Firm Performance Pay as Insurance against Promotion Risk *

Alvin Chen†

Link to current version

November 7, 2020

Abstract

The prevalence of pay based on risky firm outcomes for non-executive workers presents a puzzling departure from conventional contract theory, which predicts insurance provision by the firm. I revisit this puzzle in a framework with workers who prefer early resolution of uncertainty. When workers at the same firm compete against each other for promotions, the optimal contract features pay based on firm outcomes as insurance against unfavorable promotion prospects. The model's predictions are consistent with observed phenomena such as performance-based vesting, option-like payoffs, and overvaluation of equity pay by non-executive workers. It also generates novel predictions linking organizational structure to firm performance pay.

Keywords: Insurance, firm performance pay, year-end bonuses, stock option pay, tournament, optimal contracting, early resolution of uncertainty, Epstein-Zin

JEL Classification: D81 (Criteria for Decision-Making under Risk and Uncertainty), D86 (Economics of Contract: Theory), G32 (Corporate Finance and Governance).

*I am deeply indebted to my advisor Philip Bond for his mentorship and support throughout the Ph.D. program. I am grateful for the guidance of Edward Rice, Mark Westerfield, and Yao Zeng on my dissertation committee. I appreciate Jarrad Harford, Fahad Khalil, and James Morrow for serving on my supervisory committee. I thank Vincent Maurin, Marcus Opp, Jan Starmans, and conference participants at the Cambridge Corporate Finance Symposium and Corporate Finance Day at HEC Liege as well as seminar participants at the University of Washington, The Wharton School, York University, the University of Alberta, BI Olso, the University of Amsterdam, the Stockholm School of Economics, and Warwick Business School for their insightful comments. All mistakes are mine.

†Stockholm School of Economics / Swedish House of Finance. Email: alvin.chen at hhs.se.

The moral hazard literature highlights a tendency for firms to insure risk-averse workers against uncertain firm outcomes that are beyond their individual control. Yet, non-executive workers in the United States routinely receive performance pay based on uncertain firm outcomes (Kruse et al. 2008). As Bergman and Jenter (2007, p. 668) point out, the prevalence of such performance pay for non-executive workers "is a puzzle for standard economic theory: any positive incentive effects should be diminished by free-rider problems and overshadowed by the cost of imposing risk on employees." Adding to this puzzle, many empirical findings suggest that cash constraints cannot fully explain the use of such performance pay.[1] This paper argues that when workers compete against each other for promotions, performance pay based on firm outcomes insures against poor promotion prospects. In such a setting, performance pay based on firm outcomes is not indicative of inefficient risk sharing; rather, it *is* a form of insurance.

Promotion tournaments are common in the workplace.[2] When workers compete against each other for promotions, they face *promotion risk*—the possibility of being passed over for promotion despite having performed well. Better firm outcomes correspond to a larger pool of high-performing workers, and hence to more intense competition for promotion and worse promotion prospects for an individual worker. The optimal contract insures workers against promotion risk by providing them with *firm performance pay*—payments that increase with both individual performance and firm outcomes. This insurance motive results in an optimal contract with features consistent with common compensation practices such as performance-based vesting and option-like payoffs. The insurance role of firm performance pay helps explain phenomena such as the overvaluation of stock option pay by non-executive workers and generates novel predictions linking the organizational structure of the firm to the structure of pay for its non-executive workers.

[1] For example, Bergman and Jenter (2007) report that "firms with large amounts of cash and high cash flows grant more options, while firms with more need for cash to service debt grant fewer options." Oyer and Schaefer (2005) survey related empirical evidence and concludes that "the relation between financing constraints and option grants [for non-executives] is, at best, mixed."

[2] See Bognanno (2001), DeVaro (2006), Cichello et al. (2009) for empirical documentation of promotion tournaments in the workplace.

I study the insurance role of firm performance pay in a parsimonious two-period model featuring a firm and many workers. The model has two main ingredients: promotion risk and the worker's preference for early resolution of uncertainty. In the model, promotion risk stems from the firm's production technology. Production at the firm involves two types of tasks: lower level and higher level. In the first period, all workers engage in lower-level tasks. The individual output of a worker depends only on that worker's effort and type, which is ex-ante unknown to all. A worker's output in the first period reveals that worker's type to be either high or low. In the second period, production at the firm involves higher-level tasks that require high-type workers. Because higher-level tasks are more difficult to monitor, incentive compatibility results in higher pay for the workers promoted to positions that involve such tasks. The firm's production technology, which limits the number of higher-level tasks, results in a promotion tournament that exposes workers to promotion risk.[3]

The key economic force underlying the insurance role of firm performance pay is the worker's preference for early resolution of uncertainty. The worker's anxiety upon learning about tomorrow's unfavorable promotion prospects results in an insurance benefit of additional pay today. Absent this anxiety, promotion risk realizes only after the promotion outcome in the second period, and firm performance pay in the first period would not provide any insurance benefits. I model the worker's preferences with Epstein-Zin utility, which is routinely used in the asset-pricing and macro-finance literature to capture how people respond to news about the future. In particular, the usual parameterization of this utility function corresponds to a preference for early resolution of uncertainty, which provides the crucial link between future promotion prospects and current compensation that allows firm performance pay to play an insurance role in this framework.

Note that in this model, the realization of a poor firm outcome does not affect the

[3] There are other rationales, aside from the production technology, for limiting the number of promotions, which results in a promotion tournament. For example, Levin and Tadelis (2005) illustrate how a firm can signal its quality by limiting its output. As long as workers face promotion risk, firm performance pay can play an insurance role.

ability of the firm to meet its obligations as a going concern. This paper does not dispute that some cash-constrained firms, such as start-ups, may offer firm performance pay to their workers because they simply cannot commit to a fixed wage. Rather, I aim to illustrate why a financially unconstrained firm may want to offer firm performance pay to its workers as a form of insurance against promotion risk.

My framework sheds light on three puzzling empirical findings. First, more risk-averse workers are more likely to receive stock option pay and profit-sharing plans as part of their compensation (Kruse et al. 2008). Second, firms with more volatile stock returns are more likely to grant stock option pay to non-executive workers (Spalt 2013). Third, non-executive workers tend to overvalue stock option pay (see Oyer and Schaefer 2005, Hallock and Olson 2006, Hodge et al. 2009), while CEOs tend to undervalue it (see Hall and Murphy 2002, Bettis et al. 2005).

The first two observations are surprising given the conventional view that risky pay results from an optimal trade-off between the cost and benefit of imposing risk on workers. This perspective implies that, all else equal, an increase in the cost of imposing risk should discourage the use of firm performance pay. In contrast, my framework highlights an insurance benefit of firm performance pay when workers face promotion risk. As a result, an increase in the worker's risk aversion leads to greater demand for this insurance. In my framework, increased uncertainty about the aggregate performance of workers at a firm makes the firm riskier and at the same time increases promotion risk for its workers. My analysis implies that some parts of a firm's risk and its use of stock option pay are positively linked via optimal contracting.

My framework also helps explain the over- and undervaluation dichotomy of the third observation. Non-executive workers receive an insurance benefit from performance pay based on firm outcomes, which increases their subjective valuation of stock option pay. In contrast, the CEO derives no insurance benefit from stock option pay because the CEO is already at the top of the hierarchy. The CEO simply values stock option pay from the perspective of an underdiversified investor, which leads to a subjective valuation below the fair market price.

Prior studies have proposed separate explanations for each of these findings. For instance, Spalt (2013) posits that non-executive workers have gambling preferences for skewed, lottery-like payoffs that make stock options in riskier firms more valuable. Hodge et al. (2009) argue that non-executive employees overvalue the stock options in their pay because of bounded rationality. My model provides unifying explanations for these findings within one framework with rational, risk-averse agents who prefer early resolution of uncertainty. In addition, the model generates novel testable implications that link a firm's organizational structure to its compensation policy.

Related Literature. This paper primarily relates to the literature that considers firm performance pay for non-executive workers. Studies in this area focus on identifying a sufficiently large benefit of firm performance pay to justify the cost of imposing risk on non-executive workers, who are presumably less diversified and more risk-averse than the firm. Some benefits include reduced financial constraints (Core and Guay 2001, Kim and Ouimet 2014, Sun and Xiaolan 2019), favorable accounting treatment (Hall and Murphy 2003), employee retention (Oyer 2004), employee sorting (Lazear 2004, Bergman and Jenter 2007, Vladimirov 2019), improved product market competitiveness (Bova and Yang 2017), and lower recapitalization costs (Efing et al. 2018). I contribute to this literature by reconsidering the cost in the cost-benefit trade-off; I highlight how firm performance pay can actually *reduce* a worker's payoff uncertainty in a setting with promotion risk.

The idea that workers prefer pay that co-varies with aggregate firm outcomes because it insures them against increased competition for promotions in the future is reminiscent of a series of papers on "keeping up with the Joneses" (DeMarzo et al. 2004, DeMarzo et al. 2007, and DeMarzo et al. 2008).[4] These studies argue that investors may prefer

[4] A related literature considers agents with innate preferences about aggregate wealth distribution such as inequity aversion (Fehr and Schmidt 1999), KUJ (DeMarzo and Kaniel 2017), and envy (Goel

payouts that co-vary with aggregate wealth levels because such payouts hedge against increased competition over consumption and investment goods in the future. They investigate the asset-pricing implications of investors' relative wealth concerns due to the competition for goods. In contrast, I study the optimal contracting implications of workers' relative performance concerns due to the competition for promotion.

This paper also bridges the broad bodies of literature on optimal contracting and tournaments. Previous studies involving contracting and tournaments tend to consider each in isolation. Most dynamic contracting frameworks with career concerns do not feature promotion risk induced by tournament incentives (e.g., Gibbons and Murphy 1992, Gibbons and Waldman 1999, Holmström 1999, Terviö 2009, Axelson and Bond 2015). As a consequence, the insurance role that firm performance pay plays in my framework is absent in theirs. The tournament literature usually focuses on how to set up such a contest (e.g., Lazear and Rosen 1981, O'Keeffe et al. 1984, Lazear 1989, Chen 2003, Boudreau et al. 2016). While some authors compare contracts to tournaments (e.g., Lazear and Rosen 1981, Green and Stokey 1983, Nalebuff and Stiglitz 1983), rarely do they consider how the two interact. This paper establishes a link between the firm's organizational hierarchy and its compensation policy. A separate contribution of this paper is to study contracting under Epstein-Zin preferences.[5] In my framework, firm performance pay plays an insurance role in the optimal contract only when workers prefer early resolution of uncertainty. This result suggests that further contracting studies with Epstein-Zin utility may uncover other economic forces overlooked by conventional frameworks with time-separable utility.

and Thakor 2010).

[5] See, for example, di Tella (2017) and di Tella (2019), which study optimal contracting between households and financial intermediaries in a continuous time setup with Epstein-Zin preferences.

1 Model Setup

The model has two periods (three dates: $t = 0$, $t = 1$, $t = 2$) and features a risk-neutral firm that employs a unit-mass continuum of risk-averse workers who prefer early resolution of uncertainty. Each worker, indexed by $a \in [0, 1]$, has a type parameter $\eta_a \in \{0, 1\}$, which is ex-ante unknown to all. The average quality of workers at the firm, θ, takes values $\theta_1, \theta_2, ..., \theta_n$ with probabilities $p_1, p_2, ..., p_n$, respectively. Conditional on θ, the workers' types are i.i.d with η_a taking a value of 1 with probability θ.

The workers have lifetime utility at time t given by the Epstein-Zin formulation:

$$U_t = \left[c_t^{1-\frac{1}{\psi}} + E_t[U_{t+1}^{1-\gamma}]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}}, \qquad (1)$$

where c_t is the worker's consumption at time t. The parameters $\gamma > 0$ ($\gamma \neq 1$) and $\psi > 1$ capture the worker's relative risk aversion and elasticity of intertemporal substitution, respectively. The worker's subjective discount rate is set to 1; its exact magnitude does not qualitatively alter the main results of the paper. The workers can neither borrow nor save; they consume wages in the period in which the payments were made.[6] The lifetime utility at $t = 3$ for all workers is fixed at $U_3 > 0$ and keeps the worker's lifetime utility well defined.

The use of Epstein-Zin utility to model worker preferences is the main technical innovation in this framework. Epstein-Zin utility is routinely used in asset-pricing and macro-finance models because it captures important properties of economic behavior, namely how people respond to news.[7] Epstein-Zin utility plays the same role in this model. Specifically, the common parameterization of Epstein-Zin utility with $\gamma > \frac{1}{\psi}$ cor-

[6] Lower-level employees typically want to borrow because expected future earnings exceed current earnings, as they do in this model; however, financial constraints due to moral hazard usually prevent this consumption smoothing.

[7] For example, Ai and Bansal (2018) argue that the macroeconomic announcement premium provides evidence of "a key aspect of investors' preferences not captured by the time-separable expected utility" makes a revealed preferences argument for non-expected utilities such as the non-expected utilities such as the recursive preferences of Kreps and Porteus (1978) and Epstein and Zin (1989).

Lower-level Tasks

	Effort ($e_1 = 1$)	No effort ($e_1 = 0$)
High-type ($\eta = 1$)	High output (x_H)	No output (0)
Low-type ($\eta = 0$)	Low output (x_L)	No output (0)

Table 1
Production Technology: Lower-Level Tasks

responds to a preference for early resolution of uncertainty, which realizes promotion risk as soon as the workers learn about unfavorable promotion prospects from firm outcomes.

Production at the firm involves two types of tasks: lower level and higher level. In the first period ($t = 0$ to $t = 1$), all workers engage in lower-level tasks. A worker engaging in a lower-level task generates an output $q_{1a} \in \{0, x_L, x_H\}$ that depends on that worker's type ($\eta_a \in \{0, 1\}$) and effort ($e_{1a} \in \{0, 1\}$). To emphasize the role of promotion risk, the model mutes other sources of uncertainty. First, the workers learn their own types shortly after employment prior to making the effort decision. This assumption is equivalent to the state-by-state incentive compatibility assumption of Edmans and Gabaix (2011). Second, given their effort decisions, workers face no uncertainty about their individual outputs. A high-type worker generates high output ($q_{1a} = x_H$) with effort and no output ($q_{1a} = 0$) without effort. Shirking by a high-type worker results in private benefits equivalent to an extra b_{1H} units of consumption at $t = 1$. A low-type worker generates low output ($q_{1a} = x_L$) with effort and no output ($q_{1a} = 0$) without effort. Shirking by a low-type worker results in private benefits equivalent to an extra b_{1L} units of consumption at $t = 1$. The production technology for lower-level tasks is summarized in Table 1. These assumptions allow the model to isolate the demand for insurance against promotion risk from the demand for insurance against other risks.[8] The paper's main results concerning the insurance role of firm performance pay are robust to these assumptions.

The firm outcome in the first period, e.g., total production, is a simple aggregation

[8] For example, Prendergast (1992) and Fairburn and Malcomson (2001) study worker demand for insurance against the uncertainty about their types.

of the individual outputs of all workers engaging in lower-level tasks:

$$Q_1 = \int_0^1 q_{1a} da. \tag{2}$$

The model aims to describe a setting in which the firm outcome provides little information about an individual worker's effort because it aggregates the output of many individual workers. For ease of exposition, the model captures this setting in the extreme by assuming a continuum of workers, which implies that the firm outcome provides no information about an individual worker's effort. Moreover, the firm's production technology implies that the firm outcome also provides no information about an individual worker's type over and above that contained in the worker's individual output. These features of the model help highlight the insurance role of firm performance pay and distinguish it from the informational role of Holmström (1979).

In the second period ($t = 1$ to $t = 2$), the firm can promote up to a fraction $K \in (0, \theta_n)$ of the workers from the first period to engage in higher-level tasks. A worker engaging in a higher-level task generates an output $q_{2a} \in \{0, y_H\}$ that depends on that worker's type and effort. A low-type worker always generates no output and cannot extract private benefits from higher-level tasks. A high-type worker generates high output ($q_{2a} = y_H$) with effort and no output ($q_{2a} = 0$) without effort. If a high-type worker does not exert effort in the higher-level task, that worker receives private benefits worth an additional $b_{2H} > b_{1H}$ units of consumption at $t = 2$. This assumption captures the idea that higher-level tasks are more difficult to monitor and results in higher pay for promoted workers due to incentive compatibility.[9] The production technology for higher-level tasks is summarized in Table 2.

The firm outcome in the second period, e.g., total production, is also a simple aggre-

[9] One can model the difficulty of monitoring higher-level tasks in other ways without changing the paper's main results. For example, the production technology can be modified so that shirking in both types of tasks results in the same private benefit but shirking in higher-level tasks is more difficult to detect.

Higher-level Tasks

	Effort ($e_2 = 1$)	No effort ($e_2 = 0$)
High-type ($\eta = 1$)	High output (y_H)	No output (0)
Low-type ($\eta = 0$)	No output (0)	No output (0)

Table 2
Production Technology: Higher-Level Tasks

gation of the individual outputs of all workers engaging in higher-level tasks:

$$Q_2 = \int_0^1 \phi_a q_{2a} da, \tag{3}$$

where ϕ_a takes a value of 1 if the a^{th} worker is promoted to a higher-level task and 0 otherwise. For simplicity, the workers who are not promoted exit the firm and take up their outside option.[10]

Promotion risk in this model stems from the firm's production technology. Generating high output while engaging in lower-level tasks reveals a worker to be high-type. Because a worker's type is persistent and higher-level tasks are NPV positive only when the firm assigns high-type workers to them, the firm only promotes workers with high output in lower-level tasks. Because high-output workers are identical, the model assumes that the firm promotes high-output workers with equal probability.[11] The limited number of higher-level tasks and the pay increase for workers assigned to them result in promotion risk, which motivates firm performance pay in the optimal contract.

As is standard in the moral hazard literature, I assume that parameters are such that inducing effort is optimal for the firm ($x_L > b_{1L}$, $x_H > b_{1H}$, and $y_H > b_{2H}$). Hence, the objective of the firm is to induce effort at lowest expected wage cost.

The main specification of the model assumes that the firm enters into one-period contracts with its workers at $t = 0$ and $t = 1$. The assumption is meant to capture the

[10] This paper's results do not change if one assumes that the nonpromoted workers remain employed in the firm, engaging in lower-level tasks.

[11] The Appendix micro-founds this assumption in a set-up where a worker's output in higher-level tasks also depends on a fit parameter that is i.i.d across high-output workers.

difficulty of entering into long-term employment arrangements. The paper's main results on firm performance pay extend to the case where the firm can commit to a two-period contract at $t = 0$.[12] A feasible contract at $t = 0$, hereafter "the first-period contract," specifies the payments to the worker at $t = 1$ given the worker's individual output and the firm outcome that period. The worker's outside option at $t = 0$ pays u_1 at $t = 1$. The model assumes that u_1 is sufficiently small ($u_1 \leq b_{1L}$, $u_1 \leq b_{1H}$) such that a first-period contract that satisfies the worker's incentive compatibility constraints also satisfies the worker's participation constraint. The promotion decision occurs after $t = 1$ wages are paid. This timing assumption allows the model to capture the idea that it takes time for the firm to determine the fit between a worker and a higher-level task without adding an additional period.[13]

A feasible wage contract at $t = 1$, hereafter "the second-period contract," specifies the payments to the promoted worker at $t = 2$ given the worker's individual output and the firm outcome that period. The worker's outside option at $t = 1$ pays u_2 at $t = 2$.[14] The model assumes that b_{2H} is sufficiently large ($b_{2H} > u_2$) such that a promotion to a higher-level task pays better than the outside option. Workers are protected by limited liability so payments to the worker cannot be negative. Table 3 summarizes the timing of the model.

2 Optimal Contracts

This section characterizes the optimal first- and second-period contracts. I proceed with backward induction, first solving for the optimal second-period contract, $W_2(q_2, Q_2)$, which specifies the payment to the promoted worker contingent on that worker's indi-

[12] See Section 5.2 for a discussion of this extension.

[13] The paper's results can also be derived in a model with three periods (four dates: $t = 0$, $t = 1$, $t = 2$, $t = 3$) without this timing assumption.

[14] The worker's outside option in the second period can be made contingent on the worker's output in the first period without changing the paper's results. Better outside options for high-output workers in the second period simply increase the implicit incentives in the first period.

$t = 0$	$t = 1$	$t = 2$
· Endowment: C_0	· Output: q_1, Q_1	· Output: q_2, Q_2
· Contract	· u_1 or $W_1(q_1, Q_1)$	· u_2 or $W_2(q_2, Q_2)$
· Worker learns type: η	· Promotion / contract	
· Effort decision: e_1	· Effort decision: e_2	

Table 3
Model Timing

vidual output and the firm outcome in the second period. I then solve for the optimal first-period contract, $W_1(q_1, Q_1)$, which specifies the payment to the worker contingent on that worker's individual output and the firm outcome in the first period, taking as given the promotion tournament and the optimal second-period contract. The subscript identifying the worker is suppressed for notational convenience.

2.1 Optimal Second-Period Contract

The optimal second-period contract yields additional rent to promoted workers in order to satisfy incentive compatibility in higher-level tasks. The firm outcome in the first period is a sufficient statistic for the mass of high-output workers, who are the only candidates for promotion. Given the firm outcome in the first period, there is no uncertainty about the firm outcome in the second period if all promoted workers exert effort. A promoted high-type worker who exerts effort generates high output with certainty. A promoted high-type worker who shirks generates no output with certainty, but receives an additional b_{2H} units of consumption at $t = 2$. The incentive compatibility constraint for higher-level tasks for a promoted high-type worker can be expressed as

$$\left[c_1^{1-\frac{1}{\psi}} + W_2(y_H, Q_2)^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}} \geq \left[c_1^{1-\frac{1}{\psi}} + (W_2(0, Q_2) + b_{2H})^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}. \quad (4)$$

Similarly, a promoted worker's participation constraint can be expressed as

$$\left[c_1^{1-\frac{1}{\psi}} + W_2(y_H, Q_2)^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}} \geq \left[c_1^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}. \tag{5}$$

Given the promotion outcomes, the firm's optimization problem in the second period can expressed as

$$\underset{\{W_2(q_2,Q_2)\geq 0\}}{\text{minimize}} \quad E_0[W_2(q_2, Q_2)]$$

$$\text{subject to} \quad (4) \text{ and } (5).$$

Standard contracting arguments imply that the optimal second-period contract pays the promoted worker nothing for no output and a bonus b_{2H} for high output regardless of the firm outcome. The production technology for higher-level tasks implies that the likelihood ratio for no output is zero. The optimal second-period contract minimizes the payment to the worker given no output. Limited liability implies that the minimum is 0. The production technology also implies that the likelihood ratio for high output is infinity. As such, the firm outcome in the second period does not provide incremental information about the worker's effort; therefore, the bonus does not depend on the firm outcome. The positive pay differential between the promotion and the worker's outside option, $b_{2H} - u_2$, serves as the prize for winning a promotion and is instrumental in generating promotion risk in this setting.

2.2 Incentive Compatibility in Lower-Level Tasks

The possibility of promotion to a higher-level task in the second period provides workers with implicit incentives in the first. Let $U_2(\phi)$ be the $t = 2$ lifetime utility of a high-type worker in equilibrium, with ϕ being the firm's promotion decision. A promoted worker ($\phi = 1$) who puts forth effort always generates high output ($q_2 = y_H$) in the higher-level task; in equilibrium, that worker receives a payment of b_{2H} at $t = 2$. The promoted worker's lifetime utility at $t = 2$ is

$$U_2(1) = \left[b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}. \tag{6}$$

A worker who is not promoted to a higher-level task takes up the outside option, which pays u_2 at $t = 2$:

$$U_2(0) = \left[u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}. \tag{7}$$

Let $\bar{U_2}(Q_1)$ be the certainty equivalent $t = 2$ lifetime utility of a high-output worker at $t = 1$ prior to the promotion decision. When all workers engaging in lower-level tasks exert effort, the firm outcome in the first period is a sufficient statistic for the mass of high-type workers in the firm. Let Q_{1i} be the firm outcome in the first period when all workers exert effort and let $\theta = \theta_i$:

$$Q_{1i} = x_L + \theta_i(x_H - x_L).$$

When $Q_1 \leq x_L + K(x_H - x_L)$, implying that $\theta \leq K$, a worker who generates high output in the first period receives a promotion in the second period with certainty:

$$\bar{U_2}(Q_1) = U_2(1).$$

When $Q_1 > x_L + K(x_H - x_L)$, implying that $\theta > K$, there are more high-output workers than there are available promotions. In these instances, a high-output worker receives a promotion in the second period with probability $\frac{K}{\theta}$:

$$\bar{U_2}(Q_1) = \left(\frac{K}{\theta}U_2(1)^{1-\gamma} + \left(1 - \frac{K}{\theta}\right)U_2(0)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}.$$

Let $U_1(q_1, Q_1)$ be the $t = 1$ lifetime utility of the worker as a function of that worker's individual output and total firm production in the first period. The $t = 1$ lifetime utility of the high-output ($q_1 = x_H$) worker can be written as

$$U_1(x_H, Q_1) = \left[W_1(x_H, Q_1)^{1-\frac{1}{\psi}} + \bar{U}_2(Q_1)^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\frac{1}{\psi}}}. \tag{8}$$

A worker with low or no output in the first period is not promoted in equilibrium. Consequently, the $t = 1$ lifetime utility of a worker with low or no output in the first period can be expressed as

$$U_1(q_1, Q_1) = \begin{cases} \left[(W_1(x_L, Q_1) + b_{1L})^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\frac{1}{\psi}}} & q_1 = x_L \\ \left[(W_1(0, Q_1) + b_{1L})^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\frac{1}{\psi}}} & q_1 = 0. \end{cases} \tag{9}$$

Let p_i^H be the high-type worker's posterior that $\theta = \theta_i$, i.e., $p_i^H = Prob(\theta = \theta_i | \eta = 1)$. When all other workers exert effort, the incentive compatibility constraint of a high-type worker can be written as

$$\begin{aligned} &\left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^H [W_1(x_H, Q_{1i})^{1-\frac{1}{\psi}} + \bar{U}_2(Q_{1i})^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \\ \geq &\left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^H [(W_1(0, Q_{1i}) + b_{1H})^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}}. \end{aligned} \tag{10}$$

Let p_i^L be the low-type worker's posterior that $\theta = \theta_i$, i.e., $p_i^L = Prob(\theta = \theta_i | \eta = 0)$. With certainty, a low-type worker generates low output with effort and no output without effort. When all other workers exert effort, the incentive compatibility constraint of a low-type worker can be expressed as

$$\begin{aligned} &\left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^L [W_1(x_L, Q_{1i})^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \\ \geq &\left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^L [(W_1(0, Q_{1i}) + b_{1L})^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}}. \end{aligned} \tag{11}$$

When the worker's outside option at $t = 0$ is sufficiently low ($u_1 \leq b_{1L} \leq b_{1H}$), a first-period contract that satisfies the incentive compatibility constraints of both types of workers, (10) and (11), also satisfies the worker's participation constraint. Hence, the firm's optimization problem in the first period can expressed as

$$\begin{aligned} \underset{\{W_1(q_1,Q_1)\geq 0\}}{\text{minimize}} \quad & E_0[W_1(q_1, Q_1)] \\ \text{subject to} \quad & (10) \text{ and } (11). \end{aligned}$$

2.3 Optimal First-Period Contract

Before solving for the optimal first-period contract, I eliminate an uninteresting case in which the implicit incentives from promotion alone are enough to induce effort from high-type workers in the first period. In this case, high-type workers exert effort in lower-level tasks even if the first-period contract pays nothing. The remainder of this paper focuses on the more interesting case in which the optimal first-period contract features positive payments for high output.[15]

Lemma 1 *The optimal first-period contract pays the worker nothing for no output and b_{1L} for low output:*

$$W_1(0, Q_1)^* = 0 \textit{ for all } Q_1,$$

$$W_1(x_L, Q_1)^* = b_{1L} \textit{ for all } Q_1.$$

The results of Lemma 1 follow from standard contracting arguments. Only workers who shirk generate no output. Because the likelihood ratio for no output is 0, the optimal contract minimizes the payment to the worker for no output. Limited liability implies that this minimum is 0. Only low-type workers generate low output. Consequently,

[15] A sufficient condition to ensure that the optimal first-period contract makes positive payments for high output is $b_{2H} < [b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$.

workers with low output do not face promotion risk because they are not candidates for promotion. A bonus of b_{1L} for low output, independent of firm outcomes, minimizes the expected wage cost necessary to incentivize effort from low-type workers due to their risk-aversion. The intimate relationship between promotion risk and firm performance pay in this setting is further highlighted in the following benchmark.

Lemma 2 *In the absence of promotion risk ($K \geq \theta_n$), the optimal first-period contract does not feature bonuses that depend on firm outcomes.*

When promotions are not scarce ($K \geq \theta_n$), workers with high output do not face promotion risk. In this case, firm performance pay does not provide a high-output worker with any insurance benefits. In addition, the firm outcome in the first period does not contain information about a worker's effort and type over and above that contained in the worker's individual output. Consequently, in the absence of promotion risk, the optimal first-period contract features payments based solely on the worker's individual output, consistent with the informativeness principle of Holmström (1979). Note that the same argument also holds when there are no opportunities for promotion ($K = 0$).

When promotions are scarce ($\theta_0 < K < \theta_n$), workers face promotion risk; in some states of the world, they do not receive a promotion despite generating high individual output. When workers face promotion risk, firm performance pay plays an insurance role in the optimal first-period contract.

Proposition 1 *When workers face promotion risk, the optimal first-period contract features a bonus for high individual output that increases in the firm outcome after a threshold:*

$$W_1(x_H, Q_{11})^* = ... = W_1(x_H, Q_{1k})^* < W_1(x_H, Q_{1k+1})^* < ... < W_1(x_H, Q_{1n})^*,$$

where k is the largest integer such that $Q_{1k} \leq x_L + K(x_H - x_L)$.

When workers face promotion risk, the structure of pay for high-output workers in the optimal contract stems from the insurance motive. The firm outcome (Q_1) provides a sufficient statistic for the mass of candidates for promotion, i.e., those who generated high output in the first period. When $Q_1 \leq x_L + K(x_H - x_L)$, there are fewer high-output workers than there are higher-level tasks. All workers with high output receive a promotion. In these instances, the optimal first-period contract features a bonus for high output that does not depend on the firm outcome. When $Q_1 > x_L + K(x_H - x_L)$, there are more high-output workers than there are higher-level tasks. In these instances, some workers with high output will not receive a promotion in the second period. In particular, a better firm outcome in the first period corresponds to a larger pool of promotion candidates and more intense competition for promotion in the next period. The optimal first-period contract compensates high-output workers for the decline in promotion prospects with an increase in pay. Consequently, the optimal first-period contract features payments for high output that are option-like in the firm outcome.

The key economic force underlying the insurance role of firm performance pay is the worker's preference for early resolution of uncertainty. It provides a necessary link between promotion prospects in a future period and the value of pay in the present period.

Proposition 2 *The worker's preference for early resolution of uncertainty is a necessary condition for firm performance pay to play an insurance role in the optimal first-period contract.*

To see why the preference for early resolution of uncertainty is important for firm performance pay to play an insurance role, it is useful to consider the case where workers do not have this preference. For example, when $\gamma = \frac{1}{\psi}$, the Epstein-Zin formulation simplifies to time-separable utility and the worker is indifferent over the timing of uncertainty resolution. In this case, the firm outcome still contains information about the worker's promotion prospects in the next period but does not affect the marginal utility of consumption in the current period. Hence, when workers are indifferent over the timing of uncertainty resolution, the optimal first-period contract, which equalizes the marginal utility of pay across states for the high-output worker, features the same payment for high output regardless of the firm outcome.[16]

In contrast, when the worker prefers early resolution of uncertainty, the arrival of information about tomorrow's consumption affects the marginal utility of today's consumption. In the context of the model, the additional anxiety from learning news about unfavorable promotion prospects in the next period results in added benefits from consumption in the current period. It is this link between poor promotion prospects in a future period and increased benefits from pay in the present period that allows firm performance pay to play an insurance role in the optimal contract in this setting.

3 Comparative Statics

This section provides some intuition for how the pay sensitivity of the optimal first-period contract varies in the cross section by deriving the comparative statics of the model in the limiting case where $\frac{1}{\psi}$ goes to 0. Numerical results suggest identical intuitions for larger values of $\frac{1}{\psi}$.

Let the high-output worker's pay sensitivity to firm outcomes be defined as Δ, which captures the expected percentage pay increase for a high-output worker due to better firm outcomes:

$$\Delta = E_0\left[\tfrac{W_1(q_1,Q_1)-W_1(q_1,Q_{11})}{W_1(q_1,Q_{11})}\Big|q_1 = x_H\right]. \tag{12}$$

Proposition 3 *In the limiting case where $\frac{1}{\psi}$ goes to 0, when promotions are very scarce ($K < \theta_1$), the pay sensitivity of the optimal first-period contract (Δ) increases in K. When promotions are not very scarce ($\theta_1 < K < \theta_n$), the pay sensitivity of the optimal first-period contract (Δ) decreases in K.*

Lemma 2 and Proposition 1 together imply that the relation between the pay sen-

[16] When $\gamma < \frac{1}{\psi}$, workers prefer late resolution of uncertainty. In this case, the optimal first-period contract features bonuses for high individual output that decline in firm outcomes after a threshold.

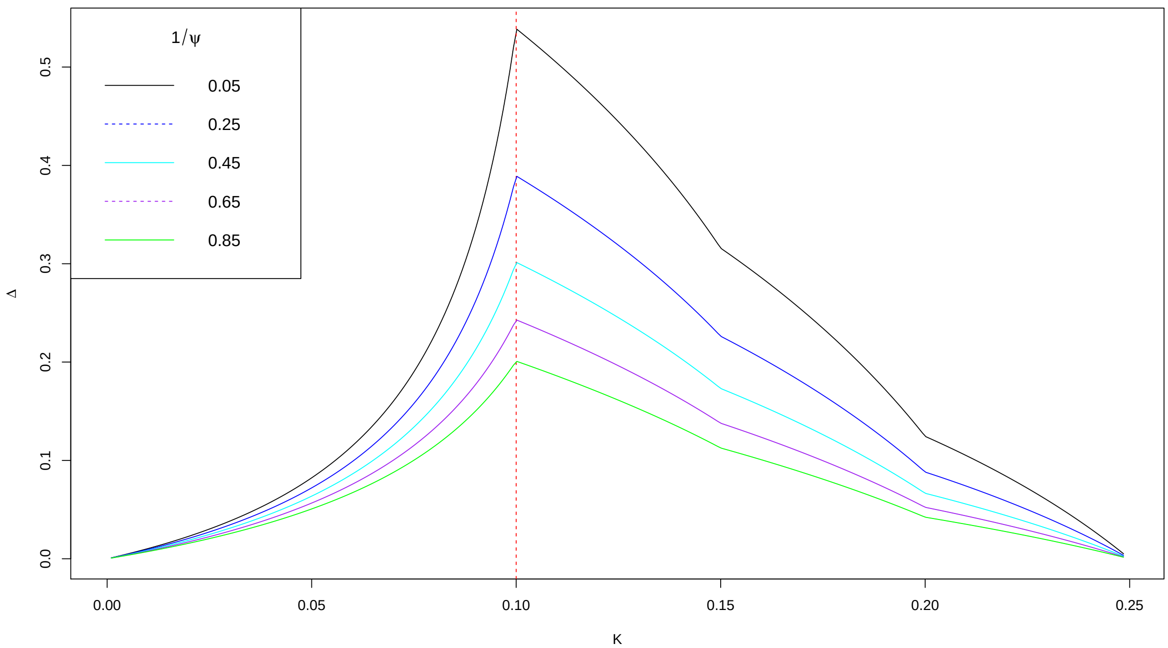

Figure 1. Pay sensitivity versus organizational structure (K). This figure plots an example of the optimal first-period contract's pay sensitivity to firm performance (Δ) against the firm's organizational structure, defined as the ratio of the number of higher-level tasks to the number of lower-level tasks (K) for different values of $\frac{1}{\psi}$. The figure was generated using parameters $\gamma = 8$, $b_{1H} = 1$, and $U_3 = 1$; θ takes values $0.1, 0.15, 0.2$, and 0.25 with equal probability; and $b_{2H} = 1.5$, $u_2 = 1$, and $K \in (0, 0.25)$.

sitivity of the optimal first-period contract (Δ) and the organizational structure (K) is non-monotonic. When $K = 0$, workers do not face promotion risk because there is no possibility of promotion. When $K \geq \theta_n$, workers also do not face promotion risk because promotions are not scarce. In both of these instances, the optimal first-period contract features no firm performance pay and $\Delta = 0$.

When promotions are available but scarce ($\theta_1 < K < \theta_n$), an increase in promotion opportunities has two effects on the worker's exposure to promotion risk. A larger value of K increases the value of implicit incentives from promotions, which means that uncertainty about promotion prospects translates into greater risk. However, an increase in K also reduces the likelihood that the worker is passed over for promotion. When promotions are very scarce ($K < \theta_1$), the first effect dominates. When promotions are

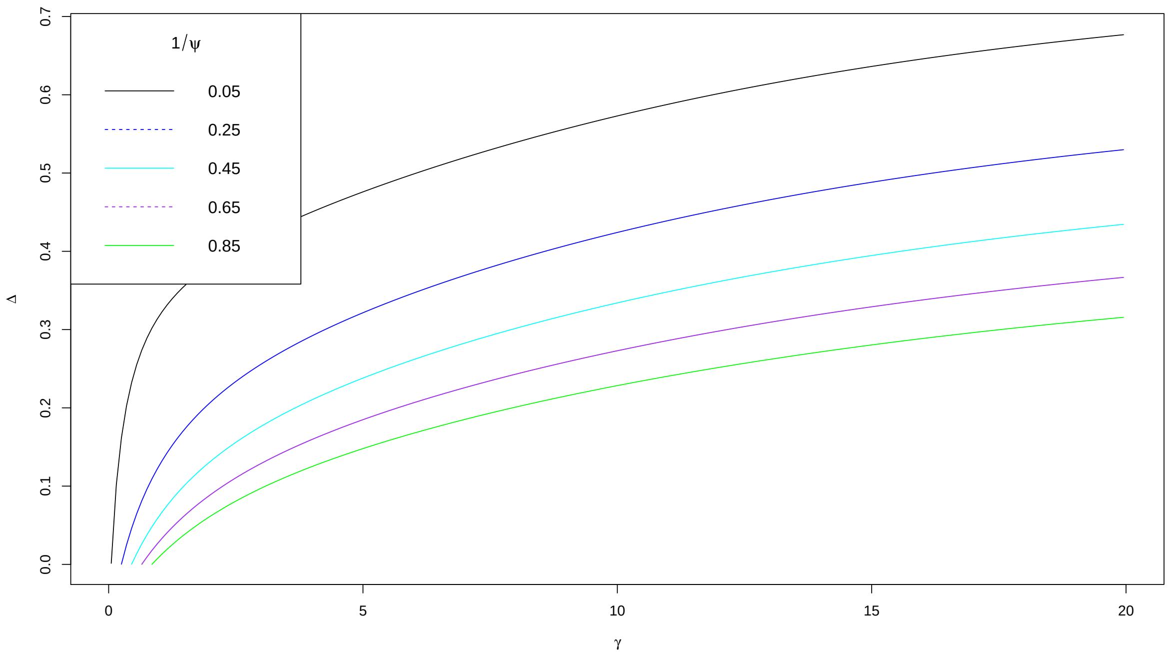

Figure 2. Pay sensitivity versus workers' relative risk aversion (γ). This figure plots an example of the optimal first-period contract's pay sensitivity to firm performance (Δ) against the worker's relative risk aversion (γ) for different values of $\frac{1}{\psi}$. The figure was generated using parameters $b_{1H} = 1$, $U_3 = 1$, and $K = 0.15$; θ takes values $0.1, 0.15, 0.2$, and 0.25 with equal probability; and $b_{2H} = 1.5$, $u_2 = 1$, and $\gamma \in (\frac{1}{\psi}, 20)$.

not very scarce ($K \geq \theta_1$), the second effect dominates. Figure 1 illustrates the non-monotonic relationship between the pay sensitivity of the optimal first-period contract and organizational structure for different values of $\frac{1}{\psi}$.

Proposition 4 *In the limiting case where $\frac{1}{\psi}$ goes to 0, when promotions are not very scarce ($K \geq \theta_1$), the pay sensitivity of the optimal first-period contract (Δ) increases in the worker's relative risk aversion (γ).*

In conventional principal-agent models with risk-averse workers, the trade-off between incentive provision and risk-sharing leads to a negative relation between measures of risk-aversion and pay sensitivity to firm outcomes.[17] In contrast, my framework features a

[17] As an exception, Inderst and Müller (2003) study a setting in which the owners of the firm cannot commit to operate in all states of the world. Workers accept firm performance pay in order to stave off the possibility of unemployment due to the owners shutting down the firm. If an increase in risk-

positive relation because firm performance pay serves as insurance against promotion risk. Intuitively, the more risk-averse workers are, the more valuable insurance against promotion risk becomes. Hence, pay sensitivity to firm outcomes in the optimal first-period contract increases in the worker's risk aversion. Figure 2 illustrates the positive relation between the pay sensitivity of the optimal first-period contract and the worker's risk aversion for different values of $\frac{1}{\psi}$.

It is worth noting that when promotions are very scarce ($K < \theta_1$), the pay sensitivity of the optimal first-period contract does not increase monotonically in the worker's risk aversion. The extreme scarcity of promotions implies that the high-output worker is not guaranteed a promotion regardless of the firm outcome. In this case, an increase in the worker's risk aversion reduces the value of implicit incentives from promotion for all realizations of θ. In the limit as γ goes to ∞, the value of implicit incentives from promotion declines to zero for all realizations of θ, which implies that the information about promotion prospects contained in the firm outcome is not payoff-relevant. Hence, in the limit as γ goes to ∞, the optimal first-period contract does not feature firm performance pay ($\Delta = 0$), implying that Δ must decline in γ for extreme values of γ when promotions are very scarce.

To study how an increase in uncertainty about the aggregate performance of workers at a firm affects the pay sensitivity of the optimal first-period contract, I adopt the concept of increased uncertainty as a mean-preserving spread following Rothschild and Stiglitz (1970).[18] Let $\hat{\theta}$, which also take values θ_1, θ_2, θ_3,..., θ_n, where $n \geq 4$, be a mean-preserving spread of θ.

A mean-preserving spread of θ has two opposing effects on the pay sensitivity of the optimal first-period contract. On the one hand, a mean-preserving spread of θ implies that the high-output worker is more likely to face intense competition for promotion. On

aversion makes unemployment more unpalatable than risky pay, then their model may also feature a positive relation.

[18] See the proof of Proposition 5 for more details on the construction of a mean-preserving spread.

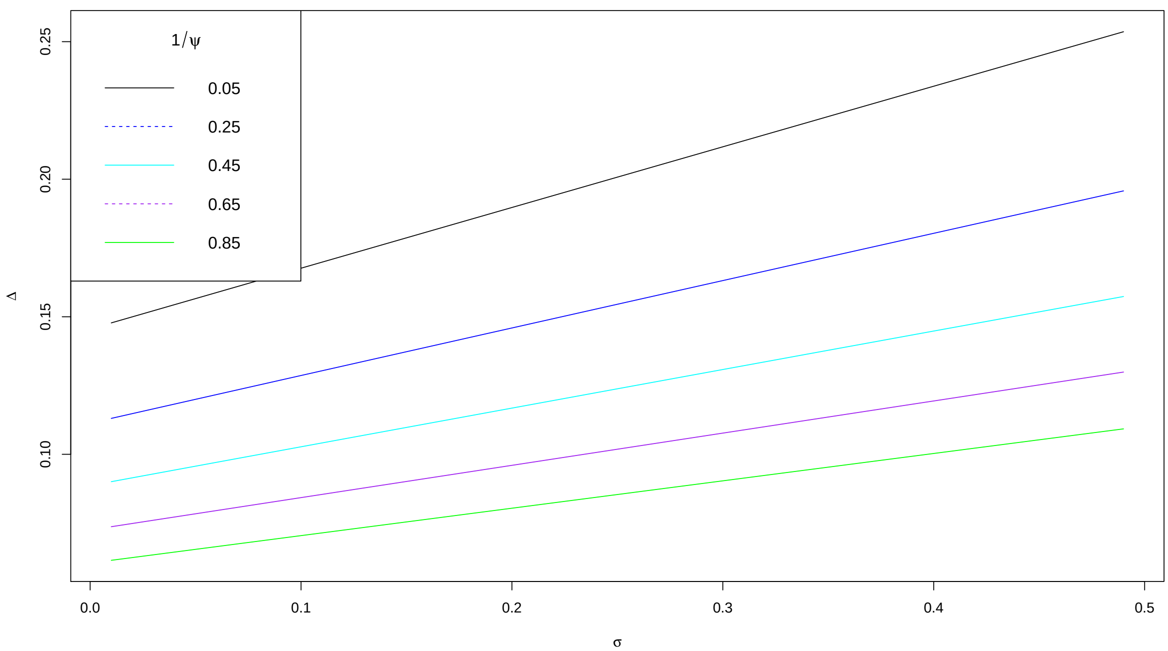

Figure 3. Pay sensitivity versus uncertainty about aggregate worker performance. This figure plots an example of the optimal first-period contract's pay sensitivity to firm performance (Δ) against the uncertainty about aggregate worker performance at the firm (σ) for different values of $\frac{1}{\psi}$. The figure was generated using parameters $\gamma = 8$, $b_{1H} = 1$, $U_3 = 1$, $K = 0.15$, $b_{2H} = 1.5$, and $u_2 = 1$; θ takes values $0.1, 0.15, 0.2$, and 0.25 with probabilities σ, $0.5 - \sigma$, $0.5 - \sigma$, and σ, respectively; and $\sigma \in (0, 0.5)$. Note that an increase in σ corresponds to a mean-preserving spread of the distribution of θ.

the other, a mean-preserving spread of θ also implies that the high-output worker is less likely to face moderate levels of competition for promotion. The following proposition provides a sufficient condition for the effects of the former to dominate those of the latter.

Proposition 5 *In the limiting case where $\frac{1}{\psi}$ goes to 0, for a given mean-preserving spread of θ, there exists some $\hat{K} \leq \theta_{n-2}$ such that $K \geq \hat{K}$ implies that*

$$\hat{\Delta} = \sum_{i=1}^{n} \hat{p}_i^H \left(\frac{W_1(x_H, Q_{1i}) - W_1(x_H, Q_{11})}{W_1(x_H, Q_{1i})} \right) > \sum_{i=1}^{n} p_i^H \left(\frac{W_1(x_H, Q_{1i}) - W_1(x_H, Q_{11})}{W_1(x_H, Q_{1i})} \right) = \Delta.$$

The crux of the argument underlying Proposition 5 is that if there are sufficient opportunities for promotion, then an increase in uncertainty about aggregate worker performance in the form of a mean-preserving spread unambiguously increases a high-output worker's exposure to promotion risk. Let $\hat{P}$ and P be the cumulative distribution function of $\hat{\theta}$ and θ, respectively. Rothschild and Stiglitz (1970) show that if $\hat{\theta}$ is a mean-preserving spread of θ, then there exists some $\hat{K}$ such that $1 - \hat{P}(z) \geq 1 - P(z)$ for all $z > \hat{K}$, with strict inequality for some z. In other words, for any value z above the cutoff $\hat{K}$, the realization of a mass of high-output workers above z is more likely under $\hat{\theta}$ than under θ. Hence, when promotions are not too scarce ($K \geq \hat{K}$), a mean-preserving spread of θ leads to more promotion risk for workers. The increase in promotion risk results in an optimal first-period contract with more firm performance pay and greater pay sensitivity to firm outcomes. Figure 3 illustrates the positive relation between pay sensitivity and uncertainty about aggregate worker performance at the firm for different values of $\frac{1}{\psi}$.

The formulation of a mean-preserving spread by Rothschild and Stiglitz (1970) is closely related to other notions of increased uncertainty. In particular, Rothschild and Stiglitz (1970) demonstrate that if $\hat{\theta}$ has second-order stochastic dominance over θ and the two random variables share the same support and mean, then $\hat{\theta}$ can be constructed from θ via a finite sequence of mean-preserving spreads. Hence, the result of Proposition 5 also extends to the case where increased uncertainty about aggregate worker performance at the firm is captured by the notion of second-order stochastic dominance.

Corollary 1 *If $\hat{\theta}$ has second-order stochastic dominance over θ and both random variables have the same support and mean, then there exists some $\hat{K} \leq \theta_{n-2}$ such that $K \geq \hat{K}$ implies that $\hat{\Delta} > \Delta$.*

4 Empirical Predictions

This section compares the predictions of the model to stylized facts about firm performance pay. While other explanations exist for each of the following phenomena, this paper provides a single framework with unifying explanations for the entire set of stylized facts.

4.1 Valuation of Stock Option Pay

Non-executive employees tend to value the stock options in their compensation above fair-market prices. For instance, Hodge et al. (2009) find that lower-level managers require an additional $38,688 in cash wages to forgo stock options in their pay package with a Black-Scholes value of $30,000 (see Oyer and Schaefer 2005 and Hallock and Olson 2006 for similar findings). Most explanations for this observation rely on relaxed assumptions about rationality such as optimism (Oyer and Schaefer 2005, Hallock and Olson 2006, Bergman and Jenter 2007), loss aversion (Devers et al. 2007), and inexperience (Hallock and Olson 2006, Hodge et al. 2009).

My framework provides a potential explanation for why non-executive workers value the stock options in their compensation above fair-market prices: the insurance role of firm performance pay. Workers who face promotion risk derive an insurance benefit from stock options in their employer. This insurance benefit increases their subjective valuations of the stock options in their compensation. This insurance motive may also help explain why some workers voluntarily buy shares in their employer at market prices. For example, Benartzi and Thaler (2001) document that non-executive workers often buy shares in their employer for their retirement plan and that they do not view equity in other firms as a substitute for equity in their employer. Non-executive workers may find it attractive to buy stock in their employer as a form of insurance against promotion risk if they are not adequately insured by their compensation contract.[19] In general, the benefit of insurance against promotion risk increases a worker's subjective valuation of firm performance pay such as stock options and equity.

The absence of insurance benefits also helps explain the opposite observation for CEOs. In stark contrast to the findings for non-executive workers, Bettis et al. (2005)

[19] For example, if workers have different degrees of risk aversion and the firm finds it too costly to offer individually tailored contracts, then some workers will be underinsured against promotion risk.

estimate that CEOs and board members value the stock options in their compensation at about 20% less than Black-Scholes value (see Hall and Murphy 2002 for similar evidence). Unlike non-executive workers, the individuals at the top of the corporate hierarchy have little need for insurance against promotion risk. After all, the CEO of a company cannot be promoted again at the same company. Consequently, CEOs and board members, who are underdiversified in firm-specific risks relative to the marginal investor, value stock option pay at less than fair market.

4.2 Risk Aversion

One surprising finding from the analysis of surveys from the National Bureau of Economic Research (NBER) on shared capitalism is a positive relation between risk-aversion and stock option pay (Kruse et al. 2008). In my framework, firm performance pay plays an insurance role. The more risk-averse the workers are, the more valuable insurance against promotion risk becomes. As long as promotion opportunities are not very scarce, the model predicts more risk-averse employees are more likely to receive firm performance pay, consistent with the NBER survey findings.

4.3 Performance-Based Vesting

One key feature of firm performance pay in this model is vesting based on individual output. Workers with low output do not face promotion risk because they are not candidates for promotion. Hence, the optimal contract features firm performance pay only for workers with high output. This characteristic matches common observations. For instance, equity-type compensation for non-executive employees routinely follows a vesting schedule over a period of four to five years. Workers lose this equity component of their pay if they are terminated. For example, DoubleClick, an internet advertising agency, famously fired many of its software developers for poor performance in the late 1990s; these employees lost their unvested stock options (Kowalski 2000). Because firms can fire bad workers, one can think of vesting periods as an implicit form of vesting based on individual performance. Contracts that contain explicit forms of performance-based vesting are also common. For instance, compensation plans for non-executive workers routinely include year-end bonuses based on two multipliers: one calculated from firm outcomes and the other from the worker's individual performance.

4.4 Option-Like Payoffs

The model predicts an option-like payoff for high-output workers. One common form of firm performance pay involves the use of employee stock options (ESOs), which feature option payoffs in the company's stock price.[20] This feature of option-like payoffs can also be seen in year-end bonus programs with minimum thresholds for firm outcomes such as a revenue target.

4.5 Volatility and Employee Stock Options

Finally, Spalt (2013) documents a puzzling empirical finding that firms with higher volatility in their stock returns are more likely to use broad-based employee stock option plans (ESO). In general, conventional contracting models have difficulty with this observation because when employees are risk-averse and hold rational expectations, the use of ESOs is more costly when there is more risk (Lambert et al. 1991, Hall and Murphy 2002, Bettis et al. 2005, Oyer and Schaefer 2005). Consequently, explanations for this phenomenon typically relax one or both of the assumptions about risk-aversion and rationality. For instance, Spalt (2013) posits that employees have gambling preferences for skewed, lottery-like payoffs that make stock options in riskier firms relatively more attractive.

In my model, the volatility of firm outcomes stems from the uncertainty about the

[20] According to the National Center of Employee Ownership, 8.5 to 13.4 million U.S. employees receive stock options in the company where they work as part of their compensation. https://www.nceo.org/articles/statistical-profile-employee-ownership

aggregate performance of its workers. The same uncertainty induces promotion risk. Because firm performance pay insures against promotion risk, such pay is more valuable at firms with more uncertainty about aggregate worker performance, which corresponds to riskier firm outcomes. Hence, my model predicts that the use of stock option pay and a portion of the risk in a firm's stock returns are positively linked via optimal contracting.

4.6 Additional Testable Implications

My model also generates a number of testable implications. First, it predicts a connection between a firm's organizational hierarchy, K, and the prevalence of firm performance pay. Proposition 3 highlights a non-monotonic relation between pay sensitivity to firm outcomes (Δ) and organizational structure, defined as the ratio of higher-level to lower-level tasks (K). To the best of my knowledge, this prediction is unique to my framework.

Second, the promotion tournament in my framework translates into promotion risk for talented employees because their skills are not fully transferable to other firms. Thus, the model predicts more firm performance pay at companies operating in industries and regions with more frictions that limit labor mobility.[21]

Third, Proposition 1 implies that the threshold for firm performance pay corresponds to available promotion opportunities. Consequently, the model predicts that the threshold for firm performance pay should be positively associated with the expected vacancy in higher-level positions.

Finally, my framework assumes a strong link between firm outcomes and the aggregate performance of workers. In industries where firm outcomes are largely driven by other stochastic factors such as macro-economic conditions or the productivity of capital, the use of firm performance pay as a form of insurance against promotion risk is less attractive because it involves significant basis risk. In other words, insuring workers

[21] For example, according to U.S. Treasury Department, nearly 30 million American workers are subject to non-compete agreements, which represent a substantial obstacle to labor mobility. The heterogeneity in the enforceability of such agreements at the state level and across time provides one possible test of the model's predictions. I thank Mark Kamstra for pointing this out.

against promotion risk with firm performance pay would at the same time expose them to other risks. Hence, the model's predictions should be more pronounced in firms and industries where human capital drives firm outcomes. To my knowledge, no paper has empirically tested these predictions.

5 Discussion

5.1 Promotion Risk

Promotion risk is a key component of this paper's results because the optimal first-period contract features firm performance pay in order to insure workers against promotion risk. In the main specification of the model, promotion risk arises because there are a limited number of higher-level tasks available at the worker's current employer. A high-output worker who is not promoted misses out on a pay increase because that worker's outside option pays less than the promotion does. In a scenario in which the firm can grow without limits to accommodate excess talent (e.g., Gibbons and Waldman 1999) or where workers can receive the same pay raise elsewhere (e.g., Harris and Holmström 1982, Gibbons and Murphy 1992, Holmström 1999), high-output workers receive the same high pay regardless of the promotion outcome. In such a scenario, firm performance pay does not play an insurance role because workers do not face promotion risk.

However, there are limits to a firm's ability to accommodate talent with new divisions and product lines. For example, Miller and Friesen (1984, p. 1171) point out in their seminal work on the organizational life cycle that mature firms "are conservative ... engage in very few efforts at diversification or acquisition ... and fail to even make many incremental changes to the products or services being offered" (see also Lester et al. 2003). Even when firms do not face bureaucratic hurdles, there are economic rationales for limiting firm expansion. For example, when human capital strongly affects the quality of a firm's goods and services, limiting expansion of the firm can serve as a signal of quality (see Levin and Tadelis 2005). As long as the firm cannot adjust the number of higher-level tasks to fully accommodate the pool of high-output workers, then there is promotion risk.

In the context of the model in this paper, suppose the number of higher-level tasks in the second period grows with the firm outcome in the first period. Recall that the firm outcome in the first period when $\theta = \theta_i$ is $Q_{1i} = x_L + \theta_i(x_H - x_L)$. Denote the number of available higher-level tasks in the second period as $K(Q_1)$. As before, define k as the largest integer such that $Q_{1k} \leq x_L + K(x_H - x_L)$.

Proposition 6 *The optimal first-period contract features firm performance pay if promotions are scarce and the relative increase in available promotions is smaller than that of firm outcomes:*

$$\frac{K(\theta_i)}{K(\theta_j)} < \frac{Q_{1i} - x_L}{Q_{1j} - x_L}$$

for all $i > j \geq k$.

The first condition in Proposition 6 guarantees that the pool of high-output workers is sufficiently large in some states of the world that some high-output workers cannot be promoted; otherwise, there is no promotion risk. The second condition ensures that the high-output worker's promotion prospects decline with better firm outcomes, which is the crux of Proposition 1's firm performance pay result.

Frictions that hinder labor mobility also contribute to promotion risk. For instance, according to the U.S. Department of the Treasury, nearly 30 million American workers are covered by non-compete agreements.[22] Consequently, even if talent is not firm-specific, workers may be contractually prevented from capitalizing on their talent at a different firm. Moreover, a worker's individual performance evaluation at one firm is typically not observable by other potential employers. As a result, a high-output worker who seeks greener pastures elsewhere after being passed over for promotion may be confronted by

[22] https://www.treasury.gov/resource-center/economic-policy/Documents/UST%20Non-competes%20Report.pdf

a labor market for lemons.

More broadly, as long as limitations to firm expansion and frictions in the labor market prevent workers from being fully rewarded for generating high output, then workers face promotion risk. If workers face promotion risk, then firm performance pay can play an insurance role in the optimal contract.

5.2 Long-Term Contracts

The main specification of the model considers one-period contracts. However, the main results hold when two-period contracts are allowed. A two-period contract at $t = 0$ specifies payments to the worker at $t = 1$ and at $t = 2$. The main difference is that a two-period contract allows the firm to commit to make a bonus payment to workers who generate high output in the first period but are not promoted in the second. This section focuses on the intuition of this extension. The Online Appendix contains further details.

As in the main specification, the performance pay result relies on two key ingredients: promotion risk and the worker's preference for early resolution of uncertainty. When no promotions are available ($K = 0$) or when promotions are not scarce ($K \geq \theta_n$), workers do not face promotion risk. In the absence of promotion risk, firm performance pay does not play an insurance role. Consequently, the optimal two-period contract rewards high output in lower-level tasks with a bonus payment in the first period that is independent of the firm outcome.

The pivotal role played by the worker's preference for early resolution of uncertainty in generating the firm performance pay result is again best understood by considering the case of $\gamma = \frac{1}{\psi}$, which corresponds to the worker being indifferent about the timing of uncertainty resolution. Using a two-period contract, the firm can insure the worker against promotion risk using firm performance pay at $t = 1$ or a bonus payment at $t = 2$, contingent on not being promoted. Firm performance pay in the first period only indirectly insures against adverse promotion outcomes in the second. The worker may receive low firm performance pay in the first period and still be passed over for promotion in the second. In contrast, a bonus payment at $t = 2$, contingent on the promotion outcome, directly insures the worker against adverse promotion outcomes. When the worker is indifferent over the timing of uncertainty resolution, direct insurance via a bonus payment at $t = 2$ dominates indirect insurance via firm performance pay at $t = 1$. Consequently, the optimal contract only insures against promotion risk via a bonus payment at $t = 2$, contingent on the high-output worker being passed over.

When the worker prefers early resolution of uncertainty, the promotion-contingent payments at $t = 2$ no longer dominate firm performance pay at $t = 1$ as a superior form of insurance against promotion risk. As noted earlier, firm outcomes in the first period contain information about a high-output worker's promotion prospects in the next period. In particular, better firm outcomes in the first period correspond to worse promotion prospects in the second. A worker who prefers early resolution of uncertainty derives an extra benefit from payments at $t = 1$ upon learning that promotion in the future is less likely. In other words, the worker benefits from insurance not only against unfavorable promotion outcomes, but also against news about unfavorable promotion prospects. The optimal two-period contract features firm performance pay in the first period as an indirect insurance against the former *and* a direct insurance against the latter. Consequently, the firm performance pay result persists even in an environment where long-term contracting is possible.

Finally, note that the optimal two-period contract equalizes the marginal utility to the high-output worker of an expected dollar paid at $t = 1$ and $t = 2$, subject to incentive compatibility in the lower- and higher-level tasks. Consequently, even if the high-output worker could borrow or save at a rate that is equivalent to that of the firm, without violating incentive compatibility, the high-output worker would not choose to do so given the optimal two-period contract. Hence, in a setting that allows for long-term contracts, the model's restriction preventing the worker from borrowing and saving is not binding under the optimal long-term contract.

5.3 Uniqueness of Equilibrium Outcome

The optimal contracts in Section 2 only specify the worker's pay on the equilibrium path. The incentive compatibility conditions in that section assume that all other workers put in effort. A contract that satisfies those conditions ensures that a worker does not have the incentive to unilaterally deviate. However, depending on how the contract is structured off the equilibrium path, alternative equilibriums exist.[23] For example, suppose the optimal contract from Section 3 pays the worker nothing at $t = 1$ when $Q_1 \notin \{Q_{11}, ..., Q_{1n}\}$. Then, all workers shirking can also be an equilibrium outcome given the contract. To induce effort from all workers as a unique equilibrium outcome, the optimal contract must also properly specify payments off the equilibrium path. For example, if the optimal first-period contract from Section 2.2 pays the high-output worker an arbitrarily large bonus when $Q_1 \notin \{Q_{11}, ..., Q_{1n}\}$, then it induces effort from all workers as a unique outcome.

6 Conclusion

This paper analyzes a principal-agent problem in a setting with promotion tournaments, which are common in the workplace. Because workers compete against each other for a limited number of promotions, when there are many high-performing workers, some are passed over for promotion despite having performed at a high level. Better firm outcomes correspond to a larger pool of high-performing workers and worse promotion prospects for an individual worker. The optimal contract features firm performance pay because it insures the worker against this promotion risk.

In this framework, the limited number of higher-level positions at the firm generates

[23] I thank Jungsuk Han, Mariassunta Giannetti, and Marcus Opp for their insights regarding multiple equilibria.

promotion risk, which results in an optimal contract with firm performance pay. Such pay, in its various forms, makes non-executive workers residual claimants to a portion of the firm's value. Empirical studies in the finance literature document many effects of equity-type pay for non-executive workers on corporate policies such as capital structure (McKeon 2015), innovation and investment (Chang et al. 2015, Babenko et al. 2011), and payout (Babenko 2009). Consequently, this paper highlights an underexplored connection between organizational economics and topics in finance.

In order to focus on the insurance role of firm performance pay for non-executive workers, I explore how one firm contracts with its workers in a partial equilibrium framework. A possible extension to this work would explore the macroeconomic implications of this insurance motive in a general equilibrium setting. In particular, my results suggest that an increase in the specialization of higher-level tasks, which magnifies promotion risk, may lead to greater ex-post income disparities between non-executive workers at *different* firms due to the ex-ante provision of insurance against promotion risk via firm performance pay.

Finally, this paper contributes to the study of optimal contracting under Epstein-Zin preferences. This paper identifies a setting in which modeling with Epstein-Zin preferences, instead of conventional time-separable utility, results in *qualitatively* different predictions about the optimal contract. Specifically, the optimal contract for non-executive workers features firm performance pay as a form of insurance only when workers prefer early resolution of uncertainty. This result suggests that further work in contracting with Epstein-Zin utility may uncover other economic forces overlooked by conventional studies using time-separable utility.

References

Ai, Hengjie and Ravi Bansal, "Risk preferences and the macroeconomic announcement premium," *Econometrica*, 2018, *86* (4), 1383–1430.

Axelson, Ulf and Philip Bond, "Wall street occupations," *Journal of Finance*, 2015, *70* (5), 1949–1996.

Babenko, Ilona, "Share repurchases and pay-performance sensitivity of employee compensation contracts," *Journal of Finance*, 2009, *64* (1), 117–150.

_ , Michael Lemmon, and Yuri Tserlukevich, "Employee stock options and investment," *Journal of Finance*, 2011, *66* (3), 981–1009.

Benartzi, Shlomo and Richard H Thaler, "Naive diversification strategies in defined contribution saving plans," *American Economic Review*, 2001, *91* (1), 79–98.

Bergman, Nittai K and Dirk Jenter, "Employee sentiment and stock option compensation," *Journal of Financial Economics*, 2007, *84* (3), 667–712.

Bettis, J Carr, John M Bizjak, and Michael L Lemmon, "Exercise behavior, valuation, and the incentive effects of employee stock options," *Journal of Financial Economics*, 2005, *76* (2), 445–470.

Bognanno, Michael L, "Corporate tournaments," *Journal of Labor Economics*, 2001, *19* (1), 290–315.

Boudreau, Kevin J, Karim R Lakhani, and Michael Menietti, "Performance responses to competition across skill levels in rank-order tournaments: field evidence and implications for tournament design," *RAND Journal of Economics*, 2016, *47* (1), 140–165.

Bova, Francesco and Liyan Yang, "Employee bargaining power, inter-firm competition, and equity-based compensation," *Journal of Financial Economics*, 2017, *126* (2), 342–363.

Chang, Xin, Kangkang Fu, Angie Low, and Wenrui Zhang, "Non-executive employee stock options and corporate innovation," *Journal of Financial Economics*, 2015, *115* (1), 168–188.

Chen, Kong-Pin, "Sabotage in promotion tournaments," *Journal of Law, Economics, and Organization*, 2003, *19* (1), 119–140.

Cichello, Michael S, C Edward Fee, Charles J Hadlock, and Ramana Sonti, "Promotions, turnover, and performance evaluation: Evidence from the careers of division managers," *The Accounting Review*, 2009, *84* (4), 1119–1143.

Core, John E and Wayne R Guay, "Stock option plans for non-executive employees," *Journal of Financial Economics*, 2001, *61* (2), 253–287.

DeMarzo, Peter M and Ron Kaniel, "Relative pay for non-relative performance: Keeping up with the Joneses with optimal contracts," *Available at SSRN 2755856*, 2017.

_ , _ , and Ilan Kremer, "Diversification as a public good: Community effects in portfolio choice," *Journal of Finance*, 2004, *59* (4), 1677–1716.

_ , _ , and _ , "Technological innovation and real investment booms and busts," *Journal of Financial Economics*, 2007, *85* (3), 735–754.

_ , _ , and _ , "Relative wealth concerns and financial bubbles," *Review of Financial Studies*, 2008, *21* (1), 19–50.

DeVaro, Jed, "Internal promotion competitions in firms," *RAND Journal of Economics*, 2006, *37* (3), 521–542.

Devers, Cynthia E, Robert M Wiseman, and R Michael Holmes Jr, "The effects of endowment and loss aversion in managerial stock option valuation," *Academy of Management Journal*, 2007, *50* (1), 191–208.

di Tella, Sebastian, "Uncertainty shocks and balance sheet recessions," *Journal of Political Economy*, 2017, *125* (6), 2038–2081.

_ , "Optimal regulation of financial intermediaries," *American Economic Review*, 2019, *109* (1), 271–313.

Edmans, Alex and Xavier Gabaix, "Tractability in incentive contracting," *The Review of Financial Studies*, 2011, *24* (9), 2865–2894.

Efing, Matthias, Harald Hau, Patrick Kampkötter, and Jean-Charles Rochet, "Bank Bonus Pay as a Risk Sharing Contract," *Swiss Finance Institute Research Paper*, 2018.

Epstein, LR and Stanley E Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica: Journal of the Econometric Society*, 1989, *57* (4).

Fairburn, James A and James M Malcomson, "Performance, promotion, and the Peter Principle," *Review of Economic Studies*, 2001, *68* (1), 45–66.

Fehr, Ernst and Klaus M Schmidt, "A theory of fairness, competition, and cooperation," *Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

Gibbons, Robert and Kevin J Murphy, "Optimal incentive contracts in the presence of career concerns: Theory and evidence," *Journal of political Economy*, 1992, *100* (3), 468–505.

_ **and Michael Waldman**, "A theory of wage and promotion dynamics inside firms," *Quarterly Journal of Economics*, 1999, *114* (4), 1321–1358.

Goel, Anand M and Anjan V Thakor, “Do envious CEOs cause merger waves?,” *Review of Financial Studies*, 2010, *23* (2), 487–517.

Green, Jerry R and Nancy L Stokey, “A comparison of tournaments and contracts,” *Journal of Political Economy*, 1983, *91* (3), 349–364.

Hall, Brian J and Kevin J Murphy, “Stock options for undiversified executives,” *Journal of accounting and economics*, 2002, *33* (1), 3–42.

_ and _ , “The trouble with stock options,” *Journal of Economic Perspectives*, 2003, *17* (3), 49–70.

Hallock, Kevin F and Craig Olson, “The value of stock options to non-executive employees,” 2006. Working Paper, NBER.

Harris, Milton and Bengt Holmström, “A theory of wage dynamics,” *Review of Economic Studies*, 1982, *49* (3), 315–333.

Hodge, Frank D, Shiva Rajgopal, and Terry Shevlin, “Do managers value stock options and restricted stock consistent with economic theory?,” *Contemporary Accounting Research*, 2009, pp. 74–91.

Holmström, Bengt., “Moral hazard and observability,” *Bell Journal of Economics*, 1979, pp. 74–91.

Holmström, Bengt, “Managerial incentive problems: A dynamic perspective,” *Review of Economic Studies*, 1999, *66* (1), 169–182.

Inderst, Roman and Holger M Müller, “A Theory of Broad-Based Option Pay,” 2003.

Kim, E Han and Paige Ouimet, “Broad-Based Employee Stock Ownership: Motives and Outcomes,” *Journal of Finance*, 2014, *69* (3), 1273–1319.

Kowalski, Robert, "As Options Spread, So Too Do Suits From Workers Fired Before They Vest," *The Street*, 2000.

Kreps, David M and Evan L Porteus, "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica: Journal of the Econometric Society*, 1978, pp. 185–200.

Kruse, Douglas L, Joseph R Blasi, and Rhokeun Park, "Shared Capitalism in the U.S. Economy? Prevalence, Characteristics, and Employee Views of Financial Participation in Enterprises," 2008. Working Paper, NBER.

Lambert, Richard A, David F Larcker, and Robert E Verrecchia, "Portfolio Considerations in Valuing Executive Compensation," *Journal of Accounting Research*, 1991, *29* (1), 129–149.

Lazear, Edward P, "Pay equality and industrial politics," *Journal of Political Economy*, 1989, *97* (3), 561–580.

_ , "Output-based pay: Incentives, retention or sorting?," in "Accounting for worker well-being," Emerald Group Publishing Limited, 2004, pp. 1–25.

_ **and Sherwin Rosen**, "Rank-order tournaments as optimum labor contracts," *Journal of Political Economy*, 1981, *89* (5), 841–864.

Lester, Donald L, John A Parnell, and Shawn Carraher, "Organizational life cycle: A five-stage empirical scale," *The International Journal of Organizational Analysis*, 2003, *11* (4), 339–354.

Levin, Jonathan and Steven Tadelis, "Profit sharing and the role of professional partnerships," *Quarterly Journal of Economics*, 2005, *120* (1), 131–171.

McKeon, Stephen B, "Employee option exercise and equity issuance motives," *Available at SSRN: http://dx.doi.org/10.2139/ssrn.1920985*, 2015.

Miller, Danny and Peter H Friesen, "A longitudinal study of the corporate life cycle," *Management Science*, 1984, *30* (10), 1161–1183.

Nalebuff, Barry J and Joseph E Stiglitz, "Prizes and incentives: towards a general theory of compensation and competition," *Bell Journal of Economics*, 1983, pp. 24–43.

O'Keeffe, Mary, W Kip Viscusi, and Richard J Zeckhauser, "Economic contests: Comparative reward schemes," *Journal of Labor Economics*, 1984, *2* (1), 27–56.

Oyer, Paul, "Why do firms use incentives that have no incentive effects?," *Journal of Finance*, 2004, *59* (4), 1619–1650.

_ **and Scott Schaefer**, "Why do some firms give stock options to all employees?: An empirical examination of alternative theories," *Journal of Financial Economics*, 2005, *76* (1), 99–133.

Prendergast, Canice, "The insurance effect of groups," *International Economic Review*, 1992, pp. 567–581.

Rothschild, Michael and Joseph E Stiglitz, "Increasing risk: I. A definition," *Journal of Economic Theory*, 1970, *2* (3), 225–243.

Spalt, Oliver G, "Probability weighting and employee stock options.," *Journal of Financial and Quantitative Analysis*, 2013, *48* (4), 1085–1118.

Sun, Qi and Mindy Z Xiaolan, "Financing intangible capital," *Journal of Financial Economics*, 2019, *133* (3), 564–588.

Terviö, Marko, "Superstars and mediocrities: Market failure in the discovery of talent," *Review of Economic Studies*, 2009, *76* (2), 829–850.

Vladimirov, Vladimir, "Financing Skilled Labor," in "WFA 2019 Meetings Paper" 2019.

A Micro-foundation for Promotion Policy

The main specification of the paper assumes that the firm promotes workers with high output in lower-level tasks with equal probability. This section micro-founds this assumption in a setting where the worker's productivity in higher-level tasks also depends on a parameter ζ_a, which captures a notion of fit between the worker and higher-level tasks at the firm. The firm learns the fit (ζ_a) between a high-output worker and the higher-level tasks at the beginning of period 2 at no cost. A high-type worker who exerts effort generates an output of $q_{2a} = y_H + \zeta_a$. As in the main specification, a high-type worker who exerts no effort receives private benefits in the form of an additional b_{2H} units of consumption at $t = 2$ but generates no output ($q_{2a} = 0$) with certainty. A low-type worker always generates no output ($q_{2a} = 0$) and receives no private benefits from shirking in higher-level tasks. For each high-output worker, the parameter ζ_a is identically and independently distributed according to a uniform distribution on $[0, \theta]$.

In equilibrium, given a mass θ_i of high-output workers in the first period, a profit-maximizing firm promotes all high-output workers in the second period if $\theta_i \leq K$ because by assumption the promotion of a high-output worker, who must be high-type, to the higher-level tasks is NPV positive. If $\theta_i > K$, a profit maximizing firm promotes the K high-output workers with the highest values of ζ_a. The i.i.d assumption on the distribution of ζ_a implies that ex-ante, workers expect to be promoted with equal probability in the second period if they generate high output in the first period.

B Proofs

Proof of Lemma 1. The optimal first-period contract must satisfy the IC of low- and high-type workers. The IC of a low-type worker is given by (11):

$$\left[c_0^{1-\frac{1}{\psi}}+\left(\sum_{i=1}^{n}p_i^L\left[W_1(x_L,Q_{1i})^{1-\frac{1}{\psi}}+u_2^{1-\frac{1}{\psi}}+U_3^{1-\frac{1}{\psi}}\right]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

$$\geq\left[c_0^{1-\frac{1}{\psi}}+\left(\sum_{i=1}^{n}p_i^L\left[(W_1(0,Q_{1i})+b_{1L})^{1-\frac{1}{\psi}}+u_2^{1-\frac{1}{\psi}}+U_3^{1-\frac{1}{\psi}}\right]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

Only workers who shirk generate no output. Hence, any payment made for no output disincentivizes effort from workers. The optimal first-period contract minimizes the payment for no output; limited liability implies that this minimum is 0. When $W_1(0,Q_{1i})=0$ for all Q_{1i}, the low-type worker's IC simplifies to

$$\left[c_0^{1-\frac{1}{\psi}}+\left(\sum_{i=1}^{n}p_i^L\left[W_1(x_L,Q_{1i})^{1-\frac{1}{\psi}}+u_2^{1-\frac{1}{\psi}}+U_3^{1-\frac{1}{\psi}}\right]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}\geq\left[c_0^{1-\frac{1}{\psi}}+b_{1L}^{1-\frac{1}{\psi}}+u_2^{1-\frac{1}{\psi}}+U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

For convenience of notation, let $W_1(x_L,Q_{1i})=w_{1i}$. Let G_L be the lifetime utility that the first-period contract yields to the worker for low output:

$$G_L=\left[c_0^{1-\frac{1}{\psi}}+\left(\sum_{i=1}^{n}p_i^L\left[w_{1i}^{1-\frac{1}{\psi}}+u_2^{1-\frac{1}{\psi}}+U_3^{1-\frac{1}{\psi}}\right]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

By construction, the simplified IC of the low-type worker is satisfied if and only if $G_L\geq\left[c_0^{1-\frac{1}{\psi}}+b_{1L}^{1-\frac{1}{\psi}}+u_2^{1-\frac{1}{\psi}}+U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$. Note that a first-period contract that pays the worker nothing for low output is not incentive compatible:

$$G_L = \left[c_0^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}} < \left[c_0^{1-\frac{1}{\psi}} + b_{1L}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

Hence, the optimal first-period contract must make a positive payment for low output in some states of the world. Let m_{1i} be the marginal utility to the worker of an expected dollar paid as w_{1i}:

$$m_{1i} = \frac{1}{(1-\bar{\theta})p_i^L}\frac{\partial G_L}{\partial w_{1i}},$$

where $1-\bar{\theta}$ is the unconditional probability that a worker is low-type. The optimal first-period contract equalizes the marginal utility of an expected dollar paid as w_{1i} for different realizations of θ_i, subject to limited liability. Otherwise, a cheaper incentive-compatible contract can be constructed by a reallocation of payments. Because m_{1i} approaches positive infinity as m_{1i} goes to zero from the right, the optimal first-period contract features $w_{1i}^* > 0$ for all i. Denote the quantities associated with the optimal contract by *. An interior solution implies that for any $i, j \in \{1, ..., n\}$,

$$\begin{aligned} m_{1i}^* &= m_{1j}^* \\ &\Leftrightarrow \\ [w_{1i}^{*1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} w_{1i}^{*-\frac{1}{\psi}} &= [w_{1j}^{*1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} w_{1j}^{*-\frac{1}{\psi}} \\ &\Leftrightarrow \\ w_{1i}^* &= w_{1j}^* \end{aligned} \tag{B.1}$$

because $[w_1^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} w_1^{-\frac{1}{\psi}}$ strictly decreases in w_1. Equation B.1 implies that $w_{1i}^* = w_1^*$ for all i and G_L^* can be expressed as

$$G_L = \left[c_0^{1-\frac{1}{\psi}} + w_1^{*1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

An interior solution for w_{1i}^* implies that the low-type worker's IC must bind under the optimal first-period contract; otherwise, a cheaper incentive compatible contract can be constructed by slightly lowering payments for low output. A binding IC for low-type workers implies that $w_{1i}^* = w_1^* = b_{1L}$ for all i.

Proof of Lemma 2. When promotions are not scarce ($K \geq \theta_n$), workers do not face promotion risk. A worker who generates high output is guaranteed a promotion. Hence, a high-output worker's certainty equivalent $t = 2$ lifetime utility is $\bar{U}_2(Q_1) = [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$ for all Q_1; call this quantity U_2. Recall from Lemma 1 that $W_1(0, Q_1)^* = 0$ for all Q_1. The high-type worker's IC simplifies to

$$\left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^H \left[W_1(x_H, Q_{1i})^{1-\frac{1}{\psi}} + U_2^{1-\frac{1}{\psi}}\right]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}} \geq \left[c_0^{1-\frac{1}{\psi}} + b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

For convenience of notation, let $W_1(x_H, Q_{1i}) = W_{1i}$. Let G_H be the lifetime utility that the first-period contract yields to the worker for high output:

$$G_H = \left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^H \left[W_{1i}^{1-\frac{1}{\psi}} + U_2^{1-\frac{1}{\psi}}\right]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

By construction, the IC of the high-type worker is satisfied if and only if

$$G_H \geq \left[c_0^{1-\frac{1}{\psi}} + b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

Denote the quantities associated with the optimal contract by *. There are two cases to consider. If $U_2 \geq [b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$, then the high-type worker's IC is satisfied even if the first-period contract makes no payment for high output. In this case, the implicit incentives from promotion alone are enough to motivate effort from high-type workers and the optimal first-period contract makes no payment for high output: $W_{1i}^* = 0$

for all i.

If $U_2 < [b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$, then a first-period contract that makes no payment for high output is not incentive compatible. Hence, the optimal first-period contract must make a positive payment for high output in some states of the world. Let M_{1i} be the marginal utility to the worker of an expected dollar paid as W_{1i}:

$$M_{1i} = \frac{1}{\bar{\theta} p_i^H} \frac{\partial G_H}{\partial W_{1i}},$$

where $\bar{\theta}$ is the unconditional probability that a worker is high-type. The optimal first-period contract equalizes the marginal utility of an expected dollar paid as W_{1i}, subject to limited liability. Otherwise, a cheaper incentive-compatible contract can be constructed by a reallocation of payments. Because M_{1i} approaches positive infinity as M_{1i} goes to zero from the right, the optimal first-period contract features $W_{1i}^* > 0$ for all i. An interior solution for W_{1i}^* implies that for any $i, j \in \{1, ..., n\}$,

$$\begin{aligned} M_{1i}^* &= M_{1j}^* \\ &\Leftrightarrow \\ [W_{1i}^{*1-\frac{1}{\psi}} + U_2^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1i}^{*-\frac{1}{\psi}} &= [W_{1j}^{*1-\frac{1}{\psi}} + U_2^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1j}^{*-\frac{1}{\psi}} \\ &\Leftrightarrow \\ W_{1i}^* &= W_{1j}^* \end{aligned}$$

because $[W_1^{1-\frac{1}{\psi}} + U_2^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_1^{-\frac{1}{\psi}}$ strictly decreases in W_1. Consequently, the optimal first-period contract features $W_{1i}^* = W_1^*$ for all i for some $W_1^* > 0$.

Note that the crux of the argument is that the high-output worker's certainty equivalent $t = 2$ lifetime utility is independent of the firm outcome in the first period. The same argument can be made when $K = 0$. In this case, high-output worker's certainty equivalent $t = 2$ lifetime utility is $\bar{U_2}(Q_1) = [u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$ for all Q_1.

Proof of Proposition 1 and Proposition 2. For convenience of notation, let $W_1(1, Q_{1i}) = W_{1i}$ and $\bar{U}_2(Q_{1i}) = U_{2i}$. Define k as the largest integer such that $Q_{1k} \leq x_L + K(x_H - x_L)$. Denote the quantities associated with the optimal contract by *.

Lemma 1 simplifies the high-type worker's IC to

$$\left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^H [W_{1i}^{1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}} \geq \left[c_0^{1-\frac{1}{\psi}} + b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}},$$

where

$$U_{2i} = \begin{cases} [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}} & i \leq k \\ \left(\frac{K}{\theta_i}[b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} + \left(1 - \frac{K}{\theta_i}\right)[u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1}{1-\gamma}} & i > k\ . \end{cases}$$

Let G_H be the lifetime utility that a first-period contract delivers to the worker for high output:

$$G_H = \left[c_0^{1-\frac{1}{\psi}} + \left(\sum_{i=1}^{n} p_i^H [W_{1i}^{1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

For $i \in \{1, ..., n\}$, let M_{1i} be the marginal utility to the worker of an expected dollar paid as W_{1i}:

$$M_{1i} = \frac{1}{\bar{\theta} p_i} \frac{\partial G_H}{\partial W_{1i}} = \frac{G_H^{\gamma}}{\bar{\theta}} \left(\sum_{i=1}^{n} p_i [W_{1i}^{1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{\gamma-\frac{1}{\psi}}{1-\gamma}} [W_{1i}^{1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1i}^{-\frac{1}{\psi}} > 0,$$

where $\bar{\theta}$ is the unconditional probability that a worker is high-type. Because M_{1i} approaches positive infinity as W_{1i} goes to zero from the right, the optimal first-period contract features an interior solution, i.e., $W_{1i}^* > 0$ for all $i \in \{1, ..., n\}$. Moreover, the optimal first-period contract equalizes the marginal utility of an expected dollar paid at $t = 1$; otherwise, a cheaper incentive compatible contract can be constructed by reallocating payments. Consequently, for any $i, j \in \{1, ..., n\}$,

$$M_{1i}^{*} = M_{1j}^{*}$$

$$\Leftrightarrow$$

$$[W_{1i}^{*1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1i}^{*-\frac{1}{\psi}} = [W_{1j}^{*1-\frac{1}{\psi}} + U_{2j}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1j}^{*-\frac{1}{\psi}}. \tag{B.2}$$

When the worker is indifferent over the timing of uncertainty resolution ($\gamma = \frac{1}{\psi}$), (B.2) simplifies to

$$W_{1i}^{*-\frac{1}{\psi}} = W_{1j}^{*-\frac{1}{\psi}},$$

which implies that $W_{1i}^{*} = W_{1j}^{*}$ for all $i, j \in \{1, ..., n\}$. In this case, the optimal first-period contract pays the worker a bonus for high output that does not depend on the firm outcome.

When the worker prefers late resolution of uncertainty ($\gamma < \frac{1}{\psi}$), (B.2) implies that

$$W_{1i}^{*} \leq W_{1j}^{*}$$

for all $i, j \in \{1, ..., n\}$ and $i > j$ because $[W_{1}^{1-\frac{1}{\psi}} + U_{2}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1}^{-\frac{1}{\psi}}$ strictly increases in U_2 and strictly decreases in W_1. In this case, the optimal first-period contract feature a payment for high output that weakly decreases with the firm outcome.

Consequently, when workers do not prefer early resolution of uncertainty ($\gamma \leq \frac{1}{\psi}$), the optimal first-period contract does not feature firm performance pay—a payment that increases with both the worker's individual output and the firm outcome. This demonstrates the claim of Proposition 2.

Finally, consider optimal first-period contract when the worker prefers early resolution of uncertainty ($\gamma > \frac{1}{\psi}$). Recall that when all workers exert effort, the firm outcome in the first period is determined by the mass of high-type workers at the firm:

$$Q_{1i} = x_L + \theta_i(x_H - x_L) \text{ for } i \in \{1, ..., n\}.$$

When the firm outcome (Q_1) is below the threshold $x_L + K(x_H - x_L)$, the corresponding mass of high-output workers at the firm is below K. In these instances, all high-output workers receive a promotion. Hence, $U_{2i} = U_{2j} = [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$ for all $i, j \leq k$. Accordingly, (B.2) implies that $W_{1i}^* = W_{1j}^*$ for all $i, j \leq k$. When the firm outcome exceeds the threshold $x_L + K(x_H - x_L)$, the corresponding mass of high-output workers at the firm is above K. In these instances, a high-output worker receives a promotion in the second period with probability $\frac{K}{\theta_i}$. Hence, for $i \geq k$, the high-output worker's certainty equivalent $t = 2$ lifetime utility is

$$U_{2i} = \left(\frac{K}{\theta_i}[b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} + \left(1 - \frac{K}{\theta_i}\right)[u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right)^{\frac{1}{1-\gamma}}$$

Note that U_{2i} declines in θ_i because $b_{2H} > u_2$. Recall that θ_i is indexed in order of ascending magnitude. Consequently, $i > j$ implies that $\theta_i > \theta_j$, which implies that $U_{2i} < U_{2j}$. For $i > j \geq k$, (B.2) implies that

$$M_{1i} = M_{1j}$$

$$\Leftrightarrow$$

$$[W_{1i}^{*1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1i}^{*-\frac{1}{\psi}} = [W_{1j}^{*1-\frac{1}{\psi}} + U_{2j}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1j}^{*-\frac{1}{\psi}}$$

$$\Leftrightarrow$$

$$W_{1i}^* > W_{1j}^*$$

because $[W_1^{1-\frac{1}{\psi}} + U_2^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_1^{-\frac{1}{\psi}}$ strictly increases in U_2 and strictly decreases in W_1.

Consequently, when workers prefer early resolution of uncertainty, the optimal first-period contract features a bonus for high individual output that increases with the firm outcome after a threshold. This demonstrates the claim of Proposition 1. In other words, the optimal first-period contract features firm performance pay as insurance against promotion risk if and only if workers prefer early resolution of uncertainty ($\gamma > \frac{1}{\psi}$).

Proof of Proposition 3. Denote the quantities associated with the optimal contract by *. For convenience of notation, let $W_1(x_H, Q_{1i})^* = W_{1i}$ and $\bar{U}_2(Q_{1i}) = U_{2i}$.

The optimal first-period contract equalizes the marginal utility of an expected dollar paid for high output for different realizations of total firm production. The equalization of marginal utility implies that

$$[W_{11}^{1-\frac{1}{\psi}} + U_{21}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{11}^{-\frac{1}{\psi}} = [W_{1i}^{1-\frac{1}{\psi}} + U_{2i}^{1-\frac{1}{\psi}}]^{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}} W_{1i}^{-\frac{1}{\psi}} \text{ for all } i. \tag{B.3}$$

In the limiting case where $\frac{1}{\psi} = 0$, the optimal contract can be solved in closed form. In particular, (B.3) implies that

$$W_{1i} = W_{11} + U_{21} - U_{2i} \tag{B.4}$$

for all i under the parameter restriction that $b_{2H} < [b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$.

In this case, Δ can be expressed as

$$\Delta = \sum_{i=1}^{n} p_i^H \frac{U_{21} - U_{2i}}{W_{11}}. \tag{B.5}$$

Note that if $b_{2H} \geq [b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}$, $W_{11} = 0$ and Δ is not well-defined.

The optimal contract also binds the IC of high-type workers:

$$\left(\sum_{i=1}^{n} p_i^H (W_{1i} + U_{2i})^{1-\gamma}\right)^{\frac{1}{1-\gamma}} = \underbrace{[b_{1H}^{1-\frac{1}{\psi}} + u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}}_{\text{Call this } \mu_1}. \tag{B.6}$$

Substituting (B.4) into (B.6) implies that

$$W_{11} + U_{21} = \mu_1. \tag{B.7}$$

Differentiating (B.7) with respect to K yields:

$$\frac{dW_{11}}{dK} = -\frac{dU_{21}}{dK}. \tag{B.8}$$

There are two region of K to consider: $K < \theta_1$ and $K \geq \theta_1$.

Region 1: $K < \theta_1$

Recall that when $K < \theta_1$, U_{2i} can be expressed as

$$U_{2i} = \left(\tfrac{K}{\theta_i} [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} + (1 - \tfrac{K}{\theta_i}) [u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1}{1-\gamma}}.$$

Moreover, note that

$$\frac{dU_{2i}}{dK} = \frac{U_{2i}^{\gamma}}{\theta_i} \frac{\Phi}{1-\gamma} > 0, \tag{B.9}$$

where

$$\frac{\Phi}{1-\gamma} = \frac{[b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} - [u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}}}{1-\gamma} > 0$$

and that $\frac{dU_{2i}}{dK}$ decreases in θ_i.

Differentiating (B.5) with respect to K yields

$$\frac{d\Delta}{dK} = \sum_{i=1}^{n} p_i^H \frac{1}{W_{11}^2} \left[\underbrace{\left(\frac{dU_{21}}{dK} - \frac{dU_{2i}}{dK} \right)}_{>0} W_{11} - \underbrace{\frac{dW_{11}}{dK}}_{<0} \underbrace{(U_{21} - U_{2i})}_{>0} \right] > 0.$$

Region 2: $K \geq \theta_1$

Recall that when $K \geq \theta_1$, U_{21} the high-output worker faces no uncertainty about promotions in the next period:

$$U_{21} = [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}},$$

which implies that $\frac{dU_{21}}{dK} = 0$. Equation (B.8) implies that $\frac{dW_{11}}{dK} = 0$.

Differentiating (B.5) with respect to K yields

$$\frac{d\Delta}{dK} = \sum_{i=1}^{n} p_i^H \frac{1}{W_{11}^2} \left[\left(\underbrace{\frac{dU_{21}}{dK}}_{=0} - \underbrace{\frac{dU_{2i}}{dK}}_{\geq 0} \right) W_{11} - \underbrace{\frac{dW_{11}}{dK}}_{=0} \underbrace{(U_{21} - U_{2i})}_{>0} \right] < 0$$

because $\frac{dU_{2i}}{dK} < 0$ for i such that $\theta_i > K$.

Proof of Proposition 4. This proof adopts the notations from the proof of Proposition 3. Differentiating (B.7) with respect to γ yields

$$\frac{dW_{11}}{d\gamma} = -\frac{dU_{21}}{d\gamma}.$$

Recall that when $K \geq \theta_1$, the high-output worker faces no uncertainty about promotions in the next period:

$$U_{21} = [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1}{1-\frac{1}{\psi}}}.$$

Hence, $\frac{dU_{21}}{d\gamma} = 0$. Moreover, for all i such that $\theta_i > K$, U_{2i} can be expressed as

$$U_{2i} = \left(\frac{K}{\theta_i} [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} + (1 - \frac{K}{\theta_i}) [u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1}{1-\gamma}},$$

which implies that $\frac{dU_{2i}}{d\gamma} < 0$.

Differentiating (B.5) with respect to γ yields

$$\frac{d\Delta}{d\gamma} = \sum_{i=1}^{n} p_i^H \frac{1}{W_{11}^2} \left[\left(\underbrace{\frac{dU_{21}}{d\gamma}}_{=0} - \underbrace{\frac{dU_{2i}}{d\gamma}}_{\leq 0} \right) W_{11} - \underbrace{\frac{dW_{11}}{d\gamma}}_{=0} \underbrace{(U_{21} - U_{2i})}_{>0} \right] > 0$$

because $\frac{dU_{2i}}{d\gamma} < 0$ for i such that $\theta_i > K$.

Proof of Proposition 5 This proof adopts the notations from the proof of Proposition 3. Let $\hat{\theta}$ be a random variable that also takes values θ_1, θ_2, ..., θ_n ($n \geq 4$), but with probabilities $\hat{p}_1$, $\hat{p}_2$, ..., $\hat{p}_n$, respectively. Let $\hat{p}_i^H$ be a worker's posterior that $\hat{\theta}$ is θ_i given that the worker is high-type: $\hat{p}_i^H = Prob(\hat{\theta} = \theta_i | \eta = 1)$. Define $R_i = \frac{W_{1i} - W_{11}}{W_{11}}$. Recall that in the limit as $\frac{1}{\psi}$ goes to 0, R_i can be expressed as

$$R_i = \frac{U_{21} - U_{2i}}{b_{1H} + u_2 + U_3 - U_{21}}.$$

Note that θ and $\hat{\theta}$ have the same support. Hence, in the limiting case as $\frac{1}{\psi}$ goes to 0, R_i is the same under both distributions. Pay sensitivity under θ can be expressed as

$$\Delta = \sum_{i=1}^{n} p_i^H R_i.$$

Pay sensitivity under $\hat{\theta}$ can be expressed as

$$\hat{\Delta} = \sum_{i=1}^{n} \hat{p}_i^H R_i.$$

Let $\hat{\theta}$ be a mean-preserving spread of θ according to the formulation of Rothschild and Stiglitz (1970). Then, $\hat{p}_i = p_i$ for all but four i's, denoted i_1, i_2, i_3, and i_4 where $i_j < i_{j+1}$. The quantities $\hat{p}_{i_j}$, for $j \in \{1, 2, 3, 4\}$, satisfy

$$\epsilon_j = \hat{p}_{i_j} - p_{i_j}, \tag{B.10}$$

$$\epsilon_1 = -\epsilon_2 > 0, \tag{B.11}$$

and

$$\epsilon_4 = -\epsilon_3 > 0. \tag{B.12}$$

Condition (B.10), (B.11), and (B.12) imply that the distribution of $\hat{\theta}$ has thicker tails than that of θ. Moreover, the quantities $\hat{p}_{i_j}$, for $j \in \{1, 2, 3, 4\}$, satisfy

$$\sum_{j=1}^{4} \theta_{i_j}(\hat{p}_{i_j} - p_{i_j}) = 0, \tag{B.13}$$

which implies that θ and $\hat{\theta}$ have the same mean:

$$\sum_{i=1}^{n} \hat{p}_i\theta_i = \sum_{i=1}^{n} p_i\theta_i. \tag{B.14}$$

By construction,

$$\hat{p}_{i_1}\theta_{i_1} + \hat{p}_{i_2}\theta_{i_2} = p_{i_1}\theta_{i_1} + p_{i_2}\theta_{i_2} + \underbrace{\epsilon_1(\theta_{i_1} - \theta_{i_2})}_{< 0 \text{ because } \theta_{i_1} < \theta_{i_2}} < p_{i_1}\theta_{i_1} + p_{i_2}\theta_{i_2},$$

which implies that

$$\sum_{i=1}^{i_2} \hat{p}_i\theta_i < \sum_{i=1}^{i_2} p_i\theta_i \tag{B.15}$$

because $\hat{p}_i = p_i$ for all $i \notin \{i_1, i_2, i_3, i_4\}$. (B.15) and (B.14) together imply that

$$\sum_{i=i_2+1}^{n} \hat{p}_i\theta_i > \sum_{i=i_2+1}^{n} p_i\theta_i. \tag{B.16}$$

Note that by construction,

$$\hat{p}_{i_3}\theta_{i_3} = p_{i_3}\theta_{i_3} + \underbrace{\epsilon_3\theta_{i_3}}_{<0} < p_{i_3}\theta_{i_3}$$

and $\hat{p}_i = p_i$ for all $i \notin \{i_1, i_2, i_3, i_4\}$. Hence,

$$\sum_{i=j}^{n} \hat{p}_i\theta_i \geq \sum_{i=j}^{n} p_i\theta_i \text{ for all } j > i_2. \tag{B.17}$$

with strict inequality if j is also strictly less than i_4. Let μ be the mean of θ and $\hat{\theta}$:

$$\mu = \sum_{i=1}^{n} \hat{p}_i \theta_i = \sum_{i=1}^{n} p_i \theta_i.$$

Then,

$$\sum_{i=j}^{n} \underbrace{\frac{\hat{p}_i \theta_i}{\mu}}_{=\hat{p}_i^H} \geq \sum_{i=j}^{n} \underbrace{\frac{p_i \theta_i}{\mu}}_{=p_i^H} \text{ for all } j > i_2 \tag{B.18}$$

with strict inequality if j is also less than i_4. Suppose $K \geq \theta_{i_2}$, then $R_i = 0$ for all $i \leq i_2$. Proposition 1 implies that R_i strictly increases in i for $i > k \geq i_2$, where k is the largest integer such that $\theta_k \leq K$. Hence, (B.18) implies

$$\hat{\Delta} = \sum_{i=i_2}^{n} \hat{p}_i^H R_i \geq \sum_{i=i_2}^{n} p_i^H R_i = \Delta,$$

with strict inequality if K is also less than θ_{i_4}. Let $\hat{K} = \theta_{i_2} < \theta_{i_3} < \theta_{i_4}$. By construction, i_2 can be no bigger than i_{n-2}. Hence, $\hat{K} \leq \theta_{n-2}$ $(n \geq 4)$ for all mean-preserving spreads.

Proof of Proposition 6. Let $K(Q_1)$ be the number of promotions available in the second period as a function of the firm outcome in the first period. The crux of the argument behind Proposition 1's firm performance pay result is that the $t = 1$ certainty equivalent lifetime utility of a high-output worker declines monotonically in the firm outcome after a threshold, i.e., $\bar{U}_2(\theta_i) > \bar{U}_2(\theta_j)$ for all $\theta_i > \theta_j > \theta_k$. Recall that the $t = 1$ certainty equivalent of the high-output worker's $t = 2$ lifetime utility when $\theta > \theta_k$ is:

$$\bar{U}_2(\theta) = \left(\tfrac{K(\theta)}{\theta} [b_{2H}^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} + \left(1 - \tfrac{K(\theta)}{\theta}\right) [u_2^{1-\frac{1}{\psi}} + U_3^{1-\frac{1}{\psi}}]^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \right)^{\frac{1}{1-\gamma}}.$$

Observe that $\bar{U}_2(\theta)$ strictly increases in $\frac{K(\theta)}{\theta}$ because $b_{2H} > u_2$. Hence, $\bar{U}_2(\theta_i) < \bar{U}_2(\theta_j)$ for all $\theta_i > \theta_j > \theta_k$ if and only if

$$\frac{K(\theta_i)}{\theta_i} < \frac{K(\theta_j)}{\theta_j} \Leftrightarrow \frac{K(\theta_i)}{K(\theta_j)} < \frac{\theta_i}{\theta_j} = \frac{Q_{1i} - x_L}{Q_{1j} - x_L}$$

for all $\theta_i > \theta_j > \theta_k$.