Incentive Design for Talent Discovery

Erik Madsen[*] Basil Williams[†] Andrzej Skrzypacz[‡]

April 25, 2022

Abstract

We study how career concerns within an organization distort employee risk-taking. When employees act to maximize their chances of promotion, aggregate risk-taking can be either too high or too low. Their choices can be influenced through incentive schemes which pay bonuses and/or reallocate promotions between groups of employees. We show that the optimal incentive tool depends on the desired power of incentives, with low-powered incentives optimally provisioned through bonuses while high-powered incentives are achieved by reallocating promotions. When asymmetric schemes are possible, the organization may further benefit from dividing employees into multiple groups and incentivizing different rates of risk-taking in each group.

Keywords: Incentive pay, promotion policies, career concerns, risk-taking

1 Introduction

The prospect of promotion is an important source of employee incentives in organizations.[1] Existing theoretical work, for instance on tournaments (Green and Stokey 1983; Lazear and Rosen 1981; Nalebuff and Stiglitz 1983; Rosen 1986), has focused on promotions as a source of incentives for effort. However, in many jobs employees decide not just how hard to work, but also how to allocate their time between alternative tasks or projects. These decisions loom large in jobs with significant autonomy, for instance management, software engineering,

[*]Department of Economics, New York University. Email: emadsen@nyu.edu

[†]Department of Economics, New York University. Email: basil.williams@nyu.edu

[‡]Stanford Graduate School of Business, Stanford University. Email: skrz@stanford.edu

[1]The seminal review of organizational incentives by Baker, Jensen, and Murphy (1988) observes that "Promotions are used as the *primary* incentive device in most organizations, including corporations, partnerships, and universities" (emphasis added).

and research positions, which are increasingly important in the modern economy due to rising research and development spending (Bloom et al. 2020) and a labor market transition toward non-routine jobs requiring creative problem-solving and interpersonal skills (Autor, Levy, and Murnane 2003; Levy and Murnane 2004). Employees in autonomous roles tend to possess strong intrinsic motivation to work hard (Hackman and Oldham 1980), and what they work on is perhaps of greater concern than how hard they work. This paper studies how the prospect of promotion influences what employees work on, and how incentive schemes can be used to guide these choices.

We focus on a natural benchmark environment in which, absent the prospect of promotion or monetary incentives, employees are indifferent between alternative uses of their time. Their preferences become misaligned when the organization attempts to solve a basic selection problem: It needs to fill a set of vacant senior positions by promoting a subset of its employees, and it relies on job performance as a signal of suitability for these positions. Allocating promotions based on performance distorts incentives, because employees can manipulate the variability of their performance at a cost to expected performance.

To model this distortion parsimoniously, we suppose that employees spend their time either on a routine "safe" task or an experimental "risky" task. These tasks could represent, for instance, alternative approaches to completing an objective or mutually exclusive projects. Performance on the risky task is more variable than on the safe task, but is not necessarily higher in expectation. Rather, (expected) relative performance varies across employees and is privately observed, so that the organization cannot achieve efficiency by simply assigning employees to tasks.

As a preliminary result, we show that the "natural" incentives generated by promoting high-performing employees distort task choices. Depending on the scarcity of promotions, some employees either take undesirable risks in order to achieve an impressive outcome (if promotions are scarce), or else stick to unprofitable safe tasks in order to hedge against a poor outcome (if promotions are plentiful). To mitigate these distortions, the organization can commit to an incentive scheme which promotes underperforming employees, pays outcome-contingent monetary bonuses, or some combination.[2] While promotions and bonuses are interchangeable as rewards to employees, they impose distinctive costs on the organization—

[2]We assume that the organization cannot directly reduce the value of a promotion below some baseline level by shrinking the position's salary or eliminating perks. This assumption is consistent with the market-signaling theory of promotions developed by Waldman (1984) and Bernhardt (1995), in which promotions signal talent to an external market and unavoidably increase the employee's wage through competition. It also captures social status or empire-building rewards associated with promotion which cannot be controlled by the organization.

reallocating promotions reduces surplus by hampering selection, while paying bonuses cedes rents to employees. The organization therefore faces a nontrivial tradeoff between the two tools.

We first analyze symmetric incentive schemes, which treat employees who choose the same task and achieve the same outcome equivalently. Symmetric schemes are an important benchmark class because they respect fairness or equity concerns which are salient in many real-world organizations. Symmetric schemes also serve as a key building block for constructing more general asymmetric schemes.

Our first main result characterizes the optimal symmetric incentive scheme implementing a given risk-taking rate across the workforce. The optimal scheme either pays bonuses to underperforming employees or reallocates promotions from outperforming to underperforming ones, but not both. The choice between the two tools hinges on the scheme's desired incentive power, i.e., how far the target risk-taking rate is from the natural rate. Bonuses turn out to be optimal for providing low-powered incentives, while promotion reallocations are better for providing high-powered incentives. One feature of note is that when the organization stimulates increased risk-taking through bonuses, payments are made specifically for failure on risky tasks, a result reminiscent of the finding in Manso (2011) that experimentation may be optimally motivated through failure bonuses.[3]

This result suggests distinctive lessons for incentivizing employees in fast- and slow-growing organizations. In a fast-growing organization, opportunities for promotion are plentiful and employees will shy away from risk-taking to maximize their chance of being promoted. In these environments, organizations should reward failure. Meanwhile in a slow-growing organization, promotions are scarce and employees will take excessive risks in order to earn a promotion. In these environments, organizations should reward employees who take on routine or unglamorous tasks.

Our second main result connects the optimal power of incentives to structural features of the internal labor market. We show that when the organization's selection concern is large or the value of promotion to employees is small, incentives are optimally low-powered and the organization incentivizes with bonuses. Conversely, when selection is relatively unimportant or employees place a high premium on promotion, incentives are optimally high-powered and the organization incentivizes by reallocating promotions. These comparative statics link the structure of optimal incentives to potentially measurable quantities. For instance, the correlation between performance in lower- and higher-level roles in order could be measured

[3]In Manso (2011), experimentation must be incentivized because it incurs an effort cost. In our setting, employees' cost of risk-taking is a reduced chance of receiving a valuable promotion.

to quantify the importance of selection to the organization,[4] while measures of labor market mobility could proxy for the reward associated with promotion.[5]

Our final results concern asymmetric incentive schemes which discriminate between observably identical employees. While fairness concerns often limit such discrimination, organizations can sometimes circumvent them by erecting social barriers between firm divisions, for instance through geographic separation between offices housing different teams. We show that when such barriers are possible, the organization may benefit from splitting employees into up to (but no more than) two groups and offering different incentive schemes to each group. A key property of optimal asymmetric schemes is that they do not reallocate promotions from high- to low-performers within any group. Instead, any inefficiency in the allocation of promotions arises due to the allotment of promotions between groups.

This analysis implies that, when discrimination between equivalent employees is possible, it is preferable to reallocate promotions "ex ante" rather than "ex post". That is, employees should be informed about whether they will be favored or disfavored for promotion before rather than after they have chosen a task. In practice, our results suggest that organizations seeking to incentivize diverse task choices may benefit from siloing groups of employees into culturally distinctive divisions, accommodating distinct incentive schemes and patterns of task choice.

The remainder of the paper is organized as follows. In Section 1.1, we discuss related literature. In Section 2 we set up the model. In Section 3 we describe the basic incentive problem faced by the organization. In section 4 we characterize the optimal symmetric incentive scheme targeting a given level of risk-taking. In section 5 we characterize the optimal level of risk-taking as a function of structural features of the internal labor market. Section 6 extends our analysis to asymmetric schemes. We offer concluding remarks in section 7. All proofs are collected in the Appendix.

1.1 Related literature

Our paper contributes to a literature studying how reputational concerns distort an employee's willingness to take risks. Holmström (1999) (section 3) analyzes this distortion in a career concerns setting where the employee's compensation is determined entirely by an external market's perception of their quality. Holmstrom and Ricart I Costa (1986), Zwiebel (1995), Hvide and Kaplan (2005), and Siemsen (2008) have built on the career-concerns

[4]Mean-reversion of performance following promotion has been widely recognized in the literature on the Peter principle, both theoretically (Lazear 2004) and empirically (Benson, Li, and Shue 2019).

[5]As suggested by the model of Waldman (1984) and Bernhardt (1995), promotions may serve as a public signal of talent and therefore boost an employee's wage through competition between potential employers.

framework to model incentive contracting for risk-taking in the shadow of career concerns.[6] In all of these papers, returns to reputation are exogenous and must be offset through incentive pay or restrictions on the employee's freedom to take risks. In our model, by contrast, the employer can directly control the employee's reputational concerns by committing to a promotion policy.

Several papers study environments in which an employee is concerned with the perceptions of their employer rather than the broader market. In Kuvalekar and Lipnowski (2020) and Kostadinov and Kuvalekar (2022), the employer prefers to separate from low-quality employees and cannot commit to a termination policy, generating distortionary returns to reputation similar to career concerns. In Aghion and Jackson (2016) the employer can commit to a replacement policy, allowing it to control the employee's returns to reputation; however, no incentive payments are allowed. In our model the employer can commit to both promotion policies and incentive payments, allowing us to compare the two tools as mechanisms for incentivizing efficient risk-taking.

Bar-Isaac and Lévy (2022) study a related risk-taking environment in which an employer motivates hidden effort by committing to make employees visible on an external labor market which can bid up their wage. Thus unlike in our model, career concerns are the solution to an incentive problem rather than the source of one.[7] However, career concerns are also distortionary because the employer cannot commit to assigning employees a particular task, leading it to distort risk-taking in an attempt to hold down future wages. A common theme of our analysis and theirs is that when an organization can design the reputational concerns of its employees, its choice has important consequences for employee risk-taking.

Our paper also contributes to a discussion regarding the relative merits of promotions and money as incentive tools. Baker, Jensen, and Murphy (1988) pose a now-classic empirical puzzle: Rewarding employees with promotions degrades selection and so is less efficient than incentivizing with money, and yet performance pay is rarely observed in practice.[8] Lemieux,

[6] A related strand of the literature studies incentives for effort under career concerns; see, e.g., Holmström (1999) (section 2), Dewatripont, Jewitt, and Tirole (1999b), and Gibbons and Murphy (1992). Most closely related is Kaarbøe and Olsen (2006), which studies incentive contracts in a multitask setting. Multitasking generates a tension between the productivity and reputational impact of a given effort allocation, as in models of risk-taking. However, this tradeoff exists alongside additional distortions caused by the potential for shirking and non-contractibility of performance signals. Studies of risk-taking, including ours, abstract from these issues to focus on the productivity-reputation tradeoff.

[7] Their model shares this feature with Dewatripont, Jewitt, and Tirole (1999a) and Hörner and Lambert (2021), who analyze how the set of performance signals available to the market affects incentives for hidden effort under career concerns.

[8] This puzzle is also posed in the classic management textbook of Milgrom and Roberts (1992) (pg. 366-367).

MacLeod, and Parent (2009) have more recently found an uptick in the use of performance pay, but they still estimate that around 60% of private-sector employees do not receive any variable pay.[9] Standard solutions to this puzzle are psychological: Performance pay could crowd out intrinsic incentives, or organizational morale might be degraded in the presence of steep pay differentials across employees with comparable responsibilities. (See Baker, Jensen, and Murphy (1988), section I.A for an overview of these solutions.) More recently, several economic explanations have been proposed.

One theory highlights the importance of influence activities (in the sense of Milgrom and Roberts (1988)). Fairburn and Malcomson (2001) demonstrate this mechanism in a model in which employees can bribe managers to distort subjective performance reviews. In their setting, incentivizing through promotions reduces the susceptibility of managers to influence due to their stake in the firm's future performance.[10] Another theory emphasizes talent signaling. Schöttner and Thiele (2010) illustrate this possibility in a model in which higher-quality employees value promotion more. In their model, incentive pay compresses effort differentials between high- and low-quality employees, degrading the informativeness of performance as a signal of talent. Our results provide an alternative rationale for incentivizing with promotions in multitask environments: Task choices can be cheaply influenced by reducing the probability of promotion associated with tasks that few employees end up choosing.

2 The model

We build a stylized model of organizational task choice in which employees' risk-taking decisions are distorted by the prospect of promotion. In our model, an organization oversees a unit mass of employees with whom it interacts over two stages. In the first stage, each employee chooses between two tasks which differ in their expected productivity, risk level, and informativeness about talent. In the second stage, the organization observes task choices and outcomes, selects a set of employees to promote, and (potentially) pays monetary bonuses.

Employees have heterogeneous but initially unknown talent, and the organization wishes to allocate promotions to the most-talented employees, as revealed by their task performance. At the same time, employees possess private information about their optimal task, which the organization wishes to incentivize them to use optimally. The tension between selection

[9]This figure is likely an underestimate of the infrequency of true performance pay, since variable pay includes bonuses tied to factors other than individual achievement, such as team, division, or companywide performance.

[10]This mechanism is also proposed informally in Milgrom and Roberts (1992) (pg. 370).

and efficient task choice lies at the heart of our model. We now discuss each of these model elements in more detail.

Types. Employees are indexed by $n \in [0,1]$ and are heterogeneous across two dimensions, summarized by a type $(\theta(n), \Gamma(n))$. The two dimensions of an employee's type capture two distinct sources of uncertainty. The employee's *quality* θ summarizes the employee's general competence, which affects both his performance in his current position and his suitability for promotion. Meanwhile, his *match type* Γ summarizes information determining his most productive task.

Quality is symmetrically unobserved by the organization and the employee and must be inferred by observing employee performance. We assume that quality takes one of two real values: $\theta(n) \in \{\overline{\theta}, \underline{\theta}\}$, where $0 \leq \underline{\theta} < \overline{\theta}$. We will refer to employees of types $\overline{\theta}$ and $\underline{\theta}$ as "high-quality" and "low-quality," respectively. Employee qualities are independently and identically distributed, with $\Pr(\theta(n) = \overline{\theta}) = \pi_0 \in (0,1)$. Without loss, we normalize quality levels so that $\pi_0\overline{\theta} + (1-\pi_0)\underline{\theta} = 1$.

Match types are privately observed by each employee. We assume that they are independently and identically distributed across employees and that qualities and match types are independent: $\theta(n) \perp\!\!\!\perp \Gamma(n')$ for all employees n, n'. Independence of quality and match type and unobservability of quality implies that employees cannot directly signal their quality through their task choice.

Without loss of generality, we assume that employees are indexed in decreasing order of match type, allowing us to summarize the distribution of match types by a function $\gamma(n)$ indicating the match type of the nth employee.[11] For simplicity, we impose the mild regularity conditions that the distribution of match types has a strictly positive density on its support, and in particular has no gaps or atoms:

Assumption 1. γ is C^1 and $\gamma'(n) < 0$ for all $n \in [0,1]$.

First Stage. In the first stage, each employee chooses to complete either a *safe* task or a *risky* task. These tasks are specific to a particular employee and may capture distinct activities for employees with different assignments (so that the outcomes of tasks are independent across employees). To focus on a novel set of agency frictions, we assume that the employee has the same private cost/benefit of performing either task and normalize that cost/benefit to zero. Absent incentives from bonuses or promotions, the employees are indifferent between the two tasks.

[11] Under this convention, an employee's index is privately observed.

The two tasks differ in their expected productivity and the variability of their outcome. The safe task produces a sure payoff of $K \in (0,1)$ to the organization. We refer to this outcome as "neutral".[12] By contrast, the risky task produces a payoff of either 1 or 0 for the organization, outcomes which we refer to as "success" and "failure". The risky task succeeds with probability

$$q(\theta,\Gamma) = \theta \cdot \Gamma.$$

That is, the probability of success is increasing in both an employee's quality and the match type, so that the results of the risky task are informative about the employee's quality. Further, since θ and Γ are independent and θ has mean 1, Γ is the employee's perceived probability of success on the risky task.[13] To ensure non-trivial incentive problems, we assume that the task which maximizes the organization's expected payoff varies across employees:

Assumption 2. $\gamma(0) > K > \gamma(1)$.

This assumption implies that the organization's first-stage payoff is maximized when employees with index $n \leq N^0$ choose the risky task, where $N^0 \in (0,1)$ is the unique solution to $\gamma(N^0) = K$.

Second Stage. The organization has a mass of unfilled positions of measure $\beta \in (0,1)$ into which employees can be promoted in the second stage. All positions are identical, and a position can be assigned to at most one employee. Promoting an employee of quality θ generates a payoff of $r(\theta)$ to the organization, with $r(\overline{\theta}) > r(\underline{\theta})$. Meanwhile, a promoted employee enjoys a private benefit of $V > 0$ regardless of quality. An unfilled position generates a payoff to the organization, which we normalize to 0, and an employee who is not promoted receives a private benefit of 0.

We assume that $r(\overline{\theta}) > 0$, so that the organization benefits from promoting high-quality employees. We further focus on settings in which $r(\underline{\theta})$ is not too negative, so that the organization prefers to allocate all available positions, even when that means promoting some underperforming employees.[14] When the organization optimally allocates all positions,

[12]Nothing would change if the safe task generated a stochastic output, so long as the outcome was uninformative about the employee's quality.

[13]To ensure that $q(\theta,\Gamma) \in [0,1]$ for all θ and Γ, we must have $\gamma(0) \leq 1/\overline{\theta}$. This upper bound is simply a normalization allowing Γ to be interpreted as a success probability.

[14]More concretely, the organization should be willing to promote employees who failed at a risky task. The lowest possible posterior belief the organization could hold about such an employee's quality is $\underline{\pi}_B = (1-\overline{\theta}\gamma(0))\pi_0/\gamma(0)$, and so a sufficient condition ensuring that all positions are allocated is $\underline{\pi}_B r(\overline{\theta}) + (1-\underline{\pi}_B) r(\underline{\theta}) > 0$.

the only payoff variable which matters for designing an incentive scheme is the difference $r(\overline{\theta}) - r(\underline{\theta})$. To economize on notation, we set $r(\overline{\theta}) = R > 0$ and $r(\underline{\theta}) = 0$ going forward. The single parameter R captures the magnitude of the organization's selection concern.

In addition to allocating promotions, the organization can pay monetary bonuses to employees. We assume that employees enjoy limited liability, so that bonuses cannot be negative.[15] The organization and all employees have utility functions that are quasilinear in money, and we assume that V, R, and all task payoffs are normalized so that they are denominated in dollars.

3 The incentive problem

We begin our analysis by showing that, in the absence of an incentive scheme, the amount of risk-taking by employees is generally suboptimal. Section 3.1 characterizes the risk-taking rate which prevails when the organization promotes employees in order of perceived quality. Section 3.2 establishes that this rate is too high when promotions are scarce and too low when they are plentiful, relative to the rate the organization would implement if it were not constrained by employees' desire for promotion. Section 5.2 proves that an optimal incentive scheme shifts risk-taking toward the unconstrained optimal rate, setting the stage for our characterization of an optimal scheme.

3.1 The no-commitment outcome

Absent commitment to an incentive scheme, the organization pays no bonuses and promotes employees in descending order of perceived quality. Since successful risk-taking is good news about quality while failure is bad news, the organization first promotes successful risk-takers, followed by risk-avoiders, and finally resorts to promoting failed risk-takers until all promotions are allocated. (Because promoting a low-quality employee yields a payoff at least as high as leaving a spot unfilled, no promotions are withheld.) This prioritization rule leaves two degrees of freedom for an optimal promotion policy. First, the organization could break ties between observationally equivalent employees in different ways. Second, the organization could promote a measure-zero subset of employees "out-of-order".

We focus on the unique optimal policy which 1) treats all observationally equivalent employees equally, i.e., promotes uniformly at random from among employees of equal perceived

[15] If promoted employees could be charged for the privilege, the organization could costlessly resolve the underlying incentive problem by extracting all rents from promotion. Limited liability therefore reflects unpledgeability of the employee's promotion payoff.

quality in case promotions need to be rationed; and 2) promotes employees strictly in order of perceived quality.[16] We refer to this allocation rule as the *natural promotion policy.* The symmetry property of the natural policy is substantive, but constitutes a natural benchmark which would be selected in organizations for which fairness is an important constraint. We will relax it when we consider asymmetric incentive schemes in Section 6. The strict prioritization property rules out spurious outcomes which would not arise in a model with a finite set of employees.[17]

We now characterize the risk-taking behavior which prevails under the natural promotion policy.

Proposition 1. Fix all model parameters except for β. For every $\beta \in (0,1)$, there exists an essentially unique[18] set of employees $\mathcal{N}^{nc}$ who select the risky task under the natural promotion policy. There exist cutoffs $\underline{\beta}$ and $\overline{\beta}$, satisfying $0 < \underline{\beta} < \overline{\beta} < 1$, such that:

- If $\beta < \underline{\beta}$, then $\mathcal{N}^{nc} = [0,1]$,
- If $\beta \in [\underline{\beta}, \overline{\beta}]$, then $\mathcal{N}^{nc} = [0, N^{nc}]$, where N^{nc} is continuous and decreasing in β and satisfies $N^{nc} = 1$ when $\beta = \underline{\beta}$ and $N^{nc} = 0$ when $\beta = \overline{\beta}$.
- If $\beta > \overline{\beta}$, then $\mathcal{N}^{nc} = \emptyset$.

To unify notation, we extend N^{nc} to all $\beta \in (0,1)$ by defining $N^{nc} = 1$ for $\beta < \underline{\beta}$ and $N^{nc} = 0$ for $\beta > \overline{\beta}$. This threshold has the property that $\mathcal{N}^{nc} = [0, N^{nc}]$ for all $\beta \leq \overline{\beta}$ and $\mathcal{N}^{nc} = (0, N^{nc})$ for $\beta > \overline{\beta}$. We refer to N^{nc} as the *natural* or *no-commitment* risk-taking rate, and illustrate it in Figure 1. Recall that lower-indexed employees are those who are best-matched to the risky task. Proposition 1 therefore establishes that only the best-matched employees select the risky-task, and that the natural risk-taking rate declines as the number of promotions increases.

A threshold risk-taking structure arises because an employee's probability of successful risk-taking rises with their match type, and success is rewarded while failure is penalized under the natural promotion policy. As a result, employees with the best match types face the largest upsides and smallest downsides from risk-taking. The drop in risk-taking as β

[16]This policy is unique given a set of posterior beliefs, which are pinned down by Bayes' rule whenever a positive measure of employees choose the risky task. If Bayes' rule does not apply, we assume the organization views success as a positive signal about θ and failure as a negative signal. Such posteriors would result, for instance, if the organization assumed that the best-matched employee(s) chose the risky task.

[17]For instance, there always exists an outcome with no risk-taking supported by a policy of declining to promote successful risk-takers whenever the set of such employees has measure zero. If employees were discrete, it would never be optimal for the organization to avoid promoting a lone successful risk-taker.

[18]This set is unique except possibly up to a single marginal employee.

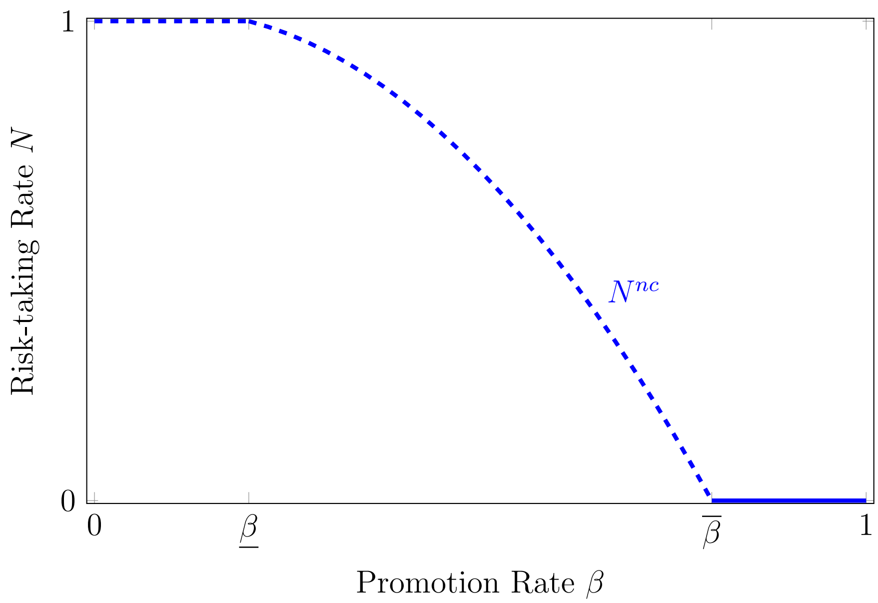

Figure 1: The natural risk-taking rate N^{nc} as a function of the measure of promotions β.

rises stems from the fact that each employee's incentive to take risks weakens as the number of promotions grows, holding fixed the task choices of all other employees. Intuitively, the bar for promotion drops as more employees are promoted, and so a given employee gains less from successful risk-taking and loses more from failure. This force pushes fewer employees to take risks as β increases.

3.2 Suboptimality of the natural risk-taking rate

We next establish that the natural risk-taking rate is generally suboptimal from the organization's perspective. That is, it deviates from the rate the organization would choose if employees did not value promotion (i.e., if $V = 0$) and would choose any task recommended by the organization.[19] Let N^{fb} be the *first-best* risk-taking rate the organization would choose in the absence of incentive constraints.[20] Note that N^{fb} need not equal N^0, the risk-taking rate maximizing the organization's first-stage payoff, since an optimal risk-taking rate balances task payoffs and talent discovery.

We now show that the natural risk-taking rate is higher than the first-best rate when β is small and lower when β is large.

[19] Lemma 1, proven below, implies that the organization optimally recommends a threshold risk-taking rule in this scenario.

[20] N^{fb} is not guaranteed to be unique. In case of non-uniqueness, Proposition 2 holds for every maximizer.

Proposition 2. There exists a $\beta^{**} \in [\underline{\beta}, \overline{\beta})$ such that

$$N^{nc}(\beta) \begin{cases} > N^{fb}(\beta), & \beta < \beta^{**} \\ = N^{fb}(\beta), & \beta = \beta^{**} \\ < N^{fb}(\beta), & \beta > \beta^{**} \end{cases}$$

One force driving this result is the simple fact that first stage payoffs are maximized by an interior amount of risk-taking (see Assumption 2). Since N^{nc} is declining in β, this force points in the direction of too much risk-taking for small β and too little for large β. A second force related to talent discovery reinforces this trend. Roughly, for small β only a small number of high-quality employees are needed to maximize promotion payoffs, and so the organization does not benefit from raising N to learn more about employees. As β increases, it becomes important to distinguish high- from low-quality employees on a large scale, and so N^{fb} is boosted by a talent discovery motive. This force also leads to too much risk-taking for small β and too little for large β.

4 Incentive schemes

Our main results concern how the distortions identified in Section 3 can be mitigated through commitment to an *incentive scheme* that specifies bonuses and promotion probabilities as a function of task choice and outcome and recommends how employees should use their private information about Γ to choose a task. We impose the following requirements on an incentive scheme:

- *Feasibility*: At most β employees are promoted.
- *Limited liability*: Every employee receives a non-negative bonus.
- *Determinism*: Aggregate promotions and transfers are non-random.
- *Symmetry*: All employees are given the same recommendation (as a function of Γ), and employees with identical task choices and outcomes are treated equally.

Feasibility and limited liability impose the constraints discussed in the model setup. Determinism focuses on schemes which are non-random "in aggregate", so that within each group of observationally identical employees, the total number of employees promoted and total transfers are deterministic. Finally, symmetry rules out schemes in which employees are split into groups and incentivized to make different task choices through different rewards for particular outcomes. We view symmetry as a realistic and important constraint in light of fairness concerns that limit unequal treatment of similar employees in many contexts. Symmetric schemes will also serve as a key building block for our analysis of more general asymmetric schemes in Section 6.

In line with the requirements above, we formally define an incentive scheme as follows:

Definition 1. An *incentive scheme* is a triple $\mathcal{S} = (\mathcal{N}, \boldsymbol{T}, \boldsymbol{\sigma})$, where:

- $\mathcal{N} \subset [0,1]$ is the set of employees to whom the organization recommends the risky task.
- $\boldsymbol{T} = (T_G, T_0, T_B) \geq 0$ are the bonuses received by an employee who, respectively, achieves a successful, neutral, or failure outcome.
- $\boldsymbol{\sigma} = (\sigma_G, \sigma_0, \sigma_B) \in [0,1]^3$ are the probabilities of promotion for an employee who, respectively, achieves a successful, neutral, or failure outcome.

We interpret an incentive scheme as promoting employees uniformly at random from within each group of observationally identical employees. More precisely, suppose that $\mathcal{N}_G, \mathcal{N}_0, \mathcal{N}_B$ are the sets of employees who, respectively, achieve a successful, neutral, or failure outcome. Then for each $i \in \{G, 0, B\}$, a fraction σ_i of employees from group $\mathcal{N}_i$ is promoted. Under such a scheme, the measure of employees promoted from each group is non-random, respecting determinism. Of course, from the perspective of any individual employee in group $\mathcal{N}_i$, promotion is random whenever $\sigma_i \in (0,1)$, even conditioning on the outcome of their chosen task.

An incentive scheme is *feasible* if it promotes at most β employees, supposing employees choose the task recommended to them.[21] This requirement is summarized by the inequality

$$\beta \geq \int_{\mathcal{N}} (\gamma(n)\sigma_G + (1-\gamma(n))\sigma_B)\, dn + (1 - |\mathcal{N}|)\sigma_0.$$

It is *incentive-compatible* if all employees find it optimal to follow the scheme's risk-taking recommendation. That is,

$$\gamma(n)(T_G + V\sigma_G) + (1-\gamma(n))(T_B + V\sigma_B) \begin{cases} \geq T_0 + V\sigma_0, & \forall n \in \mathcal{N}, \\ \leq T_0 + V\sigma_0, & \forall n \in [0,1] \setminus \mathcal{N}. \end{cases} \tag{1}$$

[21] As employees are atomistic, any feasible incentive scheme remains feasible following a deviation by a single employee. Further, such deviations do not affect bonuses or promotion probabilities under a symmetric incentive scheme, which can condition only on the measure of outcomes of each type. The organization's choice of bonuses and promotion probabilities off-path therefore do not impact employee incentives, and we do not explicitly specify them.

We call an incentive scheme *admissible* if it is both feasible and incentive-compatible. Under any admissible incentive scheme $\mathcal{S}$, the organization achieves total profits equal to

$$\begin{aligned}\Pi(\mathcal{S}) \equiv & \int_{\mathcal{N}} \gamma(n)\, dn + (1 - |\mathcal{N}|)K \\ & + R\pi_0 \left(\int_{\mathcal{N}} (\overline{\theta}\gamma(n)\sigma_G + (1 - \overline{\theta}\gamma(n))\sigma_B)\, dn + (1 - |\mathcal{N}|)\sigma_0 \right) \\ & - \left(\int_{\mathcal{N}} (\gamma(n)T_G + (1 - \gamma(n))T_B)\, dn + (1 - |\mathcal{N}|)T_0 \right).\end{aligned}$$

In this expression, the first line accounts for task payoffs, the second for promotion payoffs, and the third for bonus payments.[22]

In general an admissible incentive scheme may recommend that an arbitrary subset of employees choose the risky task. The following lemma shows that attention may be restricted to incentive schemes with a cutoff risk-taking structure, under which only employees with the best match types choose the risky task.

Lemma 1. Fix any admissible incentive scheme $\mathcal{S} = (\mathcal{N}, \boldsymbol{T}, \boldsymbol{\sigma})$ satisfying $|\mathcal{N}| = N$. If $|\mathcal{N} \setminus [0, N]| > 0$, then there exists an admissible incentive scheme $\mathcal{S}' = ([0, N], \boldsymbol{T}', \boldsymbol{\sigma}')$ such that $\Pi(\mathcal{S}') > \Pi(\mathcal{S})$.

In light of this result, going forward we will describe an incentive scheme via a triple $\mathcal{S} = (N, \boldsymbol{T}, \boldsymbol{\sigma})$ for $N \in [0, 1]$, with the understanding that such a scheme recommends employees $n \in [0, N]$ choose the risky task.

5 Optimal incentive schemes

We now characterize the organization's optimal incentive scheme. Our analysis proceeds in two steps. In Section 5.1, we hold fixed a target risk-taking rate and characterize the optimal incentive scheme implementing the target. We then endogenize the risk-taking rate and identify the optimal direction of incentives in Section 5.2 and the optimal incentive tool in Section 5.3.

Our characterization reveals several key features of an optimal scheme. First, the tradeoff between money and promotions as incentive tools exhibits a bang-bang structure: An optimal scheme reallocates promotions or pays bonuses, but not both. Second, the optimal incentive tool depends on the desired power of incentives, as measured by the extremity of the risk-taking rate. Third, the optimal power of incentives depends critically on structural features

[22] Recall that we have normalized the value of promoting a low-quality employee to zero.

of the internal labor market, in particular employees' private value of promotion and the importance of selection to the organization.

Our analysis uncovers a key economic force linking the relative incentive power of bonuses and promotions to the risk-taking rate. While bonuses affect only the payoff of the incentivized group (i.e., the employees who switch tasks in response to the incentive scheme), a reallocation of promotions additionally reduces the payoff of the disincentivized group (those employees who don't switch tasks). The relative impact of the two tools on the incentivized group is independent of the risk-taking rate. By contrast, the impact of a reallocated promotion on the disincentivized group is controlled by the relative sizes of the two groups, which varies with the risk-taking rate. In particular, when the incentivized group is large relative to the disincentivized group, the extra incentive power of promotions is large and promotions are an optimal incentive tool.

5.1 Targeting a risk-taking rate

We now derive the optimal incentive scheme implementing an exogenous target risk-taking rate $N \neq N^{nc}$. Our main result is a linkage between the optimal incentive tool and the extremity of the target risk-taking rate. We prove that if $N < N^{nc}$ is sufficiently close to zero, or $N > N^{nc}$ is sufficiently close to 1, an optimal scheme reallocates promotions from high- to low-performers; otherwise, it promotes according to the natural policy and incentivizes by paying bonuses. In other words, promotions are better at providing "high-powered" incentives, while bonuses are superior when incentives need to be "low-powered".[23] We additionally show that when the organization wishes to increase risk-taking using bonuses, it optimally pays bonuses for failure rather than for success.

5.1.1 Decreasing risk-taking

We first consider target rates below the natural rate, i.e., $0 \leq N < N^{nc}$. (Of course, such targets are relevant only if the natural risk-rate is nonzero, or equivalently if $\beta < \overline{\beta}$.) To increase the fraction of employees choosing the safe task, the organization must increase its relative payoff by either reallocating promotions toward neutral outcomes, paying bonuses for neutral outcomes, or both.

[23] Depending on parameters, the low-powered regime may be degenerate, in which case promotions are the optimal incentive tool for all N. Formally, we prove a single-crossing result whereby any change in incentive tool as N moves away from the natural rate is always from bonuses toward promotions. We also derive conditions under which the low-powered regime is non-degenerate.

In principle, reallocated promotions might be drawn from either successful or failed risk-takers. However, in the absence of an incentive scheme either no employees take risks, or else the promotion probability following a failure outcome is zero. For if failed risk-takers were promoted with positive probability under the natural promotion policy, then employees taking the safe task would be promoted with probability 1, and no employees would choose the risky task. Since we assume that $0 < N^{nc}$, reallocated promotions must therefore be drawn from successes.

The following proposition characterizes an optimal incentive scheme as a function of the risk-taking target.

Proposition 3. Suppose that the organization implements a risk-taking rate $N < N^{nc}$. Then there exists a threshold $\overline{N}_- \in (0, 1]$ such that:

1. If $N \leq \overline{N}_-$, there exists an optimal scheme which pays no bonuses and reallocates promotions from successes toward neutral outcomes.

2. If $N \geq \overline{N}_-$, there exists an optimal scheme which pays a positive bonus following failure outcomes and does not reallocate promotions.

Further, the optimal scheme is unique whenever $N \neq 0, \overline{N}_-$. If R/V is sufficiently large, then $\overline{N}_- < N^{nc}$.

This result establishes several key properties of an optimal scheme. First, only one incentive tool is used at a time: The organization either pays bonuses or reallocates promotions, but never both in conjunction.[24] In the former case, we will say that the organization "incentivizes with bonuses", while in the latter we will say that it "incentivizes with promotions". Second, the optimal tool depends on the size of N, with the optimal tool switching from bonuses to promotions as N decreases. Third, there always exist risk-taking rates for which incentivizing with promotions is optimal; and if R/V is sufficiently large, there additionally exist risk-taking rates for which incentivizing with bonuses is optimal.[25]

The comparison between promotions and bonuses as incentive tools hinges on the *incentive power-per-dollar* (or IPD) of each tool. Each tool's IPD varies with N in a way which increasingly favors promotions as N drops, a result we now demonstrate heuristically. Suppose that the organization reallocates a measure m of promotions from successes to neutral outcomes. The corresponding shifts in the promotion rates for success and neutral outcomes

[24]The one exception is the edge case $N = \overline{N}_-$, in which case any combination of the two tools is optimal.

[25]The proof of the proposition additionally establishes that $\overline{N}_-$ is strictly decreasing in R/V for sufficiently large values of the ratio, and that $\lim_{R/V \to \infty} \overline{N}_- = 0$.

are

$$\Delta\sigma_G(m) = -\frac{m}{\mu(N)}, \quad \Delta\sigma_0(m) = \frac{m}{1-N},$$

where $\mu(N) = \int_0^N \gamma(n)\,dn$ is the measure of employees achieving the success outcome. Define the *incentive power* of this scheme to be the amount by which it increases the marginal employee's utility from choosing the safe task relative to the risky task. Then the total incentive power of a promotion reallocation is

$$V\left(\Delta\sigma_0(m) - \gamma(N)\Delta\sigma_G(m)\right) = \left(\frac{1}{1-N} + \frac{\gamma(N)}{\mu(N)}\right) Vm.$$

The cost to the organization of this promotion reallocation is $R(\pi_G - \pi_0)m$, where π_G is the organization's posterior belief about the quality of an employee who succeeded on the risky task. The IPD of promotions is therefore

$$IPD^{Pr}(N) = \left(\frac{1}{1-N} + \frac{\gamma(N)}{\mu(N)}\right) \frac{V/R}{\pi_G - \pi_0}.$$

Meanwhile, a scheme offering a bonus $t \geq 0$ for safe outcomes has total incentive power t and incurs a cost to the organization of $(1-N)t$. The IPD of bonuses is therefore

$$IPD^B(N) = \frac{1}{1-N}.$$

Note that the IPD for promotions does not depend on the number of promotions reallocated, and similarly the IPD for bonuses does not depend on the size of the bonus payment. Hence the optimal scheme will exhibit a bang-bang structure, using only the tool with the larger IPD.

The ratio of the two IPDs is

$$\frac{IPD^{Pr}(N)}{IPD^B(N)} = \frac{V/R}{\pi_G - \pi_0}\left(1 + \frac{\gamma(N)(1-N)}{\mu(N)}\right).$$

Since $\gamma'(N) < 0$ while $\mu'(N) = \gamma(N) > 0$, the IPD of promotions relative to bonuses declines in N, demonstrating why bonuses perform better for large N while promotions are preferrable for small N. In particular, in the limit $N \to 0$, $\mu(N) \to 0$ and the IPD of promotions relative to bonuses grows unboundedly. Thus for sufficiently small N, promotions are used in an optimal scheme. Meanwhile as $N \to N^{nc}$, the ratio of IPDs approaches a finite limit whose size is controlled by V/R. For V/R sufficiently small, the IPD of bonuses exceeds that of promotions for N close to N^{nc}, and bonuses are used in an optimal scheme.

They key distinction between bonuses and promotions as incentive tools is that bonuses impact the payoff only of employees who choose the incentivized task, while reallocated promotions additionally impact the payoff of employees who choose the disincentivized task.

To see this, suppose that the organization could generate additional promotions at a constant marginal cost $\bar{R} = R\pi_G$. The IPD of promotions generated this way is

$$\widehat{IPD}^{Pr}(N) = \frac{V/(\bar{R} - R\pi_0)}{1 - N},$$

resulting in a relative IPD of promotions versus bonuses equal to

$$\frac{\widehat{IPD}^{Pr}(N)}{IPD^B(N)} = \frac{V/R}{\pi_G - \pi_0},$$

which is independent of N.

Compared to this benchmark, a reallocated promotion generates extra incentive power by additionally reducing the payoff of employees choosing the risky task. The (relative) IPD generated by this force depends on the relative sizes of the pools of successful risk-takers and risk-avoiders. As N increases, the pool of successful risk-takers grows larger, diminishing the IPD from withholding promotions; meanwhile the pool of risk-avoiders shrinks, boosting the IPD from paying bonuses. These two trends combine to make promotions less favorable for large N.

5.1.2 Increasing risk-taking

We now turn to environments in which the organization wishes to encourage risk-taking beyond the natural rate, i.e., $1 \geq N > N^{nc}$.[26] To boost risk-taking, the organization must increase the relative payoff of taking risks by either reallocating promotions toward risk-takers, paying them bonuses, or both.

Both reallocated promotions and bonuses could in principle be targeted at either success or failure. However, in the absence of an incentive scheme either all employees take risks, or else the promotion probability following a success outcome is 1. For if successful risk-taking did not lead to sure promotion, then under the natural policy safe tasks would yield no chance of promotion. In that case no employees would choose the safe task, a contradiction of our assumption that the natural rate of risk-taking is less than 1. Any reallocated promotions must therefore be allotted to failure.

The question of when to pay bonuses is less straightforward, since bonus payments are feasible following both success and failure. It turns out that the cost-minimizing bonus scheme pays bonuses only for failure. This is because the incentive power of a bonus is determined by its probability of being earned by the marginal agent, while its cost to the organization is measured by the total number of employees who earn it. Since inframarginal

[26]Recall that $N^{nc} < 1$ when $\beta > \underline{\beta}$.

employees fail less often than does the marginal one, success bonuses get paid more often in expectation than to the marginal employee, while failure bonuses get paid less often. The incentive power-per-dollar of bonuses is therefore maximized by paying for failure.

We now formally characterize an optimal incentive scheme as a function of the risk-taking target. Unlike in the case of decreasing the risk-taking rate, a single-crossing result is no longer ensured in general. We first prove a result which holds without any regularity conditions.

Proposition 4. Suppose that the organization implements a risk-taking rate $N > N^{nc}$. Then one of the following schemes is optimal:

(B) The organization pays a positive bonus following failure outcomes and does not reallocate promotions.

(Pr) The organization pays no bonuses and reallocates promotions from neutral outcomes toward failure outcomes.

If $N < 1$ is sufficiently close to 1, then scheme (Pr) is uniquely optimal. If R/V is sufficiently large and N is sufficiently close to N^{nc}, then scheme (B) is uniquely optimal.

Under a mild regularity condition, a stronger single-crossing result can be proven. Define

$$\Lambda(N) \equiv \frac{N - \mu(N)}{(1 - \gamma(N))(1 - N)},$$

where recall that $\mu(N) = \int_0^N \gamma(n)\, dn$ is the number of successful risk-takers when the risk-taking rate is N.

Proposition 5. Suppose that the organization implements a risk-taking rate $N > N^{nc}$. If Λ is non-decreasing, there exists a threshold $\overline{N}_+ \in [0, 1)$ such that:

1. If $N < \overline{N}_+$, there exists an optimal scheme which pays a positive bonus following failure outcomes and does not reallocate promotions.

2. If $N > \overline{N}_+$, there exists an optimal scheme which pays no bonuses and reallocates promotions from neutral outcomes toward failure outcomes.

Further, the optimal scheme is unique whenever $N \neq 1, \overline{N}_+$. If R/V is sufficiently large, then $\overline{N}_+ > N^{nc}$.

The main features of this result are very similar to the properties of an optimal scheme which decreases N, as characterized in Proposition 3. The forces shaping the two results are closely analogous, except that in the current setting the relative incentive power-per-dollar (IPD) of promotions and bonuses exhibits a more complex dependence on N.

Heuristically, reallocating m promotions changes the probability of promotion by

$$\Delta\sigma_0(m) = -\frac{m}{1-N}, \quad \Delta\sigma_B(m) = \frac{m}{N-\mu(N)},$$

yielding total incentive power

$$V((1-\gamma(N))\Delta\sigma_B(m) - \Delta\sigma_0(m)) = \left(\frac{1-\gamma(N)}{N-\mu(N)} + \frac{1}{1-N}\right) Vm$$

Since the cost to the organization of this reallocation is $R(\pi_0 - \pi_B(N))m$, where $\pi_B(N)$ is the posterior belief that an employee is high-quality following a failure on the risky task, the IPD of promotions is

$$IPD^{Pr}(N) = \left(\frac{1-\gamma(N)}{N-\mu(N)} + \frac{1}{1-N}\right) \frac{V/R}{\pi_0 - \pi_B(N)}.$$

Meanwhile, the incentive power of a failure bonus of size t is $(1-\gamma(N))t$ and its cost to the organization is $(N-\mu(N))t$, yielding IPD

$$IPD^B(N) = \frac{1-\gamma(N)}{N-\mu(N)}.$$

As in the case of decreasing risk-taking, the IPD for promotions does not depend on the number of reallocated promotions. Similarly, the IPD for bonuses does not depend on the size of the bonus payments. Hence an optimal scheme exhibits a bang-bang structure in which only the tool with a higher IPD is used.

The ratio of these IPDs is

$$\frac{IPD^{Pr}(N)}{IPD^B(N)} = \frac{V/R}{\pi_0 - \pi_B(N)} (1 + \Lambda(N)).$$

In general the organization's inference from failure becomes weaker as more employees choose the risky task, due to the increased expected probability of failure. (See Appendix B for a proof.) In other words, $\pi_B(N)$ is increasing in N. Then so long as $\Lambda(N)$ is nondecreasing, the IPD of promotions relative to bonuses increases in N, making promotions a more attractive incentive tool relative to bonuses as the scheme becomes higher-powered.

Similar to the case of decreasing risk-taking, the key distinction between the two incentive tools is that bonuses impact only the payoff of the incentivized group, while promotions additionally affect the payoff of the disincentivized group. In a benchmark model in which the organization can manufacture promotions at cost $\overline{R} = R\pi_0$ to reward failure, the resulting IPD would be

$$\frac{\widehat{IPD}^{Pr}(N)}{IPD^B(N)} = \frac{V/R}{\pi_0 - \pi_B(N)},$$

which depends on N only due to the weakened inference regarding failure as N increases. So long as the cost of promoting for failure does not vary significantly with N, the optimal incentive tool does not vary with N in this benchmark. Compared to this benchmark, reallocating a promotion generates extra incentive power due to its impact on the payoff of risk-avoiders, and this extra power varies significantly with N.

In general the expression Λ capturing this extra incentive power is not guaranteed to be monotone. In particular, while the ratio $(N - \mu(N))/(1 - N)$ is guaranteed to be increasing in N, the factor $1 - \gamma(N)$ in the denominator of Λ is also increasing, possibly leading to non-monotonicity and failure of single-crossing. Despite this complexity, $\Lambda(N)$ is guaranteed to grow without bound as N approaches 1, so that for N sufficiently large promotions are the optimal incentive tool. Additionally, so long as V/R is sufficiently small, the IPD of promotions relative to bonuses is guaranteed to be less than 1 for N close to N^{nc}, yielding bonuses as the optimal incentive tool.

The regularity condition that Λ be monotone is not particularly stringent. We conclude our analysis by demonstrating that the regularity condition is satisfied by a wide class of match type distributions. Let $\bar{\gamma}(N) \equiv \frac{1}{N}\int_0^N \gamma(n)\, dn$ be the average match type of all risk-takers. Since γ is decreasing, so is $\bar{\gamma}$. The following lemma shows that Λ is monotone so long as $\bar{\gamma}$ drops more slowly as N increases.

Lemma 2. If $\bar{\gamma}$ is convex, then Λ is nondecreasing. In particular, $\bar{\gamma}$ is convex if $\gamma(N) = A - BN^k$ for some constants $A, B > 0$ and $k \in (0, 1]$.

5.2 The optimal direction of incentives

We next identify the direction of optimal incentives, that is, whether an optimal scheme increases or decreases the risk-taking rate from the natural rate. We show that the answer depends on the direction of the risk-taking distortion identified in Proposition 2. An optimal scheme counteracts this distortion by moving risk-taking in the direction of the first-best rate.

Let N^* be the risk-taking rate induced by an optimal scheme.[27] Recall that N^{nc} is the natural risk-taking rate, while N^{fb} is the first-best risk-taking rate, as characterized in Section 3.

Lemma 3 (The optimal direction of incentives)**.** The organization optimally shifts the risk-taking rate toward the first-best rate:

- If $N^{fb} > N^{nc}$, then $N^* \geq N^{nc}$.

[27] N^* is not guaranteed to be unique. In case of non-uniqueness of N^* or N^{fb}, Lemma 3 holds for every selection from each set of maximizers.

- If $N^{fb} < N^{nc}$, then $N^* \leq N^{nc}$.
- If $N^{fb} = N^{nc}$, then $N^* = N^{nc}$.

This result is most easily demonstrated by supposing that in the absence of incentive constraints, the organization's profit function $\Pi^{fb}(N)$ is single-peaked in N. In that case, choosing a risk-taking rate N which is further from N^{fb} than N^{nc} must decrease profits: $\Pi^{fb}(N) < \Pi^{fb}(N^{nc})$. Since the profits $\Pi^*(N)$ that can be achieved while respecting incentive constraints must fall below $\Pi^{fb}(N)$ for all $N \neq N^{nc}$, we therefore have $\Pi^*(N) < \Pi^{fb}(N^{nc})$. Meanwhile at the natural rate, incentive constraints are non-binding, implying $\Pi^{fb}(N^{nc}) = \Pi^*(N^{nc})$ and therefore $\Pi^*(N) < \Pi^*(N^{nc})$. Thus any incentive scheme implementing the risk-taking rate N must reduce profits compared to the natural rate, meaning N cannot be an optimal risk-taking rate. In general the unconstrained profit function is not guaranteed to be single-peaked, but it has enough structure that similar reasoning can be used to prove Lemma 3.

5.3 Determining the optimal incentive tool

Given the direction of optimal incentives, the remaining qualitative feature of an optimal incentive scheme to be determined is the optimal incentive tool. Section 5.1 established that the answer depends on the desired power of incentives, that is, how far the organization's target level of risk-taking deviates from the natural rate. Of course, the risk-taking target is itself a choice variable for the organization. The optimal incentive tool must therefore be jointly determined along with the level of risk-taking to fully characterize an optimal scheme.

Directly characterizing the optimal risk-taking rate requires an optimization of the organization's optimal profit function $\Pi^*(N)$, calculated based on the optimal scheme identified in Section 5.1, across all N. Unfortunately, $\Pi^*(N)$ is a nonlinear function of N which is typically not quasiconcave. Its maximum is therefore not uniquely characterized by a first-order condition. Indeed, under many parameterizations the profit function exhibits local maxima in both the low-powered incentive and high-powered incentive regimes. Calculating the optimal risk-taking rate therefore requires a comparison of the maximal profits achieved using each incentive tool, which cannot be accomplished analytically.

Despite these difficulties, we can derive conditions under which the optimal risk-taking rate is guaranteed to lie in the low- or high-powered regime, yielding a characterization of the optimal incentive tool. We show that the optimal tool, accounting for endogeneity of the risk-taking target, depends critically on features of the internal labor market as measured by R and V.

In general, the incentive costs of reallocating promotions and paying bonuses increase as R and V increase, respectively. Hence when both R and V are large, no incentive scheme can profitably affect risk-taking. As one component of our results, we establish conditions under which the optimal incentive scheme implements a risk-taking rate different from the no-commitment level N^{nc}. We call such an incentive scheme *nontrivial*.

We first establish that when an employee's value of promotion V is sufficiently small, the organization optimally incentivizes with bonuses.

Proposition 6. Hold all model parameters fixed except for V. Suppose that $\beta \neq \beta^{fb}$. Then for V sufficiently small, there exists a nontrivial optimal incentive scheme, and every optimal scheme incentivizes with bonuses.

Recall from Proposition 2 that when $\beta \neq \beta^{fb}$, the natural incentives lead to suboptimal risk-taking. In that case, Proposition 6 establishes that when V is small, there exists an incentive scheme which improves on the natural incentives, and further an optimal scheme incentivizes employees to change their risk-taking behavior using bonuses. Intuitively, when V is small it becomes cheap to influence risk-taking by paying bonuses, while the cost of incentivizing with promotions remains bounded away from 0 when R is held fixed.

We next establish that when the value of selection R is sufficiently small, the organization optimally incentivizes with promotions. Let β^0 be the unique $\beta \in (\underline{\beta}, \overline{\beta})$ such that $\gamma(N^{nc}(\beta)) = K$.

Proposition 7. Hold all model parameters fixed except for R. Suppose that $\beta \neq \beta^0$. Then for R sufficiently small, there exists a nontrivial optimal incentive scheme, and every optimal scheme incentivizes with promotions.

This proposition compares the promotion rate β to the reference level β^0 rather than β^{fb}. This is because β^{fb} varies with R, and so the hypothesis $\beta \neq \beta^{fb}$ cannot be maintained independently of the value of R. It can be shown that as R goes to zero, β^{fb} approaches β^0. Hence whenever $\beta \neq \beta^0$, the risk-taking rate induced by the natural incentives is bounded away from the optimal level for small R. In that case, Proposition 7 establishes that for sufficiently small R, there exists an incentive scheme which improves on the natural incentives, and an optimal scheme incentivizes employees to change their behavior by reallocating promotions. Intuitively, when R is small it becomes cheap to influence risk-taking by reallocating promotions, while the cost of paying bonuses remains bounded away from 0 when V is held fixed.

6 Asymmetric incentive schemes

Our analysis so far has assumed that the organization uses schemes which are *symmetric*: All employees are recommended the same task (as a function of their match type), and all employees who choose the same task and achieve the same outcome are rewarded in the same way. We now examine whether and how an organization can benefit from more general asymmetric schemes. Such schemes may be relevant in large organizations which can introduce social barriers between divisions, for instance by maintaining offices in multiple locations, facilitating separate corporate cultures and incentive schemes across divisions.

We relax symmetry by allowing the organization to partition employees into multiple groups, each of which is allocated a (potentially unequal) subset of the available promotions. Within each group, the organization commits to a symmetric incentive scheme using the promotions allotted to that group, with the freedom to offer distinct schemes to different groups. Employees observe their assigned group and incentive scheme prior to choosing a task. The following definition formalizes this setup:

Definition 2. An *asymmetric incentive scheme* consists of a countable set G of employee groups and a set of triples $\mathcal{A} = \{(k^g, \beta^g, \mathcal{S}^g)\}_{g \in G}$, where

- $k^g \in (0, 1]$ is the measure of employees in group g,
- $\beta^g \in [0, 1]$ is the promotion rate within group g,
- $\mathcal{S}^g = (\mathcal{N}^g, \boldsymbol{\sigma}^g, \boldsymbol{T}^g)$ is the (symmetric) incentive scheme offered to group g.

We interpret an asymmetric scheme as dividing employees into groups through uniform random assignment. The distribution of qualities and match types within a group is therefore identical to the aggregate population, implying a natural correspondence between each group g and the full organization with promotion rate β^g.

An asymmetric scheme is *feasible* if 1) each scheme $\mathcal{S}^g$ is feasible in the sense defined in Section 4; and 2) the number of employees and promotions within each group sum to the aggregate population size and promotion rate:

$$\sum_{g \in G} k^g = 1, \qquad \sum_{g \in G} k^g \beta^g = \beta. \tag{2}$$

It is *incentive-compatible* if each scheme $\mathcal{S}^g$ is incentive-compatible in the sense defined in Section 4. It is *admissible* if it is both feasible and incentive-compatible.

Under any admissible asymmetric scheme, the organization's profit from each group g is simply $\Pi(\mathcal{S}^g)$, as defined in Section 4, scaled by the number of employees k^g in the group.

The total profit from an admissible asymmetric scheme is therefore

$$\sum_{g\in G} k^g \Pi(\mathcal{S}^g).$$

Holding fixed a group structure $\mathcal{G} = (G, \{k^g, \beta^g\}_{g\in G})$, the organization optimally offers each group the (symmetric) incentive scheme characterized in Section 5. Let $\Pi^*(\beta)$ denote the profits derived from such a scheme, as a function of the promotion rate β. Then the organization's optimal profits under a given group structure $\mathcal{G}$ are

$$\sum_{g\in G} k^g \Pi^*(\beta^g).$$

Choosing a group structure to optimize this objective is mathematically equivalent to concavifying the symmetric profit function Π^*. Viewed as a concavification problem, the group structure functions as a randomization over β, with probability k^g assigned to promotion rate β^g. The aggregate feasibility constraints (2) ensure that the probabilities sum to 1 and that the average promotion rate is equal to the ex ante promotion rate β.

We establish two main properties of an optimal asymmetric scheme. First, there always exists an optimal scheme involving at most two distinct groups. This result follows directly from the fact that a single-variable function can be concavified by randomizing over at most two points in the function's domain. Second, there always exists an optimal asymmetric scheme in which promotions within each group are allocated according to the natural policy.[28] In other words, when asymmetric schemes are possible, any reallocation of promotions should be done "ex ante" rather than "ex post"; that is, employees should be informed about whether they will be favored or disfavored for promotion before they choose tasks. The following proposition formally establishes this result.

Proposition 8. There exists an optimal asymmetric incentive scheme in which promotions are allocated according to the natural policy within each group.

Mathematically, the result follows from the fact that the symmetric profit function $\Pi^*(\beta)$ is strictly convex in β whenever an optimal symmetric scheme is non-trivial and incentivizes with promotions. As a result, any mixture over β achieving the concave envelope of profits places no weight on promotion rates in this region. An optimal group structure must therefore involve only promotion rates β at which an optimal symmetric incentive scheme is either

[28] It can additionally be shown that *all* optimal schemes exhibit this property under the regularity condition that the optimal incentive scheme is nontrivial when $\beta = \beta^0$, where β^0 is the unique β satisfying $\gamma(N^{nc}(\beta)) = K$.

trivial or incentivizes with bonuses. In either case, promotions are allocated according to the natural policy.

Intuitively, convexity of Π^* can be understood as follows. Suppose that the organization maintains a fixed risk-taking target N over a range of β. Then as β increases within that range, the organization reallocates proportionally more promotions to maintain indifference of the marginal employee. As a result, promotion payoffs, and therefore total profits $\Pi(N, \beta)$, are linear in β under a fixed target N. The optimal profit function Π^* departs from this benchmark because the organization can flexibly adjust N as β varies. Since the maximum over a set of linear functions is convex, this optionality introduces convexity to $\Pi^*(\beta) = \max_N \Pi(N, \beta)$.

This logic does not necessarily imply that optimal promotion rates within each group are 0 or 1, for two reasons. First, reallocated promotions are proportional to β only so long as N^{nc} does not cross the target rate N. (Recall that N^{nc} varies with β, as demonstrated in Section 3.) At this crossing point, the required reallocation scheme changes qualitatively, introducing a kink in $\Pi(N, \beta)$ with respect to β which turns out to be concave. Second, when an optimal scheme incentivizes using bonuses rather than promotions, total bonus payments (hence also profits) are in general not globally linear in β. Linearity fails both when N^{nc} crosses N, as when incentives are provisioned through promotions, as well as when the marginal group of employees who face rationing under the natural policy changes. An optimal asymmetric scheme could therefore involve one or more groups with an interior promotion rate who face only the natural incentives or who are incentivized with bonuses.

We conclude our analysis with a pair of examples illustrating the process of constructing an optimal asymmetric scheme. In our first example, summarized in Figure 2, promotions are the only relevant incentive tool, corresponding to an environment where R is small. The top planel of the figure plots a possible symmetric profit function Π^* and corresponding concavified profit function Π^{A*} as a function of β, with the associated optimal risk-taking rate N^* plotted in the bottom panel. Π^* is convex wherever the optimal risk-taking rate differs from the natural rate. By contrast, for values of β at which $N^*(\beta) = N^{nc}(\beta)$, optimal symmetric profits are concave in β.

Given the convex-concave-convex structure of Π^*, the concavified profit function Π^{A*} is linear for low and high values of β, and concave over the interior interval (β_*, β^*). If $\beta \in (\beta_*, \beta^*)$, then the organization does not benefit by splitting employees into multiple groups. Otherwise, an optimal asymmetric scheme splits employees into two groups. If $\beta < \beta_*$, one of these groups is promoted at rate β_* while the remaining group receives no promotions; and if $\beta > \beta^*$, one group is promoted at rate β^* while the remaining group is promoted with certainty. In each of these groups, employees face only the natural incentives associated with their promotion rate.

In our next example, summarized in Figure 3, bonuses also play a role in an optimal scheme. The panels of this figure are analogous to the panels of Figure 2. The grey region in the lower panel corresponds to (N, β) pairs for which bonuses are the optimal tool in a symmetric incentive scheme. The green dashed line $N^{\ddagger}$ identifies where the group of agents who face rationing under the natural promotion policy changes with N. Π^* is convex whenever N^* differs from N^{nc} and $N^{\ddagger}$. By contrast, Π^* has a concave kink when N^* crosses N^{nc} (at $\beta = \underline{\beta}$) and is concave when $N^*(\beta) = N^{\ddagger}(\beta)$. As a result, the concave envelope Π^{A*} is concave at the kink $\underline{\beta}$ and on the interval (β_*, β^*) and linear elsewhere.

Whenever $\beta \notin \{\underline{\beta}\} \cup (\beta_*, \beta^*)$, the organization benefits from splitting employees into two groups. In contrast to the previous example, some groups may face a nontrivial incentive scheme that awards them bonuses. In particular, if $\beta \in (\underline{\beta}, \beta_*)$, one group is promoted at rate β_* and awarded bonuses to incentivize a risk-taking rate $N^{\ddagger}(\beta_*) > N^{nc}(\beta_*)$, and the other group is promoted at rate $\underline{\beta}$ and is offered a trivial incentive scheme, resulting in a risk-taking rate $N^{nc}(\underline{\beta})$. Similarly, if $\beta \in (\beta^*, 1)$, one group is promoted at rate β^* and awarded bonuses to incentivize a risk-taking rate $N^{\ddagger}(\beta^*)$, and the other group is promoted with certainty, requiring no bonuses. Although bonuses are paid in some of these groups, employees in all groups are promoted according to the natural policy within their group.

7 Conclusion

In this paper we have analyzed how employees' career concerns in an organization may distort their choice of projects, and have characterized how organizations should design promotion and bonus policies to mitigate these distortions. In organizations with little upward mobility, employees are overly motivated to take risks in order to stand out, while in very dynamic organizations, employees become preoccupied with avoiding tarnishing failures. An optimal promotion policy addresses these issues by overpromoting certain categories of underperforming employees—when employees naturally take too many risks, those who choose routine projects are overpromoted, while when employees naturally take too few risks, those who take on risky projects and fail are overpromoted. When both bonuses and promotions can be used for motivation, we find that bonuses are optimal for inducing low-powered incentives, while promotions are better for inducing high-powered incentives. We further characterize how the optimal intensity of incentives varies with the value of ex-post selection of high-quality employees.

We show these results in the context of a decentralized organization in which employees are free to choose their own projects. We therefore abstract from the role of management in

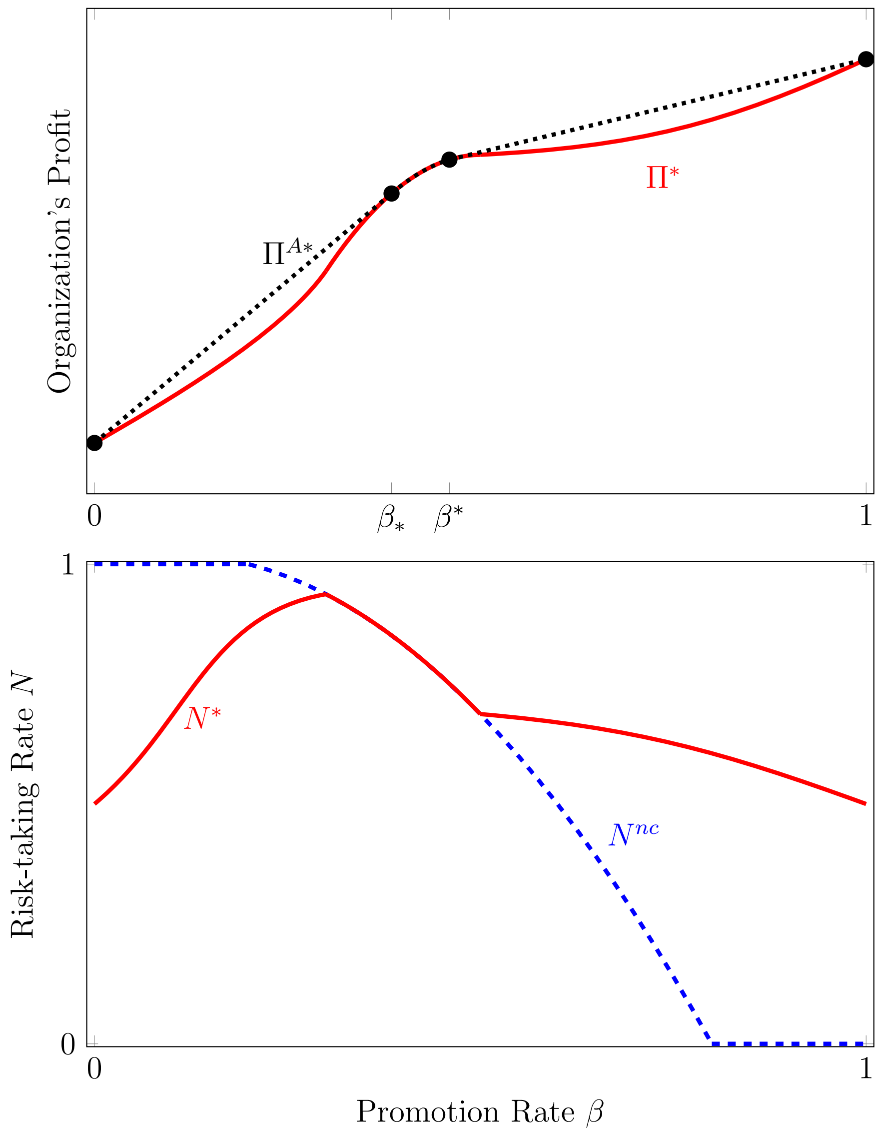

Figure 2: An example of an optimal asymmetric incentive scheme. The top panel plots an optimal symmetric profit Π^* as a function of β. The optimal asymmetric profit Π^{A*} is the concavification of $\Pi^*(\cdot)$. The bottom panel plots an optimal risk-taking rate N^* corresponding to $\Pi^*(\cdot)$. In this example, an optimal symmetric scheme incentivizes with promotions for all (N, β).

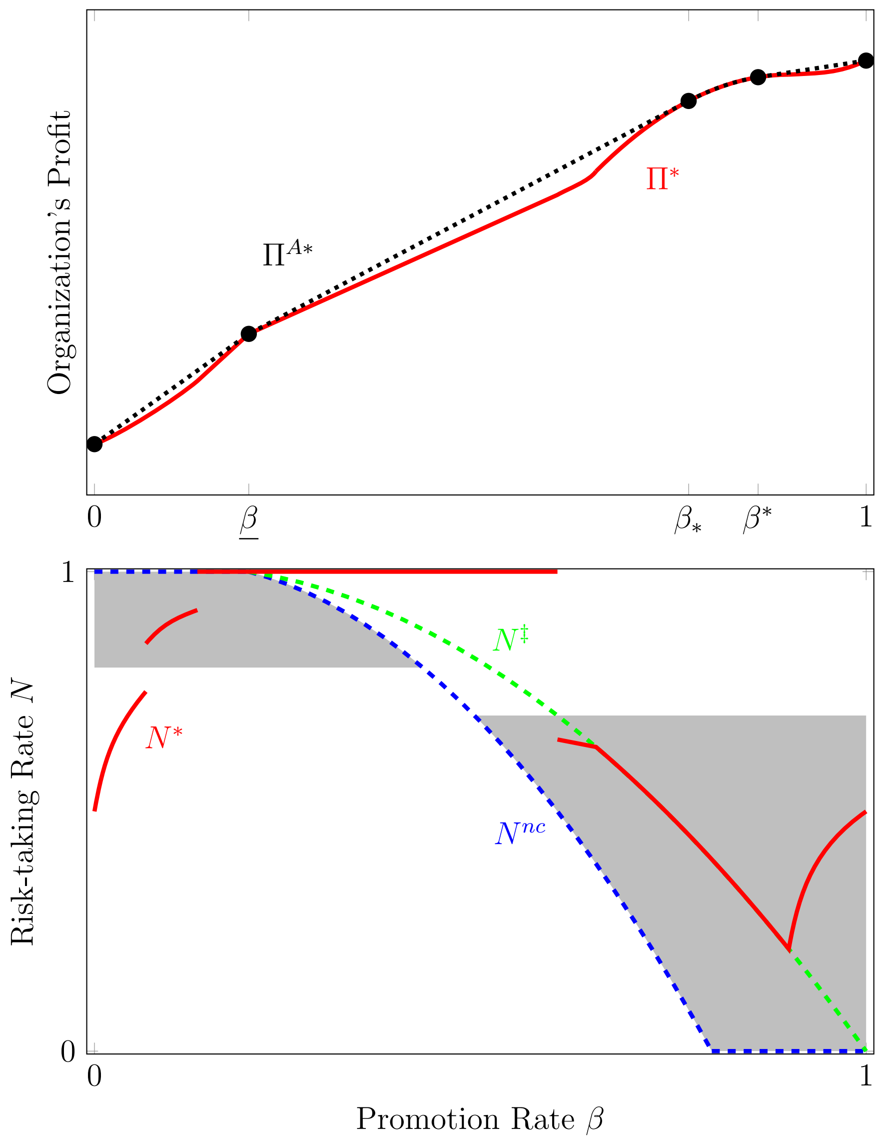

Figure 3: An example of an optimal asymmetric incentive scheme involving bonuses. The panels of this figure are as in Figure 2. The gray shading in the lower panel identifies (N, β) pairs for which an optimal symmetric incentive scheme uses bonuses. The line $N^\ddagger$ indicates when the group of rationed employees under the natural promotion policy changes with N.

directly assigning employees to projects. While such top-down decisions are indisputably an important management function, they can backfire. Since employees typically possess private information about their fit to (or excitement for) particular projects, top-down assignment of employees to projects can lead to poor matching and significant efficiency losses. Our paper is motivated by applications in which this efficiency loss is prohibitive compared to the costs of a bonus or promotion scheme. The interaction between top-down assignment and incentive schemes is left as an interesting direction for future research.

We have also abstracted from moral hazard by assuming that employees need not exert unobserved effort to complete projects. This assumption reflects organizations in which employee activities are highly visible, so that managers can directly monitor employees to ensure they're working hard on their chosen project. A natural concern is that in organizations or jobs without such visibility, incentive schemes which reward failure may create perverse incentives for employees to choose risky projects and then shirk. Whether such forces might hamstring efforts to encourage risk-taking, and whether optimal incentive schemes in such environments may resort to bonuses or other rewards to success, are important unanswered questions for further research.

Finally, in our setting all employees are viewed as having the same ex-ante quality prior to completing a project. This assumption is natural in settings where all employees are new hires or newly promoted into their role. But in some contexts, employees may either enter their roles with heterogeneous initial reputations, or may develop divergent reputations over time in a dynamic setting. It would be interesting to extend our setting to accommodate such heterogeneity, in particular by allowing employees to undertake a sequence of projects before being evaluated for promotion, in order to understand its implications for optimal design of incentives.

References

Aghion, Philippe and Matthew O. Jackson (2016). "Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits". *American Economic Journal: Microeconomics* 8 (3), pp. 1–38.

Autor, David H., Frank Levy, and Richard J. Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration". *The Quarterly Journal of Economics* 118 (4), pp. 1279–1333.

Baker, George P., Michael C. Jensen, and Kevin J. Murphy (1988). "Compensation and Incentives: Practice vs. Theory". *The Journal of Finance* 43 (3), pp. 593–616.

Bar-Isaac, Heski and Raphaël Lévy (2022). "Motivating Employees through Career Paths". *Journal of Labor Economics* 40 (1), pp. 95–131.

Benson, Alan, Danielle Li, and Kelly Shue (2019). "Promotions and the Peter Principle". *The Quarterly Journal of Economics* 134 (4), pp. 2085–2134.

Bernhardt, Dan (1995). "Strategic Promotion and Compensation". *The Review of Economic Studies* 62 (2), pp. 315–339.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb (2020). "Are Ideas Getting Harder to Find?" *American Economic Review* 110 (4), pp. 1104–1144.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999a). "The Economics of Career Concerns, Part I: Comparing Information Structures". *The Review of Economic Studies* 66 (1), pp. 183–198.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999b). "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies". *The Review of Economic Studies* 66 (1), pp. 199–217.

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

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

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

Hackman, Richard J. and Greg R. Oldham (1980). *Work Redesign*. Reading, MA: Addison-Wesley.

Holmstrom, Bengt and Joan Ricart I Costa (1986). "Managerial Incentives and Capital Management". *The Quarterly Journal of Economics* 101 (4), pp. 835–860.

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

Hörner, Johannes and Nicolas S. Lambert (2021). “Motivational Ratings”. *The Review of Economic Studies* 88 (4), pp. 1892–1935.

Hvide, Hans K. and Todd R. Kaplan (2005). “Delegated Job Design”. Unpublished.

Kaarbøe, Oddvar M. and Trond E. Olsen (2006). “Career concerns, monetary incentives and job design”. *Scandinavian Journal of Economics* 108 (2), pp. 299–316.

Kostadinov, Rumen and Aditya Kuvalekar (2022). “Learning in Relational Contracts”. *American Economic Journal: Microeconomics*, Forthcoming.

Kuvalekar, Aditya and Elliot Lipnowski (2020). “Job Insecurity”. *American Economic Journal: Microeconomics* 12 (2), pp. 188–229.

Lazear, Edward P. (2004). “The Peter Principle: A Theory of Decline”. *Journal of Political Economy* 112 (S1), pp. S141–S163.

Lazear, Edward P. and Sherwin Rosen (1981). “Rank-order tournaments as optimum labor contracts”. *Journal of Political Economy* 89 (5), pp. 841–864.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent (2009). “Performance Pay and Wage Inequality”. *The Quarterly Journal of Economics* 124 (1), pp. 1–49.

Levy, Frank and Richard J. Murnane (2004). *The New Division of Labor: How Computers Are Creating the Next Job Market*. Princeton University Press.

Manso, Gustavo (2011). “Motivating Innovation”. *The Journal of Finance* 66 (5), pp. 1823–1860.

Milgrom, Paul and John Roberts (1988). “An Economic Approach to Influence Activities in Organizations”. *American Journal of Sociology* 94 (S), pp. S154–S179.

Milgrom, Paul and John Roberts (1992). *Economics, Organization and Management*. Pearson.

Nalebuff, Barry J. and Joseph E. Stiglitz (1983). “Prizes and incentives: towards a general theory of compensation and competition”. *The Bell Journal of Economics*, pp. 21–43.

Rosen, Sherwin (1986). “Prizes and Incentives in Elimination Tournaments”. *The American Economic Review* 76.4, pp. 701–715.

Schöttner, Anja and Veikko Thiele (2010). “Promotion Tournaments and Individual Performance Pay”. *Journal of Economics & Management Strategy* 19 (3), pp. 699–731.

Siemsen, Enno (2008). “The Hidden Perils of Career Concerns in R&D Organizations”. *Management Science* 54 (5), pp. 865–877.

Waldman, Michael (1984). “Job Assignments, Signalling, and Efficiency”. *The RAND Journal of Economics* 15 (2), pp. 255–267.

Zwiebel, Jeffrey (1995). "Corporate Conservatism and Relative Compensation". *Journal of Political Economy* 103 (1), pp. 1–25.

Appendix

A Notation for proofs

Given a risk-taking set $\mathcal{N}$, define

$$\mu(\mathcal{N}) \equiv \int_{\mathcal{N}} \gamma(n)\, dn, \quad Q(\mathcal{N}) \equiv \int_{\mathcal{N}} q(\overline{\theta}, \gamma(n))\, dn.$$

Let $\pi_G(\mathcal{N})$ and $\pi_B(\mathcal{N})$ be the organization's posterior belief about the quality of an agent who succeeds and fails at risk-taking, respectively. By Bayes' rule,

$$\pi_G(\mathcal{N}) = \frac{Q(\mathcal{N})}{\mu(\mathcal{N})}\pi_0, \quad \pi_B(\mathcal{N}) = \frac{|\mathcal{N}| - Q(\mathcal{N})}{|\mathcal{N}| - \mu(\mathcal{N})}\pi_0$$

whenever $|\mathcal{N}| > 0$. Since $q(\overline{\theta}, \Gamma) = \overline{\theta}\Gamma$, we have $Q(\mathcal{N}) = \overline{\theta}\mu(\mathcal{N})$, and so π_G is independent of $\mathcal{N}$ and will be written without an argument going forward.

For $N \in [0, 1]$, let

$$\mu(N) \equiv \mu([0, N]), \quad Q(N) \equiv Q([0, N]), \quad \rho(N) \equiv \mu(N) + \gamma(N)(1 - N).$$

For $N \in (0, 1]$, define $\pi_B(N) \equiv \pi_B([0, N])$. We extend this definition to $N = 0$ by defining $\pi_B(0) \equiv \lim_{N\downarrow 0} \pi_B(N)$. (Lemma B.1 below implies that $\pi_B(N)$ is monotone in N, so this limit is well-defined.)

Given any scheme $\mathcal{S} = (\mathcal{N}, \boldsymbol{T}, \boldsymbol{\sigma})$, the principal's total profits under $\mathcal{S}$ can be written

$$\Pi(\mathcal{S}) = f(\mathcal{N}) + \Pi^{Pr}(\mathcal{S}) + \Pi^{B}(\mathcal{S}),$$

where

$$f(\mathcal{N}) \equiv \int_{\mathcal{N}} \gamma(n)\, dn + K(1 - |\mathcal{N}|)$$

are total expected task payoffs,

$$\Pi^{Pr}(\mathcal{S}) \equiv R\left(\mu(\mathcal{N})\pi_G\sigma_G + (|\mathcal{N}| - \mu(\mathcal{N}))\pi_B(\mathcal{N})\sigma_B + (1 - |\mathcal{N}|)\pi_0\sigma_0\right)$$

are total profits from promotions, and

$$\Pi^{B}(\mathcal{S}) \equiv -\mu(\mathcal{N})T_G - (|\mathcal{N}| - \mu(\mathcal{N}))T_B - (1 - |\mathcal{N}|)T_0$$

are total profits from bonus payments (which are always non-positive). Using Bayes' rule, promotion profits can also be written

$$\Pi^{Pr}(\mathcal{S}) = R\pi_0\left(Q(\mathcal{N})\sigma_G + (|\mathcal{N}| - Q(\mathcal{N}))\sigma_B + (1 - |\mathcal{N}|)\sigma_0\right).$$

Let

$$M(\mathcal{S}) \equiv \mu(\mathcal{N})\sigma_G + (|\mathcal{N}| - \mu(\mathcal{N}))\sigma_B + (1 - |\mathcal{N}|)\sigma_0$$

denote the total number of employees promoted under $\mathcal{S}$.

B Posterior belief lemma

Lemma B.1. $\pi_B'(N) > 0$ for every $N \in (0, 1]$.

Proof. By Bayes' rule,

$$\pi_B(N) = \frac{\int_0^N (1 - q(\overline{\theta}, \gamma(n)))\, dn}{\int_0^N (1 - \gamma(n))\, dn} \pi_0.$$

Differentiating this expression yields

$$\pi_B'(N) = \pi_B(N) \left(\frac{1 - q(\overline{\theta}, \gamma(N))}{\int_0^N (1 - q(\overline{\theta}, \gamma(n)))\, dn} - \frac{1 - \gamma(N)}{\int_0^N (1 - \gamma(n))\, dn} \right).$$

Since $q(\overline{\theta}, \Gamma) = \overline{\theta}\Gamma$, we have

$$\frac{1 - q(\overline{\theta}, \Gamma)}{1 - \Gamma} = 1 - \frac{q(\overline{\theta}, \Gamma) - \Gamma}{1 - \Gamma} = 1 - \frac{\overline{\theta} - 1}{\Gamma^{-1} - 1},$$

which is strictly decreasing in Γ given that $\overline{\theta} > 1$. We may therefore write

$$\begin{aligned} \int_0^N (1 - q(\overline{\theta}, \gamma(n)))\, dn &= \int_0^N \frac{1 - q(\overline{\theta}, \gamma(n))}{1 - \gamma(n)} (1 - \gamma(n))\, dn \\ &< \frac{1 - q(\overline{\theta}, \gamma(N))}{1 - \gamma(N)} \int_0^N (1 - \gamma(n))\, dn. \end{aligned}$$

Combining this bound with the previous expression for $\pi_B'(N)$ yields $\pi_B'(N) > 0$. $\square$

C Natural promotion policy

Suppose that $\mathcal{N} = [0, N]$ for $N \in (0, 1)$. In this appendix we characterize the no-commitment promotion rates $\sigma_i^{nc}(N, \beta)$ associated with the natural promotion policy for each group $i \in \{G, B, 0\}$ given any N and β.

For any $N \in [0, 1]$, let

$$\nu(N) \equiv \mu(N) + (1 - N)$$

be the number of successful risk-takers plus risk-avoiders. Note that $\mu'(N) = \gamma(N) > 0$ while $\nu'(N) = \gamma(N) - 1 < 0$ for all $N \in (0, 1)$. Let

$$N^\dagger(\beta) \equiv \sup\{N \ : \ \mu(N) \leq \beta\}$$

be the smallest risk-taking rate such that the number of successful risk-takers exceeds β, and let

$$N^\ddagger(\beta) \equiv \sup\{N \ : \ \nu(N) \geq \beta\}$$

be the largest risk-taking rate such that the number of successful risk-takers plus risk-avoiders exceeds β. Then $N^\dagger$ and $N^\ddagger$ are both strictly positive for all $\beta \in (0,1)$, $N^\dagger$ is nondecreasing in β while $N^\ddagger$ is nonincreasing in β, and each is strictly monotone whenever it is less than 1. Further $N^\dagger(\beta) < 1$ iff $\beta < \underline{\beta}$, while $N^\ddagger(\beta) < 1$ iff $\beta > \underline{\beta}$.

The properties of $N^\dagger$ and $N^\ddagger$ imply that if $\beta \le \underline{\beta}$, then $\sigma_B^{nc}(N,\beta) = 0$ for all N, while

$$\sigma_G^{nc}(N,\beta) = \begin{cases} 1, & N \le N^\dagger(\beta) \\ \beta/\mu(N), & N > N^\dagger(\beta) \end{cases}$$

$$\sigma_0^{nc}(N,\beta) = \begin{cases} 0, & N \le N^\dagger(\beta) \\ (\beta - \mu(N))/(1-N), & N > N^\dagger(\beta) \end{cases}$$

Meanwhile if $\beta > \underline{\beta}$, then $\sigma_G^{nc}(N,\beta) = 1$ for all N, while

$$\sigma_0^{nc}(N,\beta) = \begin{cases} (\beta - \mu(N))/(1-N), & N \le N^\ddagger(\beta) \\ 1, & N > N^\ddagger(\beta) \end{cases}$$

$$\sigma_B^{nc}(N,\beta) = \begin{cases} 0, & N \le N^\ddagger(\beta) \\ (\beta - \nu(N))/(N - \mu(N)), & N > N^\ddagger(\beta) \end{cases}$$

Let

$$\Pi^{fb}(N,\beta) = f(N) + R\pi_0\big\{\sigma_G^{nc}(N,\beta)Q(N) + \sigma_B^{nc}(N,\beta)(N - Q(N)) + \sigma_0^{nc}(N,\beta)(1-N)\big\}$$

be the firm's profits under the natural promotion policy supposing that the risk-taking rate is exogenously set at N. (That is, this profit function ignores incentive constraints on how much risk-taking can be implemented through efficient promotion.) Then $N^{fb}(\beta) = \arg\max_{N\in[0,1]} \Pi^{fb}(N,\beta)$.

Lemma C.1. If $\beta \le \underline{\beta}$, then $\Pi^{fb}(N,\beta)$ is strictly concave in N and $N^{fb}(\beta)$ is single-valued.

If $\beta > \underline{\beta}$, then $\Pi^{fb}(N,\beta)$ is strictly concave in N on $[0, N^\ddagger(\beta)]$, and one of the following holds:

- $N^{fb}(\beta)$ is single-valued, $N^{fb}(\beta) < N^\ddagger(\beta)$, and $\Pi^{fb}(N,\beta)$ is decreasing in N on $[N^\ddagger(\beta), 1]$,
- $\min N^{fb}(\beta) \ge N^\ddagger(\beta)$ and $\Pi^{fb}(N,\beta)$ is increasing in N on $[0, N^\ddagger(\beta)]$.

Proof. Suppose first that $N \le \min\{N^\dagger(\beta), N^\ddagger(\beta)\}$. Then

$$\Pi^{fb}(N,\beta) = f(N) + R\pi_0(Q(N) + \beta - \mu(N)).$$

Differentiating wrt N yields

$$\frac{\partial \Pi^{fb}}{\partial N}(N,\beta) = \gamma(N) - K + R\pi_0(q(\overline{\theta},\gamma(N)) - \gamma(N)).$$

Recall that $q(\overline{\theta},\Gamma) = \overline{\theta}\Gamma$, and so this expression may be simplified to read

$$\frac{\partial \Pi^{fb}}{\partial N}(N,\beta) = \gamma(N) - K + R\pi_0(\overline{\theta} - 1)\gamma(N).$$

Since $\gamma'(N) < 0$ for all N, this derivative is strictly decreasing in N, and so $\Pi^{fb}(N,\beta)$ is strictly concave in N on $[0, \min\{N^\dagger(\beta), N^\ddagger(\beta)\}]$. When $\beta = \underline{\beta}$, we have $N^\dagger(\beta), N^\ddagger(\beta) = 1$, establishing strict concavity on $[0,1]$.

Next suppose that $\beta < \underline{\beta}$, so that $N^\ddagger(\beta) = 1 > N^\dagger(\beta)$. For $N \geq N^\dagger(\beta)$, we have

$$\Pi^{fb}(N,\beta) = f(N) + R\beta\frac{Q(N)}{\mu(N)}\pi_0 = f(N) + R\beta\pi_G.$$

Differentiating this expression wrt N yields

$$\frac{\partial \Pi^{fb}}{\partial N}(N,\beta) = \gamma(N) - K,$$

which is strictly decreasing in N. So $\Pi^{fb}(N,\beta)$ is strictly concave in N on $[N^\dagger(\beta), 1]$. Further,

$$\begin{aligned}\frac{\partial \Pi^{fb}}{\partial N}(N^\dagger(\beta)-,\beta) &= \gamma(N^\dagger(\beta)) - K + R\pi_0(\overline{\theta}-1)\gamma(N^\dagger(\beta)) \\ &> \gamma(N^\dagger(\beta)) - K = \frac{\partial \Pi^{fb}}{\partial N}(N^\dagger(\beta)+,\beta),\end{aligned}$$

so $\Pi^{fb}(N,\beta)$ has a concave kink at $N = N^\dagger(\beta)$. Concavity on either side of the kink therefore implies strict concavity over the entire domain $[0,1]$. Since any strictly concave function has a unique maximizer, $N^{fb}(\beta)$ is single-valued.

Finally, suppose that $\beta > \underline{\beta}$, so that $N^\dagger(\beta) = 1 > N^\ddagger(\beta)$. For $N \geq N^\ddagger(\beta)$, we have

$$\Pi^{fb}(N,\beta) = f(N) + R\pi_0\left(Q(N) + 1 - N + \frac{\beta - \nu(N)}{N - \mu(N)}(N - Q(N))\right).$$

Using the definition of $\nu(N)$, this is equivalently

$$\Pi^{fb}(N,\beta) = f(N) + R\left(\pi_0 - (1-\beta)\pi_B(N)\right).$$

So

$$\frac{\partial \Pi^{fb}}{\partial N}(N,\beta) = \gamma(N) - K - R(1-\beta)\pi_B'(N).$$

Lemma B.1 established that $\pi_B'(N) > 0$ for all $N > 0$, so that

$$\frac{\partial \Pi^{fb}}{\partial N}(N,\beta) < \gamma(N) - K < \gamma(N^\ddagger(\beta)) - K < \frac{\partial \Pi^{fb}}{\partial N}(N^\ddagger(\beta)-,\beta)$$

for all $N > N^{\ddagger}(\beta)$. Let $\Delta \equiv \frac{\partial \Pi^{fb}}{\partial N}(N^{\ddagger}(\beta)-, \beta)$. If $\Delta < 0$, then $\Pi^{fb}(N, \beta)$ is decreasing for $N \geq N^{\ddagger}(\beta)$ as well as for $N < N^{\ddagger}(\beta)$ sufficiently large, and it must be that all maximizers of Π^{fb} are strictly smaller than $N^{\ddagger}(\beta)$. Since these maximizers are therefore also maximizers of the strictly concave function $\Pi^{fb}(\cdot, \beta)$ on the domain $[0, N^{\ddagger}(\beta)]$, there must be a unique maximizer. On the other hand, if $\Delta \geq 0$, then by strict concavity $\Pi^{fb}(N, \beta)$ must be increasing for $N \leq N^{\ddagger}(\beta)$, so that all maximizers of $\Pi^{fb}(\cdot, \beta)$ are no smaller than $N^{\ddagger}(\beta)$. □

D Proof of Lemma 1

If $N = 0$ or $N = 1$, then trivially $|\mathcal{N} \setminus [0, N]| = 0$ and there is nothing to prove, so we assume that $N \in (0, 1)$. The proof proceeds in two parts. In the first part, we fix a scheme $\mathcal{S} = (\mathcal{N}, \boldsymbol{T}, \boldsymbol{\sigma})$ satisfying $V\sigma_B + T_B = V\sigma_G + T_G$ and explicitly construct an admissible scheme $\mathcal{S}' = ([0, N], \boldsymbol{T}', \boldsymbol{\sigma}')$ such that $\Pi(\mathcal{S}') > \Pi(\mathcal{S})$. In the second part, we show that among all admissible schemes $\mathcal{S} = ([1 - N, 1], \boldsymbol{T}, \boldsymbol{\sigma})$ satisfying $V\sigma_B + T_B \geq V\sigma_G + T_G$, a profit-maximizing scheme exists, and any such scheme satisfies $V\sigma_B + T_B = V\sigma_G + T_G$.

The two results establish the lemma by the following logic. If $V\sigma_G + T_G > V\sigma_B + T_B$, then incentive-compatibility combined with $N \in (0, 1)$ imply that $\mathcal{N} = [0, N]$. Thus the hypothesis of the lemma requires $V\sigma_B + T_B \geq V\sigma_G + T_B$. In case the inequality binds, the first result produces the desired scheme. In case it is slack, incentive-compatibility requires that $\mathcal{N} = [1 - N, 1]$, since the payoff to risk-taking is strictly increasing in n. But then the second result implies that $\mathcal{S}$ is dominated by a scheme $\mathcal{S}' = ([1 - N, 1], \boldsymbol{T}', \boldsymbol{\sigma}')$ satisfying $V\sigma'_B + T'_B = V\sigma'_G + T'_G$. And by the first result, $\mathcal{S}'$ is in turn dominated by a scheme $\mathcal{S}'' = ([0, N], \boldsymbol{T}'', \boldsymbol{\sigma}'')$.

Part 1. Fix a scheme $\mathcal{S} = (\mathcal{N}, \boldsymbol{T}, \boldsymbol{\sigma})$ satisfying $V\sigma_B + T_B = V\sigma_G + T_G$, which may be equivalently written as the identity $T_B - T_G = V(\sigma_G - \sigma_B)$. In this case the expected payoff to risk-taking is the same for all employees. Then since $N \in (0, 1)$, incentive-compatibility requires that all employees must be indifferent between risk-taking and not, implying

$$V\sigma_G + T_G = V\sigma_0 + T_0 = V\sigma_B + T_B.$$

Suppose first that $\sigma_G = \sigma_B = \bar{\sigma}$ for some $\bar{\sigma}$. Then $T_G = T_B = \bar{T}$ for some $\bar{T}$, and the bonus and promotion payoffs under $\mathcal{S}$ may be written

$$\Pi^{Pr}(\mathcal{S}) = R\pi_0(N\bar{\sigma} + (1 - N)\sigma_0), \quad \Pi^B(\mathcal{S}) = -N\bar{T} - (1 - N)T_0.$$

Meanwhile the total number of employees promoted is

$$M(\mathcal{S}) = N\bar{\sigma} + (1 - N)\sigma_0$$

Promotion and bonus payoffs are independent of $\mathcal{N}$, holding fixed $|\mathcal{N}| = N$, as is the total number of employees promoted. In particular, they are unchanged under the modified scheme $\mathcal{S}' = ([0,N], \boldsymbol{T}, \boldsymbol{\sigma})$, which is therefore admissible. Meanwhile $f(N) > f(\mathcal{N})$ given that productivity increases as risk-taking shifts toward employees with higher-promise projects. Hence $\Pi(\mathcal{S}') > \Pi(\mathcal{S})$.

Next suppose that $\sigma_G > \sigma_B$. Define a family of schemes $\mathcal{S}'(\Delta) = ([0,N], \boldsymbol{T}'(\Delta), \boldsymbol{\sigma}'(\Delta))$ for $\Delta \in [0, \sigma_G - \sigma_B]$, where $\sigma'_G(\Delta) = \sigma_G - \Delta$, $T'_G(\Delta) = T_G + V\Delta$, and all remaining components of $\boldsymbol{\sigma}'(\Delta)$ and $\boldsymbol{T}'(\Delta)$ agree with $\boldsymbol{\sigma}$ and $\boldsymbol{T}$. Note that $\mathcal{S}'(\Delta)$ is incentive-compatible for any choice of Δ. Meanwhile $M(\mathcal{S}'(\Delta))$ is strictly decreasing in Δ, and $\sigma_G > \sigma_B$ along with $\mu(N) > \mu(\mathcal{N})$ imply that

$$\begin{aligned} M(\mathcal{S}'(0)) &= \mu(N)(\sigma_G - \sigma_B) + N\sigma_B + (1-N)\sigma_0 \\ &> \mu(\mathcal{N})(\sigma_G - \sigma_B) + N\sigma_B + (1-N)\sigma_0 = M(\mathcal{S}) \end{aligned}$$

while

$$M(\mathcal{S}'(\sigma_G - \sigma_B)) = N\sigma_B + (1-N)\sigma_0 < \mu(\mathcal{N})\sigma_G + (N - \mu(\mathcal{N}))\sigma_B + (1-N)\sigma_0 = M(\mathcal{S}).$$

Let $\Delta^* \in (0, \sigma_G - \sigma_B)$ be the unique Δ such that $M(\mathcal{S}'(\Delta)) = M(\mathcal{S})$. Then $\mathcal{S}'(\Delta^*)$ is admissible, and we will further show that $\Pi(\mathcal{S}'(\Delta^*)) > \Pi(\mathcal{S})$. Since $f(N) > f(\mathcal{N})$, it is sufficient to show that neither bonus nor promotion profits decrease under the modified scheme.

Solving the equation $M(\mathcal{S}'(\Delta^*)) = M(\mathcal{S})$ for Δ yields

$$\Delta^* = (\sigma_G - \sigma_B)\left(1 - \frac{\mu(\mathcal{N})}{\mu(N)}\right).$$

Calculating the difference in bonus profits between $\mathcal{S}$ and $\mathcal{S}'(\Delta^*)$ yields

$$\Pi^B(\mathcal{S}'(\Delta^*)) - \Pi^B(\mathcal{S}) = -(\mu(N) - \mu(\mathcal{N}))(T_G - T_B) - V\mu(N)\Delta^*.$$

Inserting the explicit form of Δ^* and using the identity $T_B - T_G = V(\sigma_G - \sigma_B)$ yields $\Pi^B(\mathcal{S}'(\Delta^*)) - \Pi^B(\mathcal{S}) = 0$. So the modified scheme generates the same total bonus profits as does the original scheme.

Meanwhile, calculating the difference in promotion payoffs between the two schemes yields

$$\Pi^{Pr}(\mathcal{S}'(\Delta^*)) - \Pi^{Pr}(\mathcal{S}) = R\pi_0 Q(N)\left((\sigma_G - \sigma_B)\left(1 - \frac{Q(\mathcal{N})}{Q(N)}\right) - \Delta^*\right).$$

Inserting the explicit form of Δ^* reduces this expression to

$$\Pi^{Pr}(\mathcal{S}'(\Delta^*)) - \Pi^{Pr}(\mathcal{S}) = R\pi_0 Q(N)(\sigma_G - \sigma_B)\left(\frac{\mu(\mathcal{N})}{\mu(N)} - \frac{Q(\mathcal{N})}{Q(N)}\right).$$

Since $\pi_G = Q(\mathcal{N})/\mu(\mathcal{N}) = Q(N)/\mu(N)$, it follows that $\Pi^{Pr}(\mathcal{S}'(\Delta^*)) = \Pi^{Pr}(\mathcal{S})$, and so the modified scheme yields identical promotion profits.

Finally, suppose that $\sigma_G < \sigma_B$. Define a family of schemes $\mathcal{S}'(\Delta) = ([0, N], \boldsymbol{T}'(\Delta), \boldsymbol{\sigma}'(\Delta))$ for $\Delta \in [0, \sigma_B - \sigma_G]$, where $\sigma'_G(\Delta) = \sigma_G + \Delta$, $T'_G(\Delta) = T_G - V\Delta$, and all remaining components of $\boldsymbol{\sigma}'(\Delta)$ and $\boldsymbol{T}'(\Delta)$ agree with $\boldsymbol{\sigma}$ and $\boldsymbol{T}$. Using work nearly identical to the $\sigma_G > \sigma_B$ case, it can be established that $M(\mathcal{S}'(\Delta))$ is strictly increasing in Δ, and there exists a unique $\Delta^* \in (0, \sigma_B - \sigma_G)$ such that $M(\mathcal{S}'(\Delta^*)) = M(\mathcal{S})$. In addition,

$$T'_G(\Delta^*) - T_B = T_G - T_B - V\Delta^* = V(\sigma_B - \sigma_G - \Delta^*) > 0,$$

so $T'_G(\Delta^*) > 0$. The resulting scheme $\mathcal{S}'(\Delta^*)$ is therefore admissible, and under it task payoffs strictly rise while promotion and bonus profits are unchanged as compared to $\mathcal{S}$.

Part 2. We now study the problem of maximizing $\Pi(\mathcal{S})$ among all admissible schemes $\mathcal{S} = ([1 - N, 1], \boldsymbol{T}, \boldsymbol{\sigma})$. Incentive-compatibility for this class of schemes is equivalent to

$$\begin{cases} V\sigma_B + T_B \geq V\sigma_G + T_G, \\ \gamma(1-N)(V\sigma_G + T_G) + (1 - \gamma(1-N))(V\sigma_B + T_B) = V\sigma_0 + T_0 \end{cases}$$

The first condition ensures that that the payoff to risk-taking is larger for higher-indexed employees, while the second condition ensures that the marginal employee N is indifferent between risk-taking or not.

We first argue that the optimal achievable profits are unchanged if the problem is modified to impose the bonus cap $T_G, T_0, T_B \leq \overline{T}$ for sufficiently large $\overline{T}$. For if not, then there would exist a sequence of admissible schemes $\mathcal{S}^n = ([1 - N, 1], \boldsymbol{T}^n, \boldsymbol{\sigma}^n)$ for $n = 1, 2, ...$ such that $\Pi(\mathcal{S}^{n+1}) > \Pi(\mathcal{S}^n)$ for each n and $\max\{T^n_G, T^n_0, T^n_B\} \to \infty$. But also the profit under each $\mathcal{S}^n$ can be bounded above by

$$\begin{aligned} \Pi(\mathcal{S}^n) &\leq f(\mathcal{N}) + R\beta - \mu([1-N,1])T^n_G - (N - \mu([1-N,1]))T^n_B - (1-N)T^n_0 \\ &\leq f(\mathcal{N}) + R\beta - \min\{\mu([1-N,1]), N - \mu([1-N,1]), 1-N\} \max\{T^n_G, T^n_0, T^n_B\}, \end{aligned}$$

and this upper bound approaches $-\infty$ as $n \to \infty$ given that $\min\{\mu([1-N,1]), N - \mu([1-N,1]), 1-N\} > 0$. This contradicts the hypothesis that profits are increasing in n, and so a sufficiently large upper bound on bonuses must not impact achievable profits.

Once a cap on bonuses is imposed, the resulting optimization problem involves a continuous objective function and a compact constraint set. Then by the maximum theorem, an optimal scheme must exist. By the argument of the previous paragraph, this scheme must also maximize the principal's profits absent a cap on bonuses, supposing the cap is set sufficiently large. Thus an optimal scheme exists in the original problem with no bonus cap.

In the remainder of the proof, we argue that any optimal scheme in the class of interest must satisfy $V\sigma_B + T_B = V\sigma_G + T_G$. It is sufficient to show that any scheme satisfying $V\sigma_B + T_B > V\sigma_G + T_G$ can be modified to obtain another admissible scheme in the class of interest yielding strictly higher profits. To that end, fix a scheme $\mathcal{S} = ([1-N, 1], \boldsymbol{T}, \boldsymbol{\sigma})$ satisfying $V\sigma_B + T_B > V\sigma_G + T_G$.

We begin by showing that unless $\min\{T_G, T_0\} = 0$ and $T_B = 0$, we can modify bonuses to increase profits. Suppose first that $\min\{T_G, T_0\} > 0$. Then there exists a $\Delta > 0$ sufficiently small such that the new bonus scheme $\boldsymbol{T}' = (T_G - \Delta/\gamma(1-N), T_0 - \Delta, T_B)$ satisfies the non-negativity constraints on bonuses. By construction, this new set of bonuses is fully incentive-compatible for every $\Delta > 0$. Further, this change strictly decreases total bonus payments. So the modified scheme increases total profits.

A similar argument yields profitable improvements if $\min\{T_B, T_0\} > 0$. Thus in particular if $T_B > 0$ and $T_0 > 0$, there exists a profitable improvement. Suppose instead that $T_B > 0$ and $T_0 = 0$. Consider a modified bonus scheme $\boldsymbol{T}'(\Delta)$ which sets

$$T'_G(\Delta) = T_G + \Delta\frac{1-\gamma(1-N)}{\gamma(1-N)}, \quad T'_0(\Delta) = T_0, \quad T'_B(\Delta) = T_B - \Delta.$$

This new bonus scheme preserves incentive-compatibility for the marginal employee for all Δ. Further, for $\Delta > 0$ sufficiently small, $T'_B(\Delta) > 0$ and $V\sigma_G + T'_G(\Delta) < V\sigma_B + T'_B(\Delta)$, so that the scheme satisfies the non-negativity constraints on bonuses and is incentive-compatible. Letting $\mathcal{S}'(\Delta) = ([1-N, 1], \boldsymbol{T}'(\Delta), \boldsymbol{\sigma})$, we have

$$\Pi^B(\mathcal{S}'(\Delta)) = -T_0(1-N) - T'_G(\Delta)\mu([1-N, 1]) - T'_B(\Delta)(N - \mu([1-N, 1])).$$

Differentiating wrt Δ yields

$$\frac{d}{d\Delta}\Pi^B(\mathcal{S}'(\Delta)) = -\frac{\mu([1-N, 1])}{\gamma(1-N)} + N.$$

Since γ is strictly decreasing, $\mu([1-N, 1]) < \gamma(1-N)N$, so that this derivative is strictly positive. Thus total profits from bonuses (i.e., the negative of total bonus payments) increases in Δ, meaning that for sufficiently small $\Delta > 0$ the modified scheme $\mathcal{S}'(\Delta)$ is admissible and strictly increases profits.

We have so far found a profitable modification of any scheme satisfying $\min\{T_G, T_0\} > 0$ or $T_B > 0$. It remains only to find a profitable modification in case $\min\{T_G, T_0\} = 0$ and $T_B = 0$. Incentive-compatibility for the marginal employee, combined with $V\sigma_B + T_B > V\sigma_G + T_G$ and $T_B = 0$, implies that

$$V\sigma_B > V\sigma_0 + T_0 > V\sigma_G + T_G.$$

In particular, $\sigma_B > \sigma_G, \sigma_0$, and so $\sigma_G, \sigma_0 < 1$.

Consider a family of modified schemes $\mathcal{S}'(\Delta, \alpha) = ([1-N, 1], \boldsymbol{T}, \boldsymbol{\sigma'}(\Delta, \alpha))$ for $\alpha \in (0,1)$ and $\Delta > 0$ with

$$\sigma'_G(\Delta, \alpha) = (1-\alpha)\sigma_G + \alpha\overline{\sigma} + \Delta, \quad \sigma'_B(\Delta, \alpha) = (1-\alpha)\sigma_B + \alpha\overline{\sigma} + \Delta, \quad \sigma'_0(\Delta, \alpha) = \sigma_0 + \Delta,$$

where $\overline{\sigma} \equiv \gamma(1-N)\sigma_G + (1-\gamma(1-N))\sigma_B$. All schemes in this family satisfy incentive-compatibility for the marginal employee. Full incentive-compatibility then requires

$$V\sigma'_B(\Delta, \alpha) + T_B \geq V\sigma'_G(\Delta, \alpha) + T_G$$

or

$$(1-\alpha)V(\sigma_B - \sigma_G) \geq T_G - T_B.$$

Since $V\sigma_B + T_B > V\sigma_G + T_G$, this inequality is slack when $\alpha = 0$, and so holds for $\alpha \in (0,1)$ sufficiently small.

Under $\mathcal{S}'(0, \alpha)$, the total number of promoted employees is

$$\begin{aligned} M(\mathcal{S}'(0, \alpha)) &= \mu([1-N, 1])\sigma'_G(0, \alpha) + (N - \mu([1-N, 1]))\sigma'_B(0, \alpha) + (1-N)\sigma'_0(0, \alpha) \\ &= (1-\alpha)M(\mathcal{S}) + \alpha\left(N\overline{\sigma} + (1-N)\sigma_0\right) \end{aligned}$$

Since $\gamma(1-N) > \mu([1-N, 1])/N$ and $\sigma_G < \sigma_B$, we have

$$\overline{\sigma} = \gamma(1-N)\sigma_G + (1-\gamma(1-N))\sigma_B < \frac{\mu([1-N, 1])}{N}\sigma_G + \frac{N - \mu([1-N, 1])}{N}\sigma_B,$$

and therefore $M(\mathcal{S}'(0, \alpha)) < M(\mathcal{S})$. Since $M(\mathcal{S}'(\Delta, \alpha))$ is increasing and unbounded in Δ, there exists a unique $\Delta^* > 0$ such that $M(\mathcal{S}'(\Delta^*, \alpha)) = M(\mathcal{S})$.

We next show that the promotion probabilities under $\mathcal{S}'(\Delta^*, \alpha)$ are feasible for sufficiently small $\alpha \in (0,1)$. Since each component of $\boldsymbol{\sigma'}(\Delta, \alpha)$ is a sum of non-negative and positive terms, it must be that $\boldsymbol{\sigma'}(\Delta, \alpha) > 0$ for any $\Delta > 0$ and $\alpha \in (0,1)$. It remains to check the upper bound $\boldsymbol{\sigma'}(\Delta^*, \alpha) \leq 1$. Note that Δ^* satisfies

$$M(\mathcal{S}) = (1-\alpha)M(\mathcal{S}) + \alpha(N\overline{\sigma} + (1-N)\sigma_0) + \Delta^*,$$

or $\Delta^* = \alpha\Delta^0$, where $\Delta^0 \equiv M(\mathcal{S}) - N\overline{\sigma} - (1-N)\sigma_0$. Since $\Delta^* > 0$, also $\Delta^0 > 0$. The promotion probabilities $\sigma'_G(\Delta^*, \alpha)$ and $\sigma'_0(\Delta^*, \alpha)$ may be written in terms of Δ^0 as

$$\sigma'_G(\Delta^*, \alpha) = (1-\alpha)\sigma_G + \alpha(\overline{\sigma} + \Delta^0), \quad \sigma'_0(\Delta^*, \alpha) = (1-\alpha)\sigma_0 + \alpha(\sigma_0 + \Delta^0).$$

Then as $\sigma_G, \sigma_0 < 1$ and $\overline{\sigma}, \Delta^0$ are independent of α, it must be that $\sigma'_G(\Delta^*, \alpha), \sigma'_0(\Delta^*, \alpha) < 1$ for α sufficiently small. Additionally, $\sigma'_0(\Delta^*, \alpha) > \sigma_0$ for any $\alpha(0,1)$, given that $\Delta^0 > 0$.

Meanwhile, $\overline{\sigma}$ is a weighted average of σ_G and σ_B, and so $\overline{\sigma} > \min\{\sigma_G, \sigma_B\}$. Since $\sigma_B > \sigma_G$, this implies $\overline{\sigma} > \sigma_G$, so that $\sigma'_G(\Delta^*, \alpha) > \sigma_G$ for all $\alpha \in (0,1)$. Then since $M(\mathcal{S}'(\Delta^*, \alpha)) = M(\mathcal{S})$, it must be that $\sigma'_B(\Delta^*, \alpha) < \sigma_B \leq 1$ for all $\alpha \in (0,1)$.

We have shown that for $\alpha \in (0,1)$ sufficiently small, the scheme $\mathcal{S}'(\Delta^*, \alpha)$ is admissible. The final step is to show that this modified scheme raises profits. Since risk-taking and bonuses are unchanged, we need only check that promotion profits rise. This follows from the fact that the modified scheme preserves the total number of promoted employees, while reallocating promotions from failed risk-takers to successful and non-risk-takers. Since $\pi_G > \pi_0 > \pi_B(\mathcal{N})$, this reallocation must therefore raise promotion profits.

E Proof of Proposition 1

Given a risk-taking set $\mathcal{N} \subset [0,1]$ and promotion rate β, let $\sigma_i^{nc}(\mathcal{N}, \beta)$ be the probability of promotion for employees in group $i \in \{G, 0, B\}$ under the natural promotion policy. In general

$$\sigma_G^{nc}(\mathcal{N}, \beta) \geq \sigma_0^{nc}(\mathcal{N}, \beta) \geq \sigma_B^{nc}(\mathcal{N}, \beta),$$

since employees are promoted strictly in order of perceived quality and $\pi_G(\mathcal{N}) > \pi_0 > \pi_B(\mathcal{N})$ given any risk-taking set $\mathcal{N}$.[29] Additionally, feasibility implies that $\sigma_B^{nc}(\mathcal{N}, \beta) < 1$, while $\sigma_B^{nc}(\mathcal{N}, \beta) > 0$ only if $\sigma_G^{nc}(\mathcal{N}, \beta) = 1$. Hence $\sigma_G^{nc}(\mathcal{N}, \beta) > \sigma_B^{nc}(\mathcal{N}, \beta)$. It follows that the payoff to risk-taking is strictly increasing in an employee's match type. Hence any equilibrium risk-taking set must satisfy $\mathcal{N} = \emptyset$, $\mathcal{N} = [0, N]$, or $\mathcal{N} = [0, N)$ for some risk-taking rate $N \in [0,1]$.

We first characterize all β for which $\mathcal{N} = \emptyset$ is an equilibrium. Note that $\sigma_G^{nc}(\emptyset, \beta) = 1$, $\sigma_0^{nc}(\emptyset, \beta) = \beta$, and $\sigma_B^{nc}(\emptyset, \beta) = 0$. It is an equilibrium outcome for all employees to avoid risk-taking iff this is true for the best-matched agent, who succeeds on the risky task with probability $\gamma(0)$. Given the promotion probabilities just reported, $\mathcal{N} = \emptyset$ is an equilibrium iff $\beta \geq \overline{\beta} \equiv \gamma(0)$. Note that when $\beta = \overline{\beta}$, this logic implies that $\mathcal{N} = \{0\}$ is also an equilibrium.

We next characterize all β for which $\mathcal{N} = [0,1]$ is an equilibrium. Suppose first that $\mu(1) < \beta$, where μ is as defined in Appendix A. Then $\sigma_G^{nc}([0,1], \beta) = 1$, $\sigma_0^{nc}([0,1], \beta) = 1$, and $\sigma_B^{nc}([0,1], \beta) = (\beta - \mu(1))/(1 - \mu(1))$. Since $\sigma_0^{nc}([0,1], \beta) = \sigma_G^{nc}([0,1], \beta) > \sigma_B^{nc}([0,1], \beta)$, it cannot be optimal for any employee to choose the risky task, and so $\mathcal{N} = [0,1]$ cannot be an equilibrium. If instead $\mu(1) \geq \beta$, then $\sigma_G^{nc}([0,1], \beta) = \beta/\mu(1)$, $\sigma_0^{nc}([0,1], \beta) = 0$, and $\sigma_B^{nc}([0,1], \beta) = 0$. In this case it is optimal for all employees to choose the risky task. Thus

[29] If $|\mathcal{N}| > 0$ then this ranking of posteriors is implied by Bayes' rule. If $|\mathcal{N}| = 0$, in which case $\pi_G(\mathcal{N})$ and $\pi_B(\mathcal{N})$ cannot be computed by Bayes' rule, this ranking is directly imposed, as discussed in fn 16.

$\mathcal{N} = [0,1]$ is an equilibrium iff $\beta \leq \underline{\beta} \equiv \mu(1)$. Note that

$$\mu(1) = \int_0^1 \gamma(n)\, dn < \gamma(0),$$

so that $\underline{\beta} < \overline{\beta}$. Additionally, this logic implies that $\mathcal{N} = [0,1)$ is an equilibrium when $\beta = \underline{\beta}$; and if $\gamma(1) = 0$ then $\mathcal{N} = [0,1)$ is an equilibrium for all $\beta < \underline{\beta}$.

We now characterize all β for which $\mathcal{N} = [0,N]$ is an equilibrium for $N \in (0,1)$. Suppose first that $\beta \geq \overline{\beta}$. Then

$$\mu(N) < \mu(1) = \int_0^1 \gamma(n)\, dn < \gamma(0) \leq \beta,$$

and so $\sigma_G^{nc}([0,N],\beta) = 1$. Additionally,

$$\frac{\beta - \mu(N)}{1-N} \geq \frac{\gamma(0) - \mu(N)}{1-N}.$$

Then since

$$\mu(N) = \int_0^N \gamma(n)\, dn < \int_0^N \gamma(0)\, dn = N\gamma(0),$$

we have $(\beta - \mu(N))/(1-N) > 1$. Hence $\sigma_0^{nc}([0,N],\beta) = 1$. Since $\sigma_B^{nc}([0,N],\beta) < 1$, it cannot be optimal for any employee to choose the risky task, meaning $\mathcal{N} = [0,N]$ cannot be an equilibrium.

Suppose next that $\beta \leq \underline{\beta}$. If $\mu(N) \geq \beta$, then $\sigma_0^{nc}([0,N],\beta) = 0$ and it cannot be optimal for any (except possibly the worst-matched, in case $\gamma(1) = 0$) employee to choose the safe task, meaning $\mathcal{N} = [0,N]$ cannot be an equilibrium. If on the other hand $\mu(N) < \beta$ then $\sigma_G^{nc}([0,N],\beta) = 1$. Additionally,

$$\frac{\beta - \mu(N)}{1-N} \leq \frac{\mu(1) - \mu(N)}{1-N} = \frac{1}{1-N}\int_{1-N}^1 \gamma(n)\, dn < \gamma(N).$$

Since $\gamma(N) < 1$, we therefore have $\sigma_0^{nc}([0,N],\beta) = (\beta - \mu(N))/(1-N)$ and $\sigma_B^{nc}([0,N],\beta) = 0$. The marginal employee's payoff gain from switching from the safe to the risky task is therefore

$$\gamma(N) - \frac{\beta - \mu(N)}{1-N} > 0,$$

meaning that employees slightly less well-matched than N would also strictly gain from switching to the risky task, and so $\mathcal{N} = [0,N]$ cannot be an equilibrium.

Finally, suppose that $\beta \in (\underline{\beta}, \overline{\beta})$. Since $\mu(N) < \mu(1) < \beta$, it must be that $\sigma_G^{nc}([0,N],\beta) = 1$. $\mathcal{N} = [0,N]$ is an equilibrium iff the marginal employee is indifferent between the two tasks. That is, we must have $(\beta - \mu(N))/(1-N) < 1$, so that $\sigma_0^{nc}([0,N],\beta) = (\beta - \mu(N))/(1-N)$ while $\sigma_B^{nc}([0,N],\beta) = 0$, and additionally

$$\gamma(N) = \frac{\beta - \mu(N)}{1-N}.$$

Since $\gamma(N) < 1$, the former condition is satisfied if the latter is. And the latter condition may be equivalently written $\rho(N) = \beta$, where $\rho(N) \equiv \mu(N) + \gamma(N)(1 - N)$. Note that $\rho'(N) = \gamma'(N)(1 - N) < 0$, and $\rho(0) = \overline{\beta}$ while $\rho(1) = \underline{\beta}$. Hence there exists a unique $N^{nc} \in (0, 1)$ such that $\mathcal{N} = [0, N^{nc}]$ is an equilibrium. Additionally, $\mathcal{N} = [0, N)$ is also an equilibrium whenever $\mathcal{N} = [0, N^{nc}]$ is.

The work above establishes that for each $\beta \in (0, 1)$, there exists an (essentially) unique equilibrium risk-taking set. When $\beta < \underline{\beta}$ we have $\mathcal{N} = [0, 1]$, when $\beta > \overline{\beta}$ we have $\mathcal{N} = \emptyset$, and when $\beta \in (\underline{\beta}, \overline{\beta})$ we have $\mathcal{N} = [0, N^{nc}]$, where $\rho(N^{nc}) = \beta$. The risk-taking rate N^{nc} satisfies the comparative static

$$\frac{dN^{nc}}{d\beta} = \frac{1}{\rho'(N^{nc})} < 0,$$

so that N^{nc} is decreasing in β. Additionally, $\rho(1) = \underline{\beta}$ and $\rho(0) = \overline{\beta}$, so that $N^{nc} \to 1$ when $\beta \to \underline{\beta}$ and $N^{nc} \to 0$ when $\beta \to \overline{\beta}$. Defining $N^{nc} = 1$ in the former case and $N^{nc} = 0$ in the latter case, we may write $\mathcal{N} = [0, N^{nc}]$ for $\beta \in [\underline{\beta}, \overline{\beta}]$.

F Proof of Proposition 2

Throughout this proof, we assume for simplicity that $N^{fb}(\beta)$ is single-valued. If it isn't, the proof holds for any selection from the set of maximizers.

In the absence of incentive constraints, the organization's profits under risk-taking rate N are $\Pi^{fb}(N, \beta)$, as characterized in Appendix C. We will make free use of expressions for $\partial \Pi^{fb}/\partial N$ calculated in the proof of Lemma C.1. We will also make use of the following basic fact about $N^{nc}(\beta)$.

Lemma F.1. $N^{nc}(\beta) \leq N^{\ddagger}(\beta)$ for all $\beta \in (0, 1)$, with the inequality strict if and only if $\beta > \underline{\beta}$.

Proof. For $\beta \leq \underline{\beta}$ we have $N^{nc}(\beta) = N^{\ddagger}(\beta) = 1$. Meanwhile for $\beta \geq \overline{\beta}$ we have $N^{nc}(\beta) = 0 < N^{\ddagger}(\beta)$. Finally, for $\beta \in (\underline{\beta}, \overline{\beta})$, $N^{nc}(\beta)$ satisfies $\rho(N^{nc}(\beta)) = \beta$ while $N^{\ddagger}(\beta)$ satisfies $\nu(N^{\ddagger}(\beta)) = \beta$. Note that for all $N < 1$,

$$\rho(N) = \mu(N) + \gamma(N)(1 - N) < \mu(N) + (1 - N) = \nu(N).$$

Then since $N^{nc}(\beta) < 1$ for $\beta \in (\underline{\beta}, \overline{\beta})$, we have $\nu(N^{nc}(\beta)) > \beta = \nu(N^{\ddagger}(\beta))$. Now, $\nu'(N) = \gamma(N) - 1 < 0$ for all $N \in (0, 1)$, so ν is decreasing in N. It follows that $N^{nc}(\beta) < N^{\ddagger}(\beta)$. $\square$

Suppose first that $\beta < \underline{\beta}$. Then $N^{\dagger}(\beta) < 1 = N^{nc}(\beta)$, and so for N close to $N^{nc}(\beta)$ we have

$$\frac{\partial \Pi^{fb}}{\partial N}(N, \beta) = \gamma(N) - K,$$

which under Assumption 2 is strictly negative for N sufficiently close to 1. Hence $N^{fb}(\beta) < N^{nc}(\beta)$.

Now suppose that $\beta \geq \overline{\beta}$. Then $N^{\ddagger}(\beta) > 0 = N^{nc}(\beta)$, and for N close to $N^{nc}(\beta)$ we have

$$\frac{\partial \Pi^{fb}}{\partial N}(N, \beta) = \gamma(N) - K + R\pi_0(\overline{\theta} - 1)\gamma(N).$$

Since $\gamma(N) > K$ for N close to zero, this expression is positive for N sufficiently small. Hence $N^{fb}(\beta) > N^{nc}(\beta)$.

Finally, suppose that $\beta \in [\underline{\beta}, \overline{\beta})$. Lemma F.1 established that $N^{nc}(\beta) \leq N^{\ddagger}(\beta)$. Over the range $[0, N^{\ddagger}(\beta))$,

$$\frac{\partial \Pi^{fb}}{\partial N}(N, \beta) = \gamma(N) - K + R\pi_0(\overline{\theta} - 1)\gamma(N),$$

which is continuous and decreasing in N and independent of β. Going forward, we'll suppress β from the argument of Π^{fb} when evaluating it at $N \leq N^{\ddagger}(\beta)$. Proposition 1 additionally established that $N^{nc}(\beta)$ is continuous and decreasing in β on the interval $[\underline{\beta}, \overline{\beta}]$. It follows that $\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\beta))$ is continuous and increasing in β on $[\underline{\beta}, \overline{\beta}]$. At the upper limit of this interval,

$$\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\overline{\beta})) = (1 + R\pi_0(\overline{\theta} - 1))\gamma(0) - K > 0.$$

Meanwhile at the lower end of the interval,

$$\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\underline{\beta})) = \Delta,$$

where $\Delta \equiv (1 + R\pi_0(\overline{\theta} - 1))\gamma(1) - K$.

Suppose first that $\Delta \geq 0$. Then $\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\beta)) > 0$ for all $\beta \in (\underline{\beta}, \overline{\beta})$. In that case strict concavity of $\Pi^{fb}(N)$ on the interval $[0, N^{\ddagger}(\beta)]$, established in Lemma C.1, ensures that $N^{fb}(\beta) > N^{nc}(\beta)$ for all $\beta \in (\underline{\beta}, \overline{\beta})$. Meanwhile when $\beta = \underline{\beta}$, we have $N^{nc}(\beta) = 1$ and $\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\beta)) = \Delta \geq 0$, in which case strict concavity along with the identity $N^{\ddagger}(\underline{\beta}) = 1$ implies that $N^{fb}(\beta) = 1 = N^{nc}(\beta)$. Then letting $\beta^{fb} = \underline{\beta}$, the claimed properties of $N^{fb}(\beta) - N^{nc}(\beta)$ hold.

Suppose instead that $\Delta < 0$. Then by continuity and strict monotonicity, there exists a unique $\beta^0 \in (\underline{\beta}, \overline{\beta})$ such that $\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\beta^0)) = 0$. For $\beta > \beta^0$ we have $\frac{\partial}{\partial N}\Pi^{fb}(N^{nc}(\beta), \beta) > 0$, in which case strict concavity of $\Pi^{fb}(N, \beta)$ in N over $[0, N^{\ddagger}(\beta)]$ ensures that $N^{fb}(\beta) > N^{nc}(\beta)$. Meanwhile for $\beta < \beta^0$ we have $\frac{\partial \Pi^{fb}}{\partial N}(N^{nc}(\beta)) < 0$, ensuring that also $\frac{\partial \Pi^{fb}}{\partial N}(N) < 0$ for all $N \in (N^{nc}(\beta), N^{\ddagger}(\beta))$. Then either $N^{fb}(\beta) \leq N^{\ddagger}(\beta)$ and $N^{nc}(\beta) > N^{\ddagger}(\beta)$, or else $N^{fb}(\beta) > N^{\ddagger}(\beta)$. But Lemma C.1 established that in the latter case, Π^{fb} is increasing in N on $[0, N^{\ddagger}(\beta)]$, a contradiction. So must be that $N^{fb}(\beta) > N^{\ddagger}(\beta)$. Then letting $\beta^{fb} = \beta^0$, the claimed properties of $N^{fb}(\beta) - N^{nc}(\beta)$ hold.

G Proof of Lemma 3

Throughout this proof, we assume for simplicity that $N^*(\beta)$ and $N^{fb}(\beta)$ are single-valued. If they aren't, the proof holds for any selection from each set of maximizers.

Let $\Pi^*(N,\beta)$ be the organization's profits under an optimal incentive scheme inducing risk-taking rate β, while $\Pi^{fb}(N,\beta)$ are its profits in an environment without incentive constraints. In general $\Pi^*(N,\beta) \leq \Pi^{fb}(N,\beta)$, and the inequality is strict for any $N \neq N^{nc}(\beta)$, since to induce any such N the organization must either promote inefficiency or pay strictly positive bonuses or both. It follows immediately that $N^*(\beta) = N^{fb}(\beta)$ in case $N^{fb}(\beta) = N^{nc}(\beta)$.

Suppose that $N^{fb}(\beta) > N^{nc}(\beta)$. This hypothesis implies that $N^{nc}(\beta) < 1$, so that $\beta > \underline{\beta}$. Lemma C.1 established that for $\beta \geq \underline{\beta}$, $\Pi^{fb}(N,\beta)$ is strictly concave on $[0, N^{\ddagger}(\beta)]$, and if $N^{fb}(\beta) \geq N^{\ddagger}(\beta)$ then $\Pi^{fb}(N,\beta)$ is increasing in N on $[0, N^{\ddagger}(\beta)]$. Meanwhile, Lemma F.1 established that $N^{nc}(\beta) \leq N^{\ddagger}(\beta)$. The hypothesis $N^{fb}(\beta) > N^{nc}(\beta)$ therefore implies that $\Pi^{fb}(N,\beta)$ is increasing on $[0, N^{nc}(\beta)]$, so that $\Pi^{fb}(N,\beta) < \Pi^{fb}(N^{nc}(\beta),\beta)$ for all $N < N^{nc}(\beta)$. It follows that $\Pi^*(N,\beta) < \Pi^{fb}(N^{nc}(\beta),\beta) = \Pi^*(N^{nc}(\beta),\beta)$ for all such N, establishing $N^*(\beta) \geq N^{nc}(\beta)$.

Finally, suppose that $N^{fb}(\beta) < N^{nc}(\beta)$. This hypothesis implies in particular that $N^{fb}(\beta) < N^{\ddagger}(\beta)$. If $\beta \geq \underline{\beta}$, then Lemma C.1 implies that $\Pi^{fb}(N,\beta)$ is decreasing in N on $[N^{nc}(\beta), 1]$, so that $\Pi^{fb}(N,\beta) < \Pi^{fb}(N^{nc}(\beta),\beta)$ for all $N > N^{nc}(\beta)$. It follows that $\Pi^*(N,\beta) < \Pi^{fb}(N^{nc}(\beta),\beta) = \Pi^*(N^{nc}(\beta),\beta)$ for all such N, establishing $N^*(\beta) \leq N^{nc}(\beta)$. If on the other hand $\beta < \underline{\beta}$, then Lemma C.1 established that $\Pi^{fb}(N,\beta)$ is strictly concave in N on $[0,1]$. The hypothesis $N^{fb}(\beta) < N^{nc}(\beta)$ therefore implies that $\Pi^{fb}(N,\beta) < \Pi^{fb}(N^{nc}(\beta),\beta)$ for all $N > N^{nc}(\beta)$. Hence $\Pi^*(N,\beta) < \Pi^{fb}(N^{nc}(\beta),\beta) = \Pi^*(N^{nc}(\beta),\beta)$ for all such N, establishing $N^*(\beta) \leq N^{nc}(\beta)$.

H Proof of Proposition 3

We first characterize an optimal policy for $N \in (0, N^{nc}(\beta))$, and return to the extremal case $N = 0$ at the end of the proof. When N is interior, the risk-taking set $[0, N]$ is incentive-compatible if and only if 1) the marginal employee is indifferent over his approach, and 2) the payoff from successful risk-taking exceeds the payoff from failed risk-taking. Formally, these constraints are

$$\gamma(N)(V\sigma_G + T_G) + (1-\gamma(N))(V\sigma_B + T_B) = V\sigma_0 + T_0, \qquad V\sigma_G + T_G \geq V\sigma_B + T_B.$$

We solve the relaxed problem enforcing only the marginal employee's IC constraint, and verify that the resulting optimal scheme satisfies the remaining constraint.

We first show that any scheme satisfying $T_G > 0$ is suboptimal. Fix such a scheme. Define a new bonus scheme $\boldsymbol{T}'$ by $T'_G = 0$, $T'_B = T_B + \frac{\gamma(N)}{1-\gamma(N)} T_G$, and $T'_0 = T_0$. This modification does not disturb the relaxed IC constraint, and pays out total bonuses

$$B' = (N - \mu(N)) \frac{\gamma(N)}{1-\gamma(N)} T_G + (N - \mu(N)) T_B + (1 - N) T_0.$$

Now, for any $n > 0$ we have $\gamma(n) < \mu(n)/n$ given that $\mu(n)/n$ is the average of γ over $[0, n]$, and γ is strictly decreasing. Hence

$$\frac{\gamma(N)}{1-\gamma(N)} < \frac{\mu(N)/N}{1-\mu(N)/N} = \frac{\mu(N)}{N-\mu(N)},$$

and so

$$B' < \mu(N) T_G + (N - \mu(N)) T_B + (1 - N) T_0 = B,$$

where B are total bonus payments under the original scheme. This modification to the bonus scheme therefore reduces total bonus payments and increases profits.

Going forward, we restrict attention to schemes satisfying $T_G = 0$. We next show that any scheme satisfying $T_B > 0$ is suboptimal. Fix such a scheme.
Case 1: $T_0 > 0$. Pass to the modified bonus scheme $\boldsymbol{T}'$ satisfying $T'_B = T_B - (1 - \gamma(N))\Delta$ and $T'_0 = T_0 - \Delta$, which satisfies the relaxed IC constraint for any Δ, and for $\Delta > 0$ sufficiently small also satisfies the non-negativity constraint on bonuses. This modification strictly reduces bonuses payments and therefore increases profits.
Case 2: $T_0 = 0$. In this case,

$$\gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B - \sigma_0 = -(1 - \gamma(N)) T_B / V < 0.$$

This inequality immediately implies $\sigma_0 > 0$, given that $\sigma_G, \sigma_B \geq 0$. It also implies that $\sigma_G < 1$, by the following lemma.

Lemma H.1. If the promotion scheme $\boldsymbol{\sigma}$ is feasible and $\sigma_G = 1$, then

$$\gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B - \sigma_0 > 0.$$

Proof. $N < N^{nc}(\beta) \leq N^{\ddagger}(\beta)$ implies that $\sigma_B^{nc}(N, \beta) = 0$, so $\sigma_B \geq \sigma_B^{nc}(N, \beta)$. If $\sigma_G = 1$, then feasibility additionally requires that $N \leq N^{\dagger}(\beta)$, in which case $\sigma_G = \sigma_G^{nc}(N, \beta)$. This identity combined with $\sigma_B \geq \sigma_B^{nc}(N, \beta)$ then implies $\sigma_0 \leq \sigma_0^{nc}(N, \beta)$ if feasibility is to be satisfied. Thus

$$\gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B - \sigma_0 \geq \gamma(N)\sigma_G^{nc}(N, \beta) + (1 - \gamma(N))\sigma_B^{nc}(N, \beta) - \sigma_0^{nc}(N, \beta).$$

Meanwhile $N < N^{nc}(\beta)$ implies that the rhs of the previous inequality is positive, yielding the desired bound. $\square$

Pass to the modified bonus and promotion scheme $(\boldsymbol{T}', \boldsymbol{\sigma}')$ which sets $\sigma_0' = \sigma_0 - \Delta$, $\sigma_G' = \sigma_G + \Delta(1-N)/\mu(N)$, and

$$T_B' = T_B - V \frac{\Delta}{1-\gamma(N)} \left(1 + (1-N)\frac{\gamma(N)}{\mu(N)}\right),$$

with all other promotion probabilities and bonuses unchanged. By construction, this modified scheme preserves the relaxed IC constraint and the number of employees promoted. Therefore for any $\Delta > 0$, it reallocates promotions from safe approaches to successful risk-taking, strictly increasing promotion payoffs given that $\pi_G > \pi_0$. Finally, for $\Delta > 0$ sufficiently small the modified scheme respects the non-negativity constraint on bonuses given that $T_B > 0$ and the boundary constraints on promotion probabilities given that $\sigma_0 > 0$ and $\sigma_G < 1$. So this modification is feasible and increases profits.

Going forward, we restrict attention to schemes satisfying $T_B = 0$. We now show that any scheme satisfying $\sigma_B > 0$ is suboptimal. Fix such a scheme. The relaxed IC constraint may be rearranged to read

$$\gamma(N)\sigma_G + (1-\gamma(N))\sigma_B - \sigma_0 = T_0/V \geq 0,$$

implying that $\sigma_0 < 1$ given that $\sigma_G, \sigma_B \leq 1$ and $\sigma_G = \sigma_B = \sigma_0 = 1$ violates feasibility.

Case 1: $T_0 > 0$. Pass to the modified bonus and promotion schemes $(\boldsymbol{T}', \boldsymbol{\sigma}')$, where $\sigma_0' = \sigma_0 + \Delta$, $\sigma_B' = \sigma_B - \Delta(1-N)/(N-\mu(N))$, $\sigma_G' = \sigma_G$, and

$$T_0' = T_0 - V\Delta \left(1 + (1-N)\frac{1-\gamma(N)}{N-\mu(N)}\right).$$

By construction, this modified scheme preserves the relaxed IC constraint and the number of promoted employees for any Δ. Therefore for $\Delta > 0$ it shifts promotions from failed risk-takers toward non-risk-takers, increasing total promotion payoffs given that $\pi_0 > \pi_B(N)$. It additionally reduces total bonus payments. Finally, for $\Delta > 0$ sufficiently small it respects the bonus non-negativity constraint and the boundary constraints on promotion probabilities given that $\sigma_0 < 1$ and $\sigma_B > 0$. So this modification is feasible and increases profits.

Case 2: $T_0 = 0$ and $\sigma_0 = 0$. In this case the relaxed IC constraint requires that $\sigma_G = \sigma_B = 0$ as well. But then the modified promotion scheme $\boldsymbol{\sigma}' = \boldsymbol{\sigma} + \Delta$ preserves the relaxed IC constraint for any Δ and remains feasible for sufficiently small $\Delta > 0$. Since all promoted employees yield a positive payoff for the organization, this modification raises profits.

Case 3: T_0 and $\sigma_0 > 0$. In this case Lemma H.1 additionally implies that $\sigma_G < 1$. Consider the modified promotion scheme $\boldsymbol{\sigma}'$, where

$$\sigma_B' = \sigma_B - \Delta, \quad \sigma_0' = \sigma_0 - \frac{\rho(N) - \gamma(N)}{\rho(N)}\Delta, \quad \sigma_G' = \sigma_G + \frac{1 - \rho(N)}{\rho(N)}\Delta.$$

By construction, this modification preserves both the relaxed IC constraint and the number of promoted employees. Further, $\rho(N) = \gamma(N) + \mu(N) - N\gamma(N)$, and as noted earlier $\mu(N) > N\gamma(N)$. Hence $\rho(N) > \gamma(N)$. So for all $\Delta > 0$ this modification reallocates promotions from failed and non-risk-takers to successful risk-takers, increasing total promotion payoffs given that $\pi_G > \pi_0, \pi_B(N)$. Finally, for $\Delta > 0$ sufficiently small this modification respects the boundary constraints on promotion probabilities given that $\sigma_B, \sigma_0 > 0$ while $\sigma_G < 1$. So this modification is feasible and increases profits.

Going forward, we will restrict attention to schemes satisfying $\sigma_B = 0$. We next show that any scheme which does not promote β employees is suboptimal. Fix such a scheme. Note that the relaxed IC constraint reads $\gamma(N)\sigma_G = \sigma_0 + T_0/V$, which combined with $\gamma(N) > 0$ implies that either $\sigma_G, \sigma_0 < 1$ or else $1 = \sigma_G > \sigma_0$.

Case 1: $\sigma_G, \sigma_0 < 1$. Pass to the modified scheme $(\sigma_G', \sigma_0') = (\sigma_G + \Delta, \sigma_0 + \gamma(N)\Delta)$, which preserves the relaxed IC constraint and, for sufficiently small $\Delta > 0$, remains feasible. Since this modification increases the number of promoted employees, and since every promoted employee yields a positive payoff to the organization, this modification must increase profits.

Case 2: $1 = \sigma_G > \sigma_0$. In this case Lemma H.1 implies that $T_0 > 0$. Then a modified scheme setting $\sigma_0' = \sigma_0 + \Delta$ and $T_0' = T_0 - V\Delta$ preserves the relaxed IC constraint, and for sufficiently small $\Delta > 0$ it increases the number of employees while remaining feasible. Since every promoted employee yields a positive payoff to the organization, and since the modified scheme additionally reduces bonus payments, this modification must increase profits. Going forward we restrict attention to schemes which saturate the feasibility constraint.

To complete the characterization of an optimal scheme, we enforce $T_G = T_B = \sigma_B = 0$ and solve the optimization problem

$$\max_{\sigma_G, \sigma_0, T_0} R(\mu(N)\pi_G\sigma_G + (1 - N)\pi_0\sigma_0) - (1 - N)T_0$$

subject to the boundary constraints $\sigma_G, \sigma_0 \in [0, 1]$ and $T_0 \geq 0$, the relaxed IC constraint

$$\gamma(N)\sigma_G = \sigma_0 + T_0/V,$$

and the binding feasibility constraint

$$\beta = \mu(N)\sigma_G + (1 - N)\sigma_0.$$

Notice that any solution to this problem trivially satisfies the remaining IC constraint $V\sigma_G + T_G \geq V\sigma_B + T_B$ given that $\sigma_B = T_B = 0$.

Solving the feasibility and IC constraints for σ_0 and T_0 yields

$$\sigma_0(\sigma_G) = \frac{\beta - \mu(N)\sigma_G}{1-N}, \quad T_0(\sigma_G) = V\frac{\rho(N)\sigma_G - \beta}{1-N}.$$

Using these expressions to eliminate σ_0 and T_0 from the maximization problem yields

$$\max_{\sigma_G} \beta(V + R\pi_0) + (R\mu(N)(\pi_G - \pi_0) - V\rho(N))\,\sigma_G$$

subject to the boundary constraints that $\sigma_G, \sigma_0 \in [0,1]$ and $T_0 \geq 0$.

Note that the boundary constraints on σ_0 and T_0 implicitly place additional constraints on σ_G, given that each is a function of σ_G. They collectively imply that $\sigma_G \in [\underline{\sigma}_G, \overline{\sigma}_G]$, where

$$\overline{\sigma}_G \equiv \min\left\{\frac{\beta}{\mu(N)}, 1\right\}, \quad \underline{\sigma}_G \equiv \max\left\{\frac{\beta - (1-N)}{\mu(N)}, \frac{\beta}{\rho(N)}\right\}.$$

Since the reduced objective is affine in σ_G with slope

$$\xi_-(N) \equiv R\mu(N)(\pi_G - \pi_0) - V\rho(N),$$

the optimal value of σ_G is therefore

$$\sigma_G^* = \begin{cases} \overline{\sigma}_G & \text{if } \xi_-(N) > 0 \\ \underline{\sigma}_G & \text{if } \xi_-(N) < 0. \end{cases}$$

(If $\xi_-(N) = 0$, then there exist a continuum of optimal schemes.)

We now characterize the sign of ξ_- as a function of N. Since ρ is strictly decreasing while μ is strictly increasing, ξ_- is strictly increasing. Further, $\xi_-(0) = -V\rho(0) = -V\gamma(0) < 0$, while $\xi_-(1) = \mu(1)(R(\pi_G - \pi_0) - V)$. Let

$$\overline{N}_-(R,V) \equiv \sup\{N \in [0,1] : \xi_-(N) < 0\}.$$

Given that ξ_- is strictly increasing in N, $N < \overline{N}_-(R,V)$ implies that $\xi_-(N) < 0$ while $N > \overline{N}_-(R,V)$ implies that $\xi_-(N) > 0$.

To establish the claimed comparative statics of $\overline{N}_-(R,V)$ in R and V, first note that $\overline{N}_-(R,V)$ depends on R, V only through the ratio V/R. If $V/R \geq \pi_G - \pi_0$, then $\xi_-(N) \leq 0$ for all N and $\overline{N}_-(R,V) = 1$. Meanwhile if $V/R < \pi_G - \pi_0$, then $\xi_-(1) > 0$ and $\overline{N}_-(R,V) \in (0,1)$. Further, $\xi_-(N)$ is increasing in R and decreasing in V for all $N > 0$, implying that $\overline{N}_-(R,V)$ is increasing in V/R whenever it is interior. Finally, for every $N > 0$, $\xi_-(N) > 0$ for V/R

sufficiently small. Thus $\lim_{V/R \to 0} \overline{N}_-(R,V) = 0$. The comparative statics with respect to R and V follow immediately from this analysis.

We next show that the scheme satisfying $\sigma_G^* = \overline{\sigma}_G$ corresponds to efficient promotion and a positive bonus for safe approaches. Recall the efficient promotion probabilities σ_i^{nc} characterized in Appendix C. For all N we have $\overline{\sigma}_G = \sigma_G^{nc}(N,\beta)$. Meanwhile $N < N^{nc}(\beta) \leq N^{\ddagger}(\beta)$ implies that $\sigma_0^{nc}(N,\beta) = (\beta - \mu(N)\sigma_G^{nc}(N,\beta))/(1-N) = \sigma_0(\overline{\sigma}_G)$ and $\sigma_B^{nc}(N,\beta) = 0 = \sigma_B$. So this scheme promotes efficiently.

Meanwhile $T_0(\overline{\sigma}_G) > 0$ iff $\overline{\sigma}_G > \beta/\rho(N)$. If $N \geq N^{\dagger}(\beta)$, then $\overline{\sigma}_G = \beta/\mu(N)$. Since $\rho(N) > \mu(N)$ for all $N < 1$, the desired inequality holds in this case. If instead $N < N^{\dagger}(\beta)$, then $\overline{\sigma}_G = 1$. In this case the desired inequality amounts to $\rho(N) > \beta$. Suppose first that $\beta \leq \underline{\beta}$. Then $N^{nc}(\beta) = 1$, and since ρ is strictly decreasing in N, the desired inequality holds for $N < N^{nc}(\beta)$ if $\rho(1) \geq \beta$. Since $\rho(1) = \mu(1) = \underline{\beta}$, the result follows. Next suppose that $\beta \in (\underline{\beta}, \overline{\beta})$. Then $N^{nc}(\beta)$ satisfies $\rho(N^{nc}(\beta)) = \beta$, and so $\rho(N) > \beta$ for all $N < N^{nc}(\beta)$. Finally, if $\beta \geq \overline{\beta}$ then $N^{nc}(\beta) = 0$ and there are no risk-taking rates strictly below $N^{nc}(\beta)$. So the bonus is strictly positive in all cases.

We now show that the scheme satisfying $\sigma_G^* = \underline{\sigma}_G$ underpromotes successful risk-takers, overpromotes non-risk-takers, and pays no bonuses. Since $\overline{\sigma}_G > \underline{\sigma}_G$, $\sigma_0(\sigma_G)$ is decreasing in σ_G, and $\overline{\sigma}_G$ induces efficient promotion, the results about promotion follow immediately. The zero bonus result follows from the following lemma, which ensures that $\underline{\sigma}_G = \beta/\rho(N)$.

Lemma H.2. $\frac{\beta-(1-n)}{\mu(n)} < \frac{\beta}{\rho(n)}$ for all $n \in (0,1]$.

Proof. Some algebra shows that this inequality is equivalent to $\zeta(n) > 0$, where $\zeta(n) \equiv \mu(n) + \gamma(n)(1-n-\beta)$. Differentiating this expression yields $\zeta'(n) = \gamma'(n)(1-n-\beta)$, which crosses zero from exactly once at $n = 1-\beta$, from below. Hence ζ is minimized at $n = 1-\beta$. Evaluating ζ at this point yields $\zeta(1-\beta) = \mu(1-\beta) > 0$, so ζ is positive everywhere. □

We complete the proof by returning to the extremal case $N = 0$. We analyze this case by taking the limit of the optimal scheme for $N > 0$ and invoking the maximum theorem. The organization's objective function is continuous in $(N, \boldsymbol{T}, \boldsymbol{\sigma})$, and the set of feasible, IC incentive schemes is characterized by a set of equalities and weak inequalities which are each continuous in N. Thus the constraint correspondence is continuous in N. It is not formally compact-valued, as transfers are unbounded. However, it is easy to show that placing a sufficiently large bound on transfers, uniformly for all N, does not change the optimal scheme for any N. (See the proof of Lemma 1 for a detailed argument.) Thus it is without loss to pass to the modified problem with a sufficiently large bound on transfers. The maximum theorem may then be invoked to conclude that our characterized optimal incentive scheme for $N > 0$ remains optimal in the limit $N = 0$.

I Proof of Propositions 4 and 5

We first characterize an optimal policy for $N \in (N^{nc}(\beta), 1)$, and return to the extremal case $N = 1$ at the end of the proof. When N is interior, the risk-taking set $[0, N]$ is incentive-compatible if and only if 1) the marginal employee is indifferent over his approach, and 2) the payoff from successful risk-taking exceeds the payoff from failed risk-taking. Formally, these constraints are

$$\gamma(N)(V\sigma_G + T_G) + (1 - \gamma(N))(V\sigma_B + T_B) = V\sigma_0 + T_0, \qquad V\sigma_G + T_G \geq V\sigma_B + T_B.$$

We solve the relaxed problem enforcing only the marginal employee's IC constraint, and verify that the resulting optimal scheme satisfies the remaining constraint.

We first observe that any scheme satisfying $T_G > 0$ is suboptimal. This result follows from an argument identical to the one used in the proof of Proposition 3. Going forward, we restrict attention to schemes satisfying $T_G = 0$.

We next show that any scheme satisfying $T_0 > 0$ is suboptimal. Fix any such scheme.

Case 1: $T_B > 0$. Pass to the modified bonus scheme $\boldsymbol{T}'$ satisfying $T_0' = T_0 - (1-\gamma(N))\Delta$ and $T_B' = T_B - \Delta$, which satisfies the relaxed IC constraint for any Δ, and for $\Delta > 0$ sufficiently small also satisfies the non-negativity constraint on bonuses. This modification therefore strictly reduces bonuses payments and increases profits.

Case 2: $T_B = 0$. In this case

$$\gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B - \sigma_0 = T_0/V > 0.$$

This inequality along with the bounds $\sigma_G, \sigma_B \leq 1$ imply that $\sigma_0 < 1$.

Subcase A: The feasibility constraint is slack. Pass to the modified bonus and promotion schemes $(\boldsymbol{T}', \boldsymbol{\sigma}')$, where $\sigma_0' = \sigma_0 + \Delta$, $T_0' = T_0 - V\Delta$, and all other promotion probabilities and bonuses unchanged. This modified scheme preserves the relaxed IC constraint, and for $\Delta > 0$ sufficiently small it remains feasible; satisfies the non-negativity constraint on bonuses and the boundary constraints on promotion probabilities; reduces bonus payments; and promotes more employees than the original scheme. Since every employee promoted yields a positive profit to the organization, this modification improves on the original scheme.

Subcase B: The feasibility constraint is saturated. Then the relaxed IC constraint implies that $\sigma_B > 0$, by the following lemma.

Lemma I.1. If the promotion scheme $\boldsymbol{\sigma}$ promotes β employees and $\sigma_B = 0$, then

$$\gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B - \sigma_0 < 0.$$

Proof. Since efficient promotion maximizes σ_G subject to feasibility, any feasible scheme must satisfy $\sigma_G \leq \sigma_G^{nc}(N,\beta)$. If $\sigma_B = 0$ and the feasibility constraint is saturated, then it must additionally be that $\sigma_0 \geq \sigma_0^{nc}(N,\beta)$. Thus

$$\gamma(N)\sigma_G + (1-\gamma(N))\sigma_B - \sigma_0 \leq \gamma(N)\sigma_G^{nc}(N,\beta) + (1-\gamma(N))\sigma_B^{nc}(N,\beta) - \sigma_0^{nc}(N,\beta).$$

Meanwhile $N > N^{nc}(\beta)$ implies that the rhs of the previous inequality is negative, yielding the desired bound. $\square$

Pass to the modified bonus and promotion scheme $(\boldsymbol{T'}, \boldsymbol{\sigma'})$ which sets $\sigma_0' = \sigma_0 + \Delta$, $\sigma_B' = \sigma_B - \Delta(1-N)/(N-\mu(N))$, and

$$T_0' = T_0 - V\Delta\left(1 + (1-N)\frac{1-\gamma(N)}{N-\mu(N)}\right),$$

with all other promotion probabilities and bonuses unchanged. By construction, this modified scheme preserves the relaxed IC constraint and the number of employees promoted. Therefore for any $\Delta > 0$, it reallocates promotions from failed risk-takers to non-risk-takers, strictly increasing promotion payoffs given that $\pi_0 > \pi_B(N)$. Finally, for $\Delta > 0$ sufficiently small the modified scheme respects the non-negativity constraint on bonuses given that $T_0 > 0$ and the boundary constraints on promotion probabilities given that $\sigma_0 > 0$ and $\sigma_G < 1$. So this modification is feasible and increases promotion payoffs while decreasing bonus payments, increasing total profits.

Going forward, we restrict attention to schemes satisfying $T_0 = 0$. We now show that any scheme satisfying $\sigma_G < 1$ is suboptimal. Fix any such scheme. The relaxed IC constraint may be rearranged to read

$$\gamma(N)\sigma_G + (1-\gamma(N))\sigma_B - \sigma_0 = -(1-\gamma(N))T_B/V \leq 0.$$

Case 1: $\sigma_0 = 0$. In this case the relaxed IC constraint implies that $\sigma_G = \sigma_B = 0$ as well. So pass to the modified promotion scheme $\boldsymbol{\sigma'} = \boldsymbol{\sigma} + \Delta$. This modified scheme preserves the relaxed IC constraint for any Δ, and for $\Delta > 0$ sufficiently small it promotes more employees and remains feasible. Since every promoted employee yields a positive payoff to the organization, this modification improves payoffs.

Case 2: $\sigma_0 > 0$ and $T_B > 0$. Pass to the modified bonus and promotion schemes $(\boldsymbol{T'}, \boldsymbol{\sigma'})$, where $\sigma_0' = \sigma_0 - \Delta$, $\sigma_G' = \sigma_G + \Delta(1-N)/\mu(N)$, $\sigma_B' = \sigma_B$, and

$$T_B' = T_B - \frac{V\Delta}{1-\gamma(N)}\left(1 + (1-N)\frac{\gamma(N)}{\mu(N)}\right).$$

By construction, this modified scheme preserves the relaxed IC constraint and the number of promoted employees for any Δ. Therefore for $\Delta > 0$ it shifts promotions from non-risk-takers toward successful risk-takers, increasing total promotion payoffs given that $\pi_G > \pi_0$.

It additionally reduces total bonus payments. Finally, for $\Delta > 0$ sufficiently small it respects the bonus non-negativity constraint and the boundary constraints on promotion probabilities given that $\sigma_G < 1$ and $\sigma_0 > 0$. So this modification is feasible and increases profits.
Case 3: $\sigma_0 > 0$ and $T_B = 0$. Then the relaxed IC constraint reads

$$\gamma(N)\sigma_G + (1-\gamma(N))\sigma_B - \sigma_0 = 0.$$

Since $\sigma_G < 1$, this constraint implies that also $\sigma_0 < 1$.
Subcase A: The feasibility constraint is not saturated. Pass to the modified promotion scheme $\boldsymbol{\sigma}'$ satisfying $\sigma'_G = \sigma_G + \Delta$, $\sigma'_0 = \sigma_0 + \gamma(N)\Delta$, and $\sigma'_B = \sigma_B$. This modification preserves the relaxed IC constraint, and for sufficiently small $\Delta > 0$ it raises the number of employees promoted, preserves feasibility, and satisfies the promotion probability boundary constraints given that $\sigma_G, \sigma_0 < 1$. Since every promoted employee yields positive profits for the organization, this modification increases profits.
Subcase B: The feasibility constraint is saturated. In this case Lemma I.1 along with the relaxed IC constraint imply that $\sigma_B > 0$. So pass to the modified promotion scheme $\boldsymbol{\sigma}'$, where

$$\sigma'_B = \sigma_B - \Delta, \quad \sigma'_0 = \sigma_0 - \frac{\rho(N)-\gamma(N)}{\rho(N)}\Delta, \quad \sigma'_G = \sigma_G + \frac{1-\rho(N)}{\rho(N)}\Delta.$$

By construction, this modification preserves both the relaxed IC constraint and the number of promoted employees. Further, $\rho(N) = \gamma(N) + \mu(N) - N\gamma(N)$, and since $\mu(N)/N$ is an average of γ over the interval $[0, N]$, $\mu(N) > N\gamma(N)$. Hence $\rho(N) > \gamma(N)$. So for all $\Delta > 0$ this modification reallocates promotions from failed and non-risk-takers to successful risk-takers, increasing total promotion payoffs given that $\pi_G > \pi_0, \pi_B(N)$. Finally, for $\Delta > 0$ sufficiently small this modification respects the boundary constraints on promotion probabilities given that $\sigma_G < 1$ and $\sigma_0 > 0$ by hypothesis while $\sigma_B > 0$ as established above. So this modification is feasible and increases profits.

Going forward, we will restrict attention to schemes satisfying $\sigma_G = 1$. We next show that any scheme which does not promote β employees is suboptimal. Fix such a scheme. Note that the relaxed IC constraint reads

$$\gamma(N) + (1-\gamma(N))(\sigma_B + T_B/V) = \sigma_0,$$

which combined with $\gamma(N) < 1$ implies that either $\sigma_B = \sigma_0 = 1$ or else $\sigma_0 > \sigma_B$. The first possibility violates feasibility, so assume the latter inequality.
Case 1: $\sigma_0 < 1$. Pass to the modified scheme $(\sigma'_B, \sigma'_0) = (\sigma_B + \Delta, \sigma_0 + (1-\gamma(N))\Delta)$, which preserves the relaxed IC constraint and, for sufficiently small $\Delta > 0$, remains feasible. Since this modification increases the number of promoted employees, and since every promoted employee yields a positive payoff to the organization, this modification must increase profits.

Case 2: $\sigma_0 = 1$. In this case the relaxed IC constraint combined with $\sigma_B < \sigma_0$ imply that $T_B > 0$. Passing to a modified scheme setting $\sigma'_B = \sigma_B + \Delta$ and $T'_B = T_B - V\Delta$ preserves the relaxed IC constraint, and for sufficiently small $\Delta > 0$ it increases the number of employees promoted while remaining feasible and respecting the promotion probability boundary constraints. Since every promoted employee yields a positive payoff to the organization, and since the modified scheme additionally decreases bonuses, this modification must increase profits. Going forward we restrict attention to schemes which saturate the feasibility constraint.

To complete the characterization of an optimal scheme, we enforce $T_G = T_0 = 0$ and $\sigma_G = 1$ and solve the optimization problem

$$\max_{\sigma_B, \sigma_0, T_B} R(\mu(N)\pi_G + (N - \mu(N))\pi_B(N)\sigma_B + (1 - N)\pi_0\sigma_0) - (N - \mu(N))T_B$$

subject to the boundary constraints $\sigma_B, \sigma_0 \in [0, 1]$ and $T_B \geq 0$, the relaxed IC constraint

$$\gamma(N) + (1 - \gamma(N))(\sigma_B + T_B/V) = \sigma_0,$$

and the binding feasibility constraint

$$\beta = \mu(N) + (N - \mu(N))\sigma_B + (1 - N)\sigma_0.$$

Note that the relaxed IC constraint combined with the upper bound $\sigma_0 \leq 1$ implies that $\sigma_B + T_B/V \leq 1 = \sigma_G$. Hence any solution to this problem automatically satisfies the remaining IC constraint $V\sigma_G + T_G \geq V\sigma_B + T_B$, as claimed.

Solving the feasibility and IC constraints for σ_0 and T_B yields

$$\sigma_0(\sigma_B) = \frac{\beta - \mu(N) - (N - \mu(N))\sigma_B}{1 - N}, \quad T_B(\sigma_B) = V\frac{\beta - \rho(N) - (1 - \rho(N))\sigma_B}{(1 - N)(1 - \gamma(N))}.$$

Using these expressions to eliminate σ_0 and T_B from the maximization problem yields, up to an additive constant which does not affect the solution,

$$\max_{\sigma_B} \frac{N - \mu(N)}{(1 - N)(1 - \gamma(N))} \left(V(1 - \rho(N)) - R(\pi_0 - \pi_B(N))(1 - N)(1 - \gamma(N))\right)\sigma_G$$

subject to the boundary constraints that $\sigma_B, \sigma_0 \in [0, 1]$ and $T_B \geq 0$.

The boundary constraints on σ_0 and T_B implicitly place additional constraints on σ_B, given that each is a function of σ_B. They collectively imply that $\sigma_B \in [\underline{\sigma}_B, \overline{\sigma}_B]$, where

$$\overline{\sigma}_B \equiv \min\left\{\frac{\beta - \mu(N)}{N - \mu(N)}, \frac{\beta - \rho(N)}{1 - \rho(N)}\right\}, \quad \underline{\sigma}_B \equiv \max\left\{\frac{\beta - \nu(N)}{N - \mu(N)}, 0\right\},$$

where $\nu(N) = \mu(N) + (1 - N)$ is as defined in Appendix C. Since the reduced objective is linear in σ_B with a slope of the same sign as

$$\xi_+(N) \equiv V(1 - \rho(N)) - R(\pi_0 - \pi_B(N))(1 - N)(1 - \gamma(N)),$$

the optimal value of σ_B is therefore

$$\sigma_B^* = \begin{cases} \overline{\sigma}_B & \text{if } \xi_+(N) > 0 \\ \underline{\sigma}_B & \text{if } \xi_+(N) < 0. \end{cases}$$

(If $\xi_+(N) = 0$, then there exist a continuum of optimal schemes.)

We now characterize the sign of ξ_+ as a function of N. Let $\Delta\pi(N) \equiv \pi_0 - \pi_B(N)$. Lemma B.1 established that $\pi_B(N)$ is increasing in N, and so $\Delta\pi_+(N)$ is decreasing in N. Note that ξ_+ satisfies the boundary conditions $\xi_+(1) = 1 - \mu(1) > 0$ and $\xi_+(0) = (1 - \gamma(0))(V - R\Delta\pi(0))$. Suppose first that $R \leq V/\Delta\pi(0)$. Then for all $N \in (0, 1)$,

$$\begin{aligned} \xi_+(N) &\geq V(1 - \rho(N)) - V\frac{\Delta\pi(N)}{\Delta\pi(0)}(1 - N)(1 - \gamma(N)) \\ &> V(1 - \rho(N) - (1 - N)(1 - \gamma(N))) \\ &= V(N - \mu(N)) > 0. \end{aligned}$$

Suppose instead that $R > V/\Delta\pi(0)$. Then $\xi_+(0) < 0$ given that $\gamma(0) < 1$, meaning ξ_+ is negative for N sufficiently close to 0 and positive for N sufficiently close to 1. Suppose further that $(1 - N)(1 - \gamma(N))$ is nonincreasing. Then the fact that ρ and $\Delta\pi(N)$ are both decreasing implies ξ_+ is increasing in N. Given the boundary conditions $\xi_+(0) < 0 < \xi_+(1)$, it follows that ξ_+ crosses zero exactly once.

Let

$$\overline{N}_+(R, V) \equiv \inf\{N \in [0, 1] : \xi_+(N) > 0\}.$$

To establish the claimed comparative statics of $\overline{N}_+(R, V)$ in R and V, first note that $\overline{N}_+(R, V)$ depends on R, V only through the ratio V/R. If $V/R \geq \Delta\pi(0)$, then $\xi_+(N) \geq 0$ for all N and $\overline{N}_+(R, V) = 0$. Meanwhile if $V/R < \Delta\pi(0)$, then $\xi_+(0) < 0$ and $\overline{N}_+(R, V) \in (0, 1)$. Further, $\xi_+(N)$ is increasing in V and decreasing in R for all $N < 1$, implying that $\overline{N}_+(R, V)$ is decreasing in V/R whenever it is interior. Finally, for every $N > 0$, $\xi_+(N) < 0$ for V/R sufficiently small. Thus $\lim_{V/R \to 0} \overline{N}_+(R, V) = 1$. The comparative statics with respect to R and V follow immediately from this analysis.

We next show that the scheme satisfying $\sigma_B^* = \underline{\sigma}_B$ corresponds to efficient promotion and a positive bonus for failed risk-taking. Recall the efficient promotion probabilities σ_i^{nc} characterized in Appendix C. Given that $N > N^{nc}(\beta)$, we must have $N^{nc}(\beta) < 1$, implying

$\beta > \underline{\beta}$. For all such β, $\sigma_B^{nc}(N,\beta) = \underline{\sigma}_B$ and $\sigma_G^{nc}(N,\beta) = 1 = \sigma_G$. The fact that the optimal promotion scheme and efficient promotion both saturate the feasibility constraint then ensures that $\sigma_0(\underline{\sigma}_B) = \sigma_0^{nc}(N,\beta)$. So this scheme promotes efficiently.

Meanwhile $T_B(\underline{\sigma}_B) > 0$ iff $\underline{\sigma}_B < (\beta - \rho(N))/(1-\rho(N))$. We first establish that the rhs of this bound is strictly positive for all $N > N^{nc}(\beta)$. If $\beta \geq \overline{\beta}$, then $N^{nc}(\beta) = 0$, and since ρ is strictly decreasing, $\beta - \rho(N) > \overline{\beta} - \rho(0) = \overline{\beta} - \gamma(0) = 0$. If $\beta \in (\underline{\beta}, \overline{\beta})$, then $\rho(N^{nc}(\beta)) = \beta$, and since ρ is strictly decreasing we have $\beta - \rho(N) > 0$. Finally, if $\beta \leq \underline{\beta}$, then as observed above there exist no $N > N^{nc}(\beta)$. So $\beta > \rho(N)$ in all cases, ensuring that the rhs of the bound is positive. If $N \leq N^{\ddagger}(\beta)$, then $\underline{\sigma}_B = 0$, and so the desired bound holds given positivity of the rhs. On the other hand, if $N > N^{\ddagger}(\beta)$, then $\underline{\sigma}_B = (\beta - \nu(N))/(N - \mu(N))$, and the following lemma establishes the desired result that the bonus is strictly positive.

Lemma I.2. $(\beta - \nu(n))/(n - \mu(n)) < (\beta - \rho(n))/(1 - \rho(n))$ for all $n \in (0, 1)$.

Proof. Fix $n \in (0, 1]$, and define

$$\Delta(\beta) \equiv \frac{\beta - \rho(n)}{1 - \rho(n)} - \frac{\beta - \nu(n)}{n - \mu(n)}.$$

Note that $\Delta(\beta)$ is affine in β. Additionally, $\Delta(1) = 0$ while $\Delta(0) = (1 - \gamma(n))(1 - n)/((n - \mu(n))(1 - \rho(n)) > 0$. Hence $\Delta(\beta) > 0$ for all $\beta \in (0, 1)$. $\square$

We now show that the scheme satisfying $\sigma_B^* = \overline{\sigma}_B$ underpromotes non-risk-takers, overpromotes failed risk-takers, and pays no bonuses. Since $\overline{\sigma}_B > \underline{\sigma}_B$, $\sigma_0(\sigma_B)$ is decreasing in σ_B, and $\underline{\sigma}_B$ induces efficient promotion, the results about promotion follow immediately. The zero bonus result follows from the following lemma, which ensures that $\overline{\sigma}_B = (\beta - \rho(N))/(1 - \rho(N))$ for $\beta \geq \underline{\beta}$. As observed above, when $\beta < \underline{\beta}$ there are no risk-taking rates greater than $N^{nc}(\beta)$. So it is without loss to restrict attention to $\beta \geq \underline{\beta}$.

Lemma I.3. $(\beta - \rho(n))/(1 - \rho(n)) \leq (\beta - \mu(n))/(n - \mu(n))$ for all $n \in (0, 1]$ and $\beta \geq \underline{\beta}$.

Proof. Fixing $n \in (0, 1]$, let

$$\Delta(\beta) \equiv \frac{\beta - \mu(n)}{n - \mu(n)} - \frac{\beta - \rho(n)}{1 - \rho(n)}.$$

Note that $\Delta(\beta)$ is affine in β. We first show that $\Delta(\underline{\beta}) \geq 0$. To see this, observe that $\underline{\beta} = \mu(1) \geq \mu(n)$, while ρ is a strictly decreasing function of n and so $\rho(n) \geq \rho(1) = \mu(1) = \underline{\beta}$. Meanwhile some algebra reveals that

$$\Delta(1) = \frac{1 - n}{n - \mu(n)} \geq 0.$$

It follows that $\Delta(\beta) \geq 0$ for all $\beta \in [\underline{\beta}, 1)$, yielding the desired result. $\square$

The extremal case $N = 1$ follows by taking the limit of the optimal scheme for $N < 1$ and invoking the maximum theorem, in a manner analogous to the treatment of the $N = 0$ case in the proof of Proposition 3.

J Proof of Lemma 2

Note that Λ is nondecreasing so long as $(N - \mu(N))/(1 - \gamma(N))$ is nondecreasing, or equivalently if $\log(N - \mu(N))$ is concave. Write

$$\log(N - \mu(N)) = \log N + \log(1 - \bar{\gamma}(N)).$$

The first term on the rhs is immediately concave, while when $\bar{\gamma}$ is convex the second is a composition of two concave functions, the outer of which is concave. Hence the composition is also concave. $\log(N - \mu(N))$ is therefore a sum of two concave functions and so concave. If $\gamma(N) = A - BN^k$, then

$$\bar{\gamma}(N) = A - \frac{B}{k+1} N^k.$$

So long as $k \in (0, 1]$, this expression is convex.

K Proof of Proposition 6

Let $\Pi^{Pr}(N)$ be the organization's profits under an optimal promotion-reallocation scheme implementing target rate N, with $\Pi^B(N, V)$ defined similarly for an optimal bonus scheme. (Π^{Pr} is independent of V, while Π^B in general depends on V, and our notation reflects this fact.) These profit functions can be decomposed as

$$\Pi^{Pr}(N) = \Pi^{fb}(N) - \Delta^{Pr}(N), \quad \Pi^B(N, V) = \Pi^{fb}(N) - \Delta^B(N)V,$$

where $\Pi^{fb}(N)$ is the organization's profit under risk-taking rate N and the natural promotion policy (as defined in Appendix C) and Δ^{Pr} and Δ^B are incentive costs which are continuous, non-negative for all N, and strictly positive whenever $N \neq N^{nc}$.

Define

$$\Pi^{*,Pr} \equiv \max_N \Pi^{Pr}(N), \quad \Pi^{*,B}(V) \equiv \max_N \Pi^B(N, V), \quad \overline{\Pi} \equiv \max_N \Pi^{fb}(N)$$

Note that $\Pi^{Pr}(N) < \Pi^{fb}(N) \leq \overline{\Pi}$ for all N. Meanwhile the hypothesis $\beta \neq \beta^{fb}$ implies that N^{nc} is not a maximizer of Π^{fb} and therefore $\Pi^{Pr}(N^{nc}) = \Pi^{fb}(N^{nc}) < \overline{\Pi}$. Then since $\Pi^{Pr}(N)$ is continuous in N over the compact domain $[0, 1]$, it must be that $\Pi^{*,Pr} < \overline{\Pi}$.

Meanwhile, $\Pi^B(N,0) = \Pi^{fb}(N)$ for all N, so that $\Pi^{*,B}(0) = \overline{\Pi}$. Further, the maximum theorem implies that $\Pi^{*,B}(V)$ is continuous in V, so for V sufficiently close to 0 we must have $\Pi^{*,B}(V) > \Pi^{*,Pr}$. For such values of V, an optimal bonus scheme outperforms an optimal promotion scheme, and so is a globally optimal incentive scheme. Since $\Pi^{*,Pr} \geq \Pi^{fb}(N^{nc})$, this scheme must further satisfy $\Pi^{*,B}(V) > \Pi^{fb}(N^{nc})$. Hence there exists a nontrivial optimal incentive scheme.

L Proof of Proposition 7

Let $\Pi^{Pr}(N,R)$ be the organization's profits under an optimal promotion-reallocation scheme implementing target rate N, with $\Pi^B(N,R)$ defined similarly for an optimal bonus scheme. (Both functions depend on R in general, and our notation reflects this dependence.) These profit functions can be decomposed as

$$\Pi^{Pr}(N,R) = \Pi^{fb}(N,R) - \Delta^{Pr}(N)R, \quad \Pi^B(N,R) = \Pi^{fb}(N,R) - \Delta^B(N),$$

where $\Pi^{fb}(N,R)$ is the organization's profit under risk-taking rate N and the natural promotion policy (as defined in Appendix C) and Δ^{Pr} and Δ^B are incentive costs which are continuous, non-negative for all N, and strictly positive whenever $N \neq N^{nc}$.

Define

$$\Pi^{*,Pr}(R) \equiv \max_N \Pi^{Pr}(N,R), \quad \Pi^{*,B}(R) \equiv \max_N \Pi^B(N,R), \quad \overline{\Pi}(R) \equiv \max_N \Pi^{fb}(N,R)$$

Note that $\Pi^{fb}(N,0) = f([0,N])$, which is uniquely maximized by the risk-taking rate N^0 which satisfies $\gamma(N^0) = K$. Recall that β^0 is characterized by $\gamma(N^{nc}(\beta^0)) = K$. By hypothesis $\beta \neq \beta^0$, and therefore $\Pi^{fb}(N^{nc},0) < \overline{\Pi}(0)$.

Since $\Pi^{Pr}(N,0) = \Pi^{fb}(N,0)$ for all N, we must have $\Pi^{*,Pr}(0) = \overline{\Pi}(0)$. Meanwhile, since $\Pi^B(N,0) < \Pi^{fb}(N,0) \leq \overline{\Pi}(0)$ for all $N \neq N^{nc}$, while $\Pi^B(N^{nc},0) = \Pi^{fb}(N^{nc},0) < \overline{\Pi}(0)$, continuity of Π^B in N over the compact set $[0,1]$ implies that $\Pi^{*,B}(0) < \overline{\Pi}(0)$. By the maximum theorem, $\Pi^{*,Pr}, \Pi^{*,B}$, and $\overline{\Pi}$ are each continuous in R. It follows that $\overline{\Pi}(R) - \Pi^{*,B}(R) > \overline{\Pi}(R) - \Pi^{*,Pr}(R)$ for sufficiently small R. Equivalently, $\Pi^{*,Pr}(R) > \Pi^{*,B}(R)$ for sufficiently small R. Thus an optimal promotion scheme outperforms an optimal bonus scheme for small R, and so it must be a globally optimal incentive scheme. Further, since $\Pi^{*,B}(R) \geq \Pi^{fb}(N^{nc},R)$, it must also be that $\Pi^{*,Pr}(R) > \Pi^{fb}(N^{nc},R)$. Hence there exists a nontrivial optimal incentive scheme.

M Proof of Proposition 8

Given some asymmetric scheme $(G, \{(k^g, \beta^g, \mathcal{S}^g)\}_{g \in G})$, suppose there exists a $g \in G$ such that $\mathcal{S}^g$ reallocates promotions in a way that differs from the natural policy. We show that the organization, apart from a specific case addressed at the end of the proof, can do strictly better by splitting the employees in group g into two new groups x and y and allocating promotions appropriately between them. In particular, we specify sizes and promotion rates (k^x, β^x) and (k^y, β^y) such that the combined profit of the two new groups x and y exceeds the profit from the unsplit group g:

$$k^x \Pi^*(\beta^x) + k^y \Pi^*(\beta^y) > k^g \Pi^*(\beta^g).$$

At the end of the proof, we show that a specific case requires a mild condition for this result, but then show that if the condition is violated, the organization can still do weakly better by splitting group g into two groups that implement the natural promotion policy.

We begin by making two observations which simplify the proof. First, observe that given (β^g, k^g), specifying particular values for β^x and β^y, together with feasibility conditions $k^x + k^y = k^g$ and $k^x\beta^x + k^y\beta^y = k^g\beta^g$, fully pins down k^x and k^y. Second, defining $\Pi^{*,Pr}(\beta)$ and $\Pi^{*,B}(\beta)$ analogously to $\Pi^{*,Pr}(R)$ and $\Pi^{*,B}(R)$ in the proof of Propositions 7 and 6, we have $\Pi^*(\beta) = \max\{\Pi^{*,Pr}(\beta), \Pi^{*,B}(\beta)\}$. So for any β such that an optimal symmetric scheme entails promotion reallocation, we must have $\Pi^{*,Pr}(\beta) \geq \Pi^{*,B}(\beta)$, so that $\Pi^*(\beta) = \Pi^{*,Pr}(\beta)$. Therefore, it suffices to find β^x and β^y such that for the implied k^x and k^y,

$$k^x \Pi^{Pr}(\beta^x) + k^y \Pi^{Pr}(\beta^y) > k^g \Pi^{Pr}(\beta^g), \tag{M.1}$$

because then

$$k^x \Pi^*(\beta^x) + k^y \Pi^*(\beta^y) \geq k^x \Pi^{Pr}(\beta^x) + k^y \Pi^{Pr}(\beta^y) > k^g \Pi^{Pr}(\beta^g) = k^g \Pi^*(\beta^g).$$

Before showing how to characterize whether (M.1) holds, we note that by the promotion reallocation schemes identified in Propositions 3, 4, and 5, the function $\Pi^{*,Pr}(\beta)$ is the maximized value of

$$\Pi^{Pr}(N, \beta) \equiv \begin{cases} \Pi^{Pr}_{-}(N, \beta) & \text{if } N \leq N^{nc}(\beta) \\ \Pi^{Pr}_{+}(N, \beta) & \text{if } N \geq N^{nc}(\beta), \end{cases}$$

where

$$\Pi^{Pr}_{-}(N, \beta) \equiv f(N) + R\beta \left(\pi_0 + \frac{\mu(N)}{\rho(N)}(\pi_G - \pi_0) \right),$$

$$\Pi^{Pr}_{+}(N, \beta) \equiv f(N) + R \left(\beta\pi_0 + (1-\beta)\frac{\mu(N)}{1-\rho(N)}(\pi_G - \pi_0) \right).$$

We now give sufficient conditions for (M.1) in the following lemma.

Lemma M.1. Given (k^g, β^g), let $N^*(\beta^g)$ denote the argmax of $\Pi^{Pr}(N, \beta^g)$. If there exist $\beta^x < \beta^g$, $\beta^y > \beta^g$, and $N^g \in N^*(\beta^g)$ satisfying

(i) $N^{nc}(\beta^g)$, $N^{nc}(\beta^x)$, and $N^{nc}(\beta^y)$ are on the same side of N^g

(ii) $N^g \notin N^*(\beta^x) \cap N^*(\beta^y)$

then β^x and β^y, along with their implied k^x and k^y, satisfy (M.1). If (i) holds but (ii) is violated, then (M.1) holds with equality.

Proof. Let N^x and N^y be arbitrary elements of $N^*(\beta^x)$ and $N^*(\beta^y)$, respectively. Then $\Pi^{Pr}(N^x, \beta^x) \geq \Pi^{Pr}(N^g, \beta^x)$ and $\Pi^{Pr}(N^y, \beta^y) \geq \Pi^{Pr}(N^g, \beta^y)$, and condition (ii) implies at least one inequality is strict. Without loss of generality, suppose $N^g \leq N^{nc}(\beta^g)$ so that $\Pi^{Pr}(N^g, \beta^g) = \Pi^{Pr}_-(N^g, \beta^g)$. We establish (M.1) as follows:

$$\begin{aligned}
k^x \Pi^{Pr}(\beta^x) + k^y \Pi^{Pr}(\beta^y) &= k^x \Pi^{Pr}(N^x, \beta^x) + k^y \Pi^{Pr}(N^y, \beta^y) \\
&> k^x \Pi^{Pr}(N^g, \beta^x) + k^y \Pi^{Pr}(N^g, \beta^y) \\
&= k^x \Pi^{Pr}_-(N^g, \beta^x) + k^y \Pi^{Pr}_-(N^g, \beta^y) \\
&= k^g \left(\frac{k^x}{k^g} \Pi^{Pr}_-(N^g, \beta^x) + \frac{k^y}{k^g} \Pi^{Pr}_-(N^g, \beta^y) \right) \\
&= k^g \Pi^{Pr}_- \left(N^g, \frac{k^x}{k^g}\beta^x + \frac{k^y}{k^g}\beta^y \right) \\
&= k^g \Pi^{Pr}_-(N^g, \beta^g) = k^g \Pi^{Pr}(\beta^g),
\end{aligned} \tag{M.2}$$

where the second equality follows from condition (i) and the fourth equality follows from the linearity of $\Pi^{Pr}_-(N, \beta)$ in β. On the other hand, if $N^g \geq N^{nc}(\beta^g)$, replace $\Pi^{Pr}_-(\cdot, \cdot)$ with $\Pi^{Pr}_+(\cdot, \cdot)$ and the proof is identical. Finally, if condition (ii) is violated, then the inequality in (M.2) changes to an equality. □

If β^g is such that the optimal symmetric scheme entails a promotion scheme that differs from the natural one, then there exists $N^g \in N^*(\beta^g)$ such that $N^g \neq N^{nc}(\beta^g)$. We consider the cases $N^g < N^{nc}(\beta^g)$ and $N^g > N^{nc}(\beta^g)$.

Case 1: $N^g < N^{nc}(\beta^g)$. Then because $N^{nc}(\cdot)$ is continuous and decreasing in β, setting $\beta^x = 0$ satisfies condition (i), and there exists $\beta^y > \beta^g$ satisfying condition (ii). Furthermore, $N^*(\beta^x) = N^*(0) = \{N^0\}$, where N^0 maximizes $f(\cdot)$. So we have

$$0 = f'(N^0) < f'(N^0) + R\beta^g(\pi_G - \pi_0) \frac{\partial}{\partial N} \frac{\mu(N)}{\rho(N)}\Big|_{N=N^0} = \frac{\partial}{\partial N} \Pi^{Pr}_-(N^0, \beta^g).$$

which implies $N^0 \notin N^*(\beta^g)$, and therefore $N^g \notin \{N^0\} = N^*(0) = N^*(\beta^x)$, establishing condition (ii).

Case 2: $N^g > N^{nc}(\beta^g)$. Then because $N^{nc}(\cdot)$ is continuous and decreasing in β, setting $\beta^y = 1$ satisfies condition (i) and there exists $\beta^x < \beta^g$ satisfying condition (i). Observe that $N^*(\beta^y) = N^*(1) = N^0$. If $N^g \neq N^0$, then clearly $N^g \notin N^*(\beta^y)$, satisfying condition (ii). So suppose that $N^g = N^0$. Then $N^{nc}(\beta^g) < N^g = N^0$, so $\beta^0 < \beta^g$, and we can set $\beta^x = \beta^0$. From here, we consider two subcases.

Subcase 2.1: $N^g = N^0$, $N^{nc}(\beta^0) \notin N^*(\beta^0)$. Then $N^g = N^0 = N^{nc}(\beta^0) \notin N^*(\beta^0) = N^*(\beta^x)$, satisfying condition (ii). We have therefore shown that any asymmetric scheme assigning a scheme to some group which reallocates promotions can be strictly improved upon by splitting that group into two smaller groups. As a result, all optimal asymmetric schemes must promote each group according to the natural policy.

Subcase 2.2: $N^g = N^0$, $N^{nc}(\beta^0) \in N^*(\beta^0)$. Then $N^{nc}(\beta^x) \in N^*(\beta^x)$, so it is weakly optimal to promote employees in group x according to the natural policy. Furthermore, since $\beta^y = 1$, every employee in group y is promoted, which is the natural promotion scheme for that measure of promotions. Finally, observe that $N^g = N^0 = N^{nc}(\beta^0) \in N^*(\beta^0) = N^*(\beta^x)$ and $N^g = N^0 = N^*(1) = N^*(\beta^y)$, violating condition (ii), which by the lemma indicates that (M.1) holds with equality. This establishes that for any asymmetric scheme that assigns a scheme to some group which reallocates promotions, the organization can weakly improve by splitting that group into two separate groups that are promoted naturally. As a result, there exists an optimal asymmetric scheme that promotes each group according to the natural policy.