the right-hand side of this inequality equals $2 - 3/(A^2+A+1)$, where $A = 1/\theta^*$, as shown in (40). Thus $b^{-1} > 2 - 3/(A^2+A+1)$, which implies

$$1 - 2b > -(1+b)(\theta^*)^2 - (1-2b)\theta^*. \tag{63}$$

Combining this with (62) yields

$$1 + 4b < (\theta^*)^2(5+2b) - \theta^*(5-b). \tag{64}$$

By (37), $\operatorname{sgn}(\partial b/\partial r) = \operatorname{sgn}(\Pi_{b\theta^*}\Pi_{\theta^* r} - \Pi_{\theta*\theta^*}\Pi_{br})$. Because

$$\Pi_{b\theta^*}\Pi_{\theta^* r} - \Pi_{\theta^*\theta^*}\Pi_{br} = \frac{b\sigma^2}{2c}\{b^2cr\sigma^2 + b\gamma^2[(3+b)(\theta^*)^2 - 2(2+b)\theta^* + 2b] - 2cm(1-\theta^*)\}$$

and $2cm(1-\theta^*)$,

$$\operatorname{sgn}(\Pi_{b\theta^*}\Pi_{\theta^* r} - \Pi_{\theta^*\theta^*}\Pi_{br}) < \operatorname{sgn}(b^2cr\sigma^2 + b\gamma^2[(3+b)(\theta^*)^2 - 2(2+b)\theta^* + 2b]).$$

It follows that a necessary condition for $\partial b/\partial r \geq 0$ is that

$$bcr\sigma^2 + \gamma^2[(3+b)(\theta^*)^2 - 2(2+b)\theta^* + 2b] \geq 0.$$

Combining this with (53) changes the necessary condition to

$$(\theta^*)^2(10+4b) - \theta^*(11+8b) + (1+4b) < 0. \tag{65}$$

Note that (64) and (65) form a system of inequalities in two variables, b and θ^*. It is easy to verify numerically that this system has no solution where $b > 0$ and $0 < \theta^* < 1$. This proves that $\partial b/\partial r \geq 0$ is not possible, which completes the proof of the corollary. □

References

Armstrong C, Larcker D, Su C-L (2010) Endogenous selection and moral hazard in executive compensation contracts. Working paper, Rock Center for Corporate Governance, Stanford University, Stanford, CA.

Arya A, Mittendorf B (2005) Offering stock options to gauge managerial talent. *J. Accounting Econom.* 40(1–3):189–210.

Baiman S (1991) Agency research in managerial accounting: A second look. *Accounting, Organ. Soc.* 15(4):341–371.

Baker GP, Jorgensen B (2003) Volatility, noise, and incentives. Working paper, Harvard Business School, Boston.

Bernardo AE, Cai H, Luo J (2001) Capital budgeting and compensation with asymmetric information and moral hazard. *J. Financial Econom.* 61(3):311–344.

Bushman R, Indjejikian R, Penno M (2000) Private predecision information, performance measure congruity, and the value of delegation. *Contemporary Accounting Res.* 17(4):562–587.

Darrough MN, Stoughton NM (1986) Moral hazard and adverse selection: The question of financial structure. *J. Finance* 41(2): 501–513.

Dutta S (2008) Managerial expertise, private information, and pay-performance sensitivity. *Management Sci.* 54(3):429–442.

Edmans A, Gabaix X, Landier A (2009) A multiplicative model of optimal ceo incentives in market equilibrium. *Rev. Financial Stud.* 22(12):4881–4917.

Eggleston K, Posner EA, Zeckhauser R (2000) The design and interpretation of contracts: Why complexity matters. *Northwestern University Law Rev.* 95(1):91–132.

Gabaix X, Landier A (2008) Why has CEO pay increased so much? *Quart. J. Econom.* 123(1):49–100.

Hagerty KM, Siegal DR (1988) On the observational equivalence of managerial contracts under conditions of moral hazard and self-selection. *Quart. J. Econom.* 103(2):425–428.

Hart O, Holmstrom B (1987) The theory of contracts. Bewley TF, ed. *Advances in Economic Theory, 5th World Congress of the Econometric Society* (Cambridge University Press, New York), 71–155.

Haubrich JG (1994) Risk aversion, performance pay, and the principal-agent problem. *J. Political Econom.* 102(2):258–276.

Himmelberg CP, Hubbard RG (2000) Incentive pay and the market for CEOs: An analysis of pay for performance sensitivity. Working paper, Columbia University, New York.

Jensen M, Murphy K (1990) Performance pay and top management incentives. *J. Political Econom.* 98(2):225–264.

Joskow PL (1987) Contract duration and relationship-specific investments: Empirical evidence from coal markets. *Amer. Econom. Rev.* 77(1):168–185.

Jullien B, Salani B, Salani F (1999) Should more risk-averse agents exert more effort? *The Geneva Papers on Risk and Insurance Theory* 24:19–28.

La Porta R, de Silanes FL, Shleifer A, Vishny RW (1997) Legal determinants of external finance. *J. Finance* 52(3):49–100.

La Porta R, de Silanes FL, Shleifer A, Vishny RW (1998) Law and finance. *J. Political Econom.* 106(6):1113–1155.

Liang P, Rajan M, Ray K (2008) Optimal team size and monitoring in organizations. *Accounting Rev.* 83(3):789–822.

Melumad N, Mookherjee D, Reichelstein S (1997) Contract complexity, incentives, and the value of delegation. *J. Econom. Management Strategy* 6(2):257–289.

Morck R, Shleifer A, Vishny R (1988) Management ownership and market valuation. *J. Financial Econom.* 20:293–315.

Murphy KJ (1999) Executive compensation. Ashenfelter O, Card D, eds. *Handbook of Labor Economics* (Elsevier Science, Amsterdam), 2485–2563.

Prendergast C (2002) The tenuous trade-off between risk and incentives. *J. Political Econom.* 110(5):1071–1102.

Ray K (2007a) Performance evaluations and efficient sorting. *J. Accounting Res.* 45(4):8319–882.

Ray K (2007b) The retention effect of withholding performance information. *The Accounting Rev.* 82(2):389–425.

Rosen S (1981) The economics of superstars. *Amer. Econom. Rev.* 71(5):845–858.

Sung J (2005) Optimal contracts under adverse selection and moral hazard: A continuous-time approach. *Rev. Financial Stud.* 18(3):1021–1073.