Designing Career Concerns

Erik Madsen, Basil Williams, and Andrzej Skrzypacz

October 1, 2020

Preliminary and Incomplete

Abstract

In many organizations, employees enjoy significant discretion regarding task selection. If tasks differ in their informativeness about an employee's quality, task choices will be distorted whenever career concerns are important. We analyze a model in which an organization can shape its employees' career concerns by committing to a system for allocating a limited set of promotions. We show that the organization optimally overpromotes certain categories of underperforming employees, trading off efficient matching of employees to promotions in return for superior task selection. When organizations can additionally pay monetary bonuses, we find that overpromotion is a superior incentive tool when the organization needs to offer high-powered incentives; otherwise, bonuses perform better.

1 Introduction

In this paper, we analyze how employees' career concerns distort task selection in organizations, and how organizations can design reward systems to mitigate these distortions. Our starting point is the observation that in many organizations, employees exercise significant discretion when deciding what projects to take on or how to complete projects they have been assigned. For instance, academics and many scientists have broad freedom to set their own research agendas; engineers may be handed design challenges and given authority to pursue the solution they judge most promising; and managers may be asked to craft strategic proposals for their division, with alternatives including pitching a novel strategy of their own or supporting a consensus option.

The choices available to an employee, which for simplicity we will refer to as tasks, in general differ across two key dimensions: their expected returns to the organization, and the

amount of information their outcome reveals about the employee's quality. An employee's choice of task will therefore have important implications for her career whenever the organization uses her performance to evaluate her suitability for promotion. Employees who are highly motivated by such career concerns will make task decisions which are inefficient from the organization's perspective. Some may seek tasks which will generate large losses for the organization in case of failure, but that could allow them to distinguish themselves and propel their careers if successful. Others may elect to stay away from risky but profitable tasks in order to avoid rocking the boat and undermining their progress in the organization.

As a prelude to our main analysis, we show formally how career concerns can lead to inefficient task selection absent a well-designed incentive scheme. The nature of the inefficiency depends fundamentally on the amount of upward mobility in the organization. In organizations with significant upward mobility, which we represent by a large pool of promotions to be filled by high performers, the inefficiency takes the form of employees shying away from highly productive but informative (and hence risky to their careers) projects. In organizations with little upward mobility, i.e. a small pool of promotions, the opposite problem appears: employees take too many risks in the hopes of distinguishing themselves by an impressive outcome and shy away from productive but "routine" tasks.

A substantial literature in social psychology studying the phenomenon of "psychological safety" emphasizes the first problem—situations in which employees refrain from drawing attention to problems, offering feedback, or taking other risks for fear of negative consequences, including to their careers (Edmondson 1999; Edmondson and Lei 2014; Kahn 1990). Lack of psychological safety has been associated with serious negative outcomes in a variety of contexts, for instance unreported nursing errors leading to patient deaths in hospitals (Edmondson 1999, p.352); unwillingness to offer feedback and suggestions leading to dysfunction in teams at Google (Google 2014); and reluctance to ask for information leading to slow dissemination of financial, health, and agricultural knowledge in rural India (Chandrasekhar et al. 2019). The problem of employees shying away from routine tasks is has been less systematically documented, but is illustrated, for example, by employees' neglect of dull but essential projects at Sun Hydraulics (Hill and Suesse 2003); and by the reluctance of research scientists to perform and publish replication studies of other researchers' publications, leading to the well-known "replication crisis" in a number of fields.

Our first main result demonstrates how commitment to an incentive scheme implementing an ex-post inefficient promotion policy can improve the performance of an organization. We show that in organizations with significant upward mobility, it is optimal to commit to overpromote employees who tackle risky projects and fail. That is, the organization should

act as if it discounts the value of the information gained from failure.[1] Such a scheme creates "psychological safety" to take risks and trades off more profitable task selection for less efficient matching of employees to promotions. On the other hand, in organizations with little upward mobility, the organization should commit to a very different policy: overpromoting those that devote time to routine tasks and undepromoting those who succeed at risky tasks. Instead of inducing "psychological safety" to take risks, such a policy rewards conservatism to ensure that a sufficient number of routine but important tasks are undertaken.

In many organizations, promotions are the primary incentive tool employed in practice.[2] However, in some organizations incentives can be effectively provided by monetary rewards as well as promotion policies. Our second set of results characterize the optimal incentive scheme when the organization can commit to both promotion policies and bonuses. We find that when an organization incentivizes through bonuses, it should deploy them in a very similar way to an optimal promotion scheme—when employees take too few risks, bonuses are paid for failure on risky tasks, while when employees take too many risks, bonuses are paid to employees undertaking routine tasks.

We further find that the organization should deploy only one of the two tools, with the optimal tool depending on how far the organization wishes to push task selection away from its no-commitment distribution. Bonuses turn out to be the superior tool for providing "low-powered" incentives, inducing only a small number of employees to switch tasks. On the other hand overpromotion is more effective for providing "high-powered" incentives, under which a significant fraction of employees switch tasks. We further show that the optimal power of incentives depends on the organization's payoff from matching the best employees to promotions. When this payoff is large, incentives should be low-powered and bonuses should be used, while when the payoff is small, high-powered incentives with an overpromotion scheme are optimal.

[1]Commitment to such policies bears similarities to corporate mantras urging employees to "celebrate failure" or "fail fast" in innovative organizations. For instance, Facebook adopted the internal motto "Move fast and break things" in its early years, and Google X's CEO Astro Teller has discussed how his organization recognizes and celebrates failures (`https://youtu.be/3SsnY2BvzeA`).

[2]For instance, Baker et al. (1988) note that "Promotions are used as the primary incentive device in most organizations, including corporations, partnerships, and universities. The empirical importance of promotion-based incentives, combined with the virtual absence of pay-for-performance compensation policies, suggests that providing incentives through promotion opportunities must be less costly or more effective than providing incentives through transitory financial bonuses."

1.1 Related literature

Our paper studies how an organization should structure its promotion system to influence employees' task choices in the face of career concerns. While a number of papers have studied multitasking and career concerns, to the best of our knowledge ours is the first to treat career concerns as a design variable rather than an exogenous incentive. Two foundational related papers are the career concerns model of Holmström (1999), which demonstrates how an agent's concern for their future reputation shapes incentives for exerting effort in situations of moral hazard; and the multitasking model of Holmström and Milgrom (1991), which characterizes optimal pay-for-performance schemes when an agent may split effort among multiple tasks. The basic career concerns framework has been extended by Gibbons and Murphy (1992) to allow for pay-for-performance contracts, and by Dewatripont et al. (1999) to a multitask environment. Kaarbøe and Olsen (2006) combine all three features by allowing a principal to write pay-for-performance contracts in a multitask environment with career concerns. A key feature of these papers is that the agent's reputational concerns are taken to be exogenous,[3] while we allow the organization to directly control career concerns via its choice of a promotion system.

Closely related to the career concerns literature are a pair of recent papers by Kuvalekar and Lipnowski (2020) and Kostadinov and Kuvalekar (2018), studying task selection in a dynamic agency setting. Both papers highlight that career concerns may lead agents to choose inefficient, uninformative tasks at some stages of their career in order to slow learning about their quality. The focus of those papers is quite different from ours, as they model a single agent's task dynamics, while we study how task selection varies across heterogeneous agents in a static multi-agent setting. Moreover, while those papers analyze equilibrium outcomes absent commitment, a central focus of our paper is the optimal joint design of promotions and monetary rewards.

Our paper is also related to the tournaments literature—for instance Lazear and Rosen (1981), Green and Stokey (1983), and Nalebuff and Stiglitz (1983)—which compare individual pay-for-performance contracts with tournaments where prizes are awarded based on rank-order comparisons between agents. The prizes in a tournament are often interpreted as promotions, framing the exercise as a form of promotion policy design.[4] However, the cost to

[3]Some papers, for instance Dewatripont et al. (1999), consider how an organization should respond to an agent's career concerns by assigning tasks to different agents or aggregating the performance measures that the market observes. However, such interventions do not account for any efficiency impact of changing how much the market knows about the agent's quality. Further, they do not allow the organization to tailor incentives as effectively as direct design of returns to reputation.

[4]For instance, Lazear and Rosen (1981) argue, "On the day that a given individual is promoted from

the organization of awarding a prize is assumed to be independent of the employee receiving it. As a result, analyses of tournaments typically abstract from the problem of matching high-quality workers to promotions, which is a central problem faced by the organization in our model. One exception is Rosen (1986), which considers the problem of selecting the most talented agent from a population via a sequence of pairwise tournaments. However, that paper does not design tournaments to optimize effort, instead confining attention to schemes which induce an arbitrary stationary effort level. By contrast, in our model the organization simultaneously optimizes over employee task choice and allocation of employees to promotions.

2 The model

An organization wishes to guide innovation by its employees in a decentralized environment in which innovation serves dual roles, determining short-run profits and providing information about which employees are most suitable for promotion into roles of greater responsibility. The organization's life unfolds over two stages. In the first, employees choose and complete projects, which provide an immediate productive payoff to the organization. In the second, the organization selects a subset of employees to promote, yielding a further payoff from future production.

The organization oversees a continuum of employees of measure 1, as well as a large stock of potential projects. Projects are heterogeneous across two dimensions—their risk profile and their expected productivity. Some projects are *routine*, and if completed generate a guaranteed payoff to the organization. Other projects are *innovative*, and if attempted may either succeed and generate a payoff of 1, or fail and produce no payoff. The organization possesses a continuum of measure 1 of each project type . Routine projects are homogeneous, and each generates a payoff of $K \in (0,1)$ to the organization. Innovative projects, on the other hand, differ in their productivity, as characterized by their probability of success. Specifically, the nth innovative project succeeds with probability $\gamma(n) \in [0,1)$, where γ is assumed to be a C^2 function satisfying $\gamma' < 0$.

Project assignment is decentralized, and is determined by a combination of random matching and employee choice. Each employee is first randomly matched with an inno-

vice-president to president, his salary may triple. It is difficult to argue that his skills have tripled in that 1-day period... It is not a puzzle, however, when interpreted in the context of a prize. The president of a corporation is viewed as the winner of a contest in which he receives the higher prize". Similarly, Green and Stokey (1983) note "[I]n most organizations one of the most important rewards is promotion... [E]mployees at any one echelon are competing for a fixed, smaller number of positions at the next higher echelon."

vative project. Employees are ex ante homogeneous, and so without loss we will label each employee by the innovative project to which she is matched. After being matched, each employee chooses whether to undertake a routine project or the innovative project available to them. The organization's total project payoff is then the sum of the output of all projects undertaken. Specifically, if a set $\mathcal{N} \subset [0,1]$ of employees choose innovative projects, the organization's total expected project payoff is

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

We will assume that the organization's optimal project mix includes both routine and innovative projects:

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

Let $N^\dagger \in (0,1)$ be the unique solution to $\gamma(N) = K$. Then the organization's project payoff is maximized when employees in the set $[0, N^\dagger]$ innovate.

Each employee's choice of project also impacts how much information is generated about their suitability for promotion. Specifically, we assume that each employee generates an uncertain payoff from promotion, and that succeeding at an innovative project raises the organization's estimate of that payoff, while failing lowers it. We microfound this inference through a quality type which governs both the employee's probability of success on an innovative project and the payoff they generate by being promoted. Each employee n has quality $\theta(n) \in \{H, L\}$, where qualities are drawn independently and identically with $\Pr(\theta(n) = G) = \pi \in (\gamma(0), 1)$. A low-quality employee never succeeds at their innovative project, while a high-quality employee succeeds with probability $q(n) \equiv \gamma(n)/\pi$. (These probabilities are calibrated so that the expected productivity of the nth employee's innovative project is exactly $\gamma(n)$.) An employee's quality type does not impact the output of a routine project. Quality types are not directly observed by either the employee or the organization. As a result, the organization can learn about an employee's type only by observing their performance on an innovative project.

The organization possesses a continuum of promotions of measure $\beta \in (0,1)$ to which it can assign employees. An employee who is promoted receives a private payoff of $V > 0$ regardless of their type, and receives a private payoff normalized to zero otherwise. Meanwhile a promoted employee generates a payoff of $R > 0$ to the organization if they are high-quality, and a payoff of 0 otherwise. An employee who is not promoted also generates a payoff of 0 for the organization.

Finally, we assume the organization and all employees are risk-neutral over outcomes and money, and that the payoffs $f(\mathcal{N})$, V, and R are all measured in dollar terms.

2.1 Incentive schemes

As task assignment is decentralized, the organization cannot directly determine how much innovation takes place. However, the organization can guide employee decisions through the use of two incentive tools: bonuses and promotion systems. Specifically, we assume that the organization can commit to how much it pays each employee and how it prioritizes them for promotion as a function of their project choice and outcome. We will refer to a joint choice of bonuses and promotions as an *incentive scheme*.

We impose the following requirements on an incentive scheme:

- *Feasibility*: At most β employees may be promoted.
- *Limited liability*: Every employee must receive a non-negative bonus in every state of the world.
- *Anonymity*: All employees choosing the same project type and producing the same outcome must receive the same bonus and be promoted with the same probability.

The requirements of feasibility and limited liability are routine. The requirement of anonymity amounts to an assumption that the organization cannot observe the productivity of the innovative project undertaken by a given employee.

Given the requirements above, we formally define an incentive scheme as follows:

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

- $\mathcal{N} \subset [0,1]$ is the set of innovative projects the organization recommends be implemented.
- $\boldsymbol{T} = (T_G, T_\emptyset, T_B) \geq 0$ are the bonuses received by an employee who, respectively, succeeds at an innovative project, completes a routine project, or fails at an innovative project.
- $\boldsymbol{\sigma} = (\sigma_G, \sigma_\emptyset, \sigma_B) \in [0,1]^3$ are the probabilities of promotion for an employee who, respectively, succeeds at an innovative project, chooses a routine project, or fails at an innovative project.

We interpret an incentive scheme as promoting employees uniformly at random from each bin of all employees who chose a particular project type and obtained the same outcome, with a fraction σ_i of employees from bin $i \in \{G, \emptyset, B\}$ promoted. As a result, the number of employees promoted from each bin is deterministic. Our notion of an incentive scheme therefore restricts attention to schemes satisfying a form of aggregate non-randomness. Of course,

from the perspective of any individual employee, promotion may be random conditional on the outcome of their chosen project.

An incentive scheme is *feasible* if it promotes at most β employees, supposing employees follow the recommended innovation policy.[5] This requirement is summarized by the inequality

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

It is *incentive-compatible* if all employees find it optimal to follow the scheme's recommended innovation policy. That is,

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

The following lemma shows that attention may be restricted to incentive schemes which recommend that only the best projects be implemented, given an overall number of innovative projects recommended.

Lemma 1. Fix any feasible and IC incentive scheme satisfying $|\mathcal{N}| = N$, under which $|[0, N] \setminus \mathcal{N}| > 0$. Then there exists another feasible, IC incentive scheme implementing N projects yielding strictly higher profits.

Going forward, we will describe innovation policies via the amount of innovation they recommend, without explicitly specifying the set of innovative projects implemented. For $N \in [0, 1]$, we will also define the project payoff $f(N) \equiv f([0, N])$.

3 Equilibrium innovation rates

We first characterize outcomes when the organization cannot commit to an incentive scheme. Formally, our two-stage timeline becomes a two-stage game, with employees choosing projects in the first stage and the organization subsequently choosing bonuses and promotions in the second stage.

Absent commitment, the organization pays no bonuses and promotes employees efficiently given observed outcomes. Recall that successes on innovative projects are good news about

[5] 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 an anonymous incentive scheme, which can condition only on the measure of project 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.

an employee's quality, failures are bad news, and routine projects provide no information. The organization therefore first promotes all employees who have succeeded at innovative projects, followed by employees who chose routine projects, and finally resorts to promoting employees who failed at innovated projects, until all promotions are filled. (Because even a low-quality employee provides a payoff no lower than leaving a spot unfilled, all promotion slots are filled.)

We now establish that the equilibrium innovation rate is unique in the game without commitment, and is declining in β whenever it is interior. For $N \in (0, 1]$, define

$$p_G(N) \equiv \frac{1}{N} \int_0^N \gamma(n)\, dn$$

to be the fraction of successes among all employees choosing innovative projects, supposing that employees in the set $[0, N]$ innovate. Further define $p_G(0) = \gamma(0)$. Finally, let $\underline{\beta} \equiv p_G(1)$ and $\overline{\beta} \equiv p_G(0)$. Note that

$$0 < \underline{\beta} < \overline{\beta} < 1$$

given that γ is strictly decreasing and $\gamma(0) \in (0, 1)$.

Proposition 1. For each $\beta \in (0, 1)$, there exists a unique perfect Bayesian equilibrium. In this equilibrium employees in the set $[0, N^{eq}(\beta)]$ innovate, and:

- $N^{eq}(\beta) = 1$ for $\beta \leq \underline{\beta}$,
- $N^{eq}(\beta)$ is the unique solution in $(0, 1)$ to

  $$\beta = p_G(N)N + \gamma(N)(1 - N)$$

  for $\beta \in (\underline{\beta}, \overline{\beta})$,
- $N^{eq}(\beta) = 0$ for $\beta \geq \overline{\beta}$,
- $N^{eq}(\beta)$ is continuous everywhere and strictly decreasing on $[p_G(1), p_G(0)]$.

Recall that lower-index employees are those possessing the most productive innovative projects. The fact that the equilibrium innovation set takes the form $[0, N^{eq}(\beta)]$ for some $N^{eq}(\beta)$ therefore implies that only employees with the best projects choose to innovate. This result is straightforward, because all employees receive the same payoff from routine projects while employees with good innovative projects have the most incentive to gamble for success.

The drop in innovation with β arises because each employee's incentives to innovate weaken, holding fixed the innovation decisions of all other employees. Intuitively, the bar for promotion drops as more employees are promoted, and so a given employee gains less

upside from succeeding at innovation and faces a larger downside from failing to innovate. This force pushes fewer employees to innovate in equilibrium as β increases. Equilibrium uniqueness is not immediate, because the game is one of strategic complements—as other employees innovate more, the bar for promotion rises and increases incentives for a given employee to innovate. The proof of the proposition establishes that nonetheless, a unique equilibrium exists for every value of β.

When β lies outside the range $[\underline{\beta}, \overline{\beta}]$, the equilibrium innovation rate is unambiguously suboptimal from the organization's perspective. To see this, recall that the organization's project payoff is maximized at the interior innovation rate $N^\dagger \in (0,1)$, so that absent considerations of promotion efficiency the organization would choose an interior innovation rate. Now suppose that no employees innovate. Then increasing innovation slightly would both raise project payoffs and reveal information about the quality of some employees, allowing for a more efficient allocation of employees to promotions. So the optimal innovation rate is strictly positive for all β, and therefore equilibrium innovation is inefficiently low when $\beta > \overline{\beta}$. On the other hand, suppose that all employees innovate. In this case the organization sees $p_G(1)$ employees who succeeded on innovative projects, which is strictly more than the organization needs to promote when $\beta < \underline{\beta} = p_G(1)$. As a result, when $\beta < \underline{\beta}$ equilibrium innovation is inefficiently high, because the organization can achieve the same promotion payoff and a strictly higher project payoff if slightly fewer employees innovate.

In this paper we are interested in how incentive schemes can improve on equilibrium outcomes when the equilibrium is inefficient. In light of the discussion of the previous paragraph, we will focus on incentive design when $\beta < \underline{\beta}$ and $\beta > \overline{\beta}$, which correspond to cases when equilibrium innovation is unambiguously suboptimal.

4 Optimal Policy, Small Beta

We first study optimal incentive schemes in organizations that have relatively few promotions available for their employees. In particular, we consider the case $\beta < \underline{\beta}$. In Section 3 we showed that absent commitment, for this range of β all employees would choose innovative projects. From the organization's perspective, this innovation rate is always higher than would be optimal if it could directly control innovation.

We now determine how the organization optimally mitigates this inefficiency when it can commit to an incentive scheme. We characterize the optimal incentive scheme in two steps. We first ask, for a given desired level of innovation, what incentive scheme optimally induces that amount of innovation. With this result in hand, we then optimize over the amount of innovation.

Definition 2. An incentive scheme *promotes efficiently* if its promotion probabilities maximize the organization's expected promotion payoff given the scheme's recommended innovation policy. An incentive scheme *overpromotes* (*underpromotes*) a set of employees if more (fewer) employees from that set are promoted than would occur under the promotion payoff-maximizing policy.

Proposition 2. Suppose that $\beta < \underline{\beta}$ and the organization wishes to implement N innovative projects. Then there exists $\overline{N} \in (0, 1]$ such that:

1. If $N < \overline{N}$, the organization optimally pays no bonuses, overpromotes employees who undertake routine projects, and underpromotes employees who succeed at innovative projects.

2. If $N > \overline{N}$, the organization optimally pays a positive bonus to employees who undertake routine projects and promotes efficiently.

$\overline{N}$ is decreasing in R and increasing in V, and $\overline{N} = 1$ for sufficiently small R or sufficiently large V, while $\lim_{R\to\infty} \overline{N} = \lim_{V\to 0} \overline{N} = 0$.

Because employees are incentivized to undertake too much innovation absent intervention, the organization must use promotions and bonuses to make routine tasks more attractive. The proposition shows that whether promotions or bonuses are the better incentive tool depends how much innovation N the organization wishes to target.[6]

Implementing a given number of safe projects requires the organization to pay an incentive cost for every employee choosing one. This cost is in general comprised of a combination of 1) bonus payments and 2) inefficient promotions, due to promoting some employees choosing routine projects over others who succeeded at innovative ones. The balance of these two costs must make the Nth employee just indifferent between working on an innovative versus a routine project. The optimal mix of incentive tools is determined by how both the organization and the marginal employee view the tradeoff between these two forms of reward.

It turns out that the organization's tradeoff is independent of N. Note that the incremental cost of a bonus is exactly 1; meanwhile the incremental cost of promoting an employee who chose a routine project over one who succeeded at innovating is $R(1 - \pi)$, reflecting the decreased average quality of those employees. Thus regardless of N, the organization

[6] In the proof of Proposition 2, we further establish that the optimal incentive scheme is unique whenever $N \notin \{0, \overline{N}, 1\}$. In the edge cases $N \in \{0, 1\}$, the optimal scheme is uniquely determined "on-path", that is, for employees who choose the on-path project type. When $\overline{N}$ is interior and $N = \overline{N}$, bonuses and promotion distortion are equally efficient incentive tools, and a continuum of optimal schemes exist using a combination of the two tools.

faces the same marginal rate of substitution between these two tools. How the optimal mix changes with N is therefore driven entirely by changes in how the marginal employee trades off bonuses against promotions.

Imagine lowering the marginal employee's bonus payment by an amount ΔT, and raising her probability of being promoted by an amount $\Delta\sigma$ to maintain indifference. The employee's total payoff from choosing a routine project is

$$U_R = T_\emptyset - \Delta T + V(\sigma_\emptyset + \Delta\sigma),$$

which is independent of N. Meanwhile the employee's total payoff from choosing her innovative project is

$$U_I = V\gamma(N)\sigma_G.$$

It turns out that changing $\sigma_\emptyset$ also implies an offsetting decrease in σ_G, and therefore U_I, in order to maintain the same total number of promotions. To the extent that U_I shrinks as $\sigma_\emptyset$ rises, this force decreases the $\Delta\sigma$ needed to maintain $U_R = U_I$ when the bonus drops. The key to the proof is establishing that U_I is less sensitive to $\sigma_\emptyset$ when N is larger. One component of this result is the fact that $\gamma(N)$ is decreasing in N, so that for larger N, U_I is less sensitive to changes in σ_G. Further, it turns out that for large N, σ_G is itself less sensitive to $\sigma_\emptyset$. The result is that as N increases, saving a dollar on bonus payments becomes increasingly expensive in terms of the compensating rise in promotion of employees working on routine projects. Hence bonuses are optimally employed only for large N, while promotion distortion is optimally used only for small N.

The next proposition jointly characterizes the optimal number N^* of innovative projects along with the optimal incentive scheme.

Proposition 3. Suppose that $\beta < \underline{\beta}$. Then there exists a unique $R^* > 0$ such that:

1. if $R < R^*$, the organization distorts promotions and pays no bonuses;

2. if $R > R^*$, the organization promotes efficiently and pays bonuses to employees who undertake safe projects.

Further, the optimal number of innovative projects N^* is increasing in R, and there exists an $\overline{R} > 0$ such that N^* is strictly increasing in R whenever $R \leq \overline{R}$ and N^* is constant for $R > \overline{R}$.[7]

[7]In general N^* may be set-valued for some values of R. We define the correspondence $N^*(R)$ to be *increasing* in R if for every $R > R'$ and $n \in N^*(R), n' \in N^*(R')$, either $n, n' \in N^*(R) \cap N^*(R')$ or else $n > n'$. It is *strictly increasing* in R if $R > R'$ and $n \in N^*(R), n' \in N^*(R')$ implies $n > n'$.

The parameter R captures the importance to the organization of matching promotions with high-quality employees relative to other priorities such as efficient project allocation and monetary savings. To this end, innovation is the sole tool by which the organization can discover talent, and more innovation reveals more high-quality employees.[8] Thus naturally the optimal amount of innovation pursued by the organization increases with R until the organization incentivizes enough innovation that all promotions are filled by employees succeeding at innovative projects, at which point the amount of innovation becomes insensitive to R.

The switchover from promotions to bonuses occurs for two reasons. First, fixing the amount of innovation, bonuses become cheaper relative to promotion distortion as R rises. Thus promotions are optimal for small R while bonuses are optimal for large R, holding N fixed. This logic is reflected in the fact that $\overline{N}$ is decreasing in R in Proposition 2. Second, Proposition 2 established that bonuses are optimal for large N while promotions are optimal for small N. Since the optimal amount of innovation rises in R, this force further encourages a switchover from promotions to bonuses as R rises.

5 Optimal Policy, Large Beta

We now study optimal incentive schemes in organizations that have many promotions to fill. In particular, we consider the case $\beta > \overline{\beta}$. Absent commitment, for this range of β all employees would choose safe projects. From the organization's perspective, this innovation rate is always lower than would be optimal if it could directly control innovation.

As in the small-β case, we first characterize how the organization optimally implements a given amount of innovation, and then analyze the optimal amount of innovation.

Proposition 4. Suppose that $\beta > \overline{\beta}$ and the organization wishes to implement N innovative projects. Assume that $(1-\gamma(N))(1-N)$ is nonincreasing. Then there exists $\overline{N} \in [0,1)$ such that:

1. If $N < \overline{N}$, the organization optimally pays a positive bonus to employees who attempt innovative projects but don't succeed, and promotes efficiently.

2. If $N > \overline{N}$, the organization optimally pays no bonuses, overpromotes employees who

[8] A caveat to this logic is that more innovation shrinks the pool of "indeterminate-talent" employees who have pursued routine projects. This fact is irrelevant when $\beta < \underline{\beta}$, as the organization is never short of such employees. However, as we will see in Section 5, when the organization has many promotions to hand out, reducing the pool of indeterminate-talent employees will have important consequences.

attempt innovative projects but don't succeed, and underpromotes employees who choose routine projects.

$\overline{N}$ is increasing in R and decreasing in V, and $\overline{N} = 0$ for sufficiently small R or sufficiently large V, while $\lim_{R\to\infty} \overline{N} = \lim_{V\to 0} \overline{N} = 1$.

When β is large, the organization faces the opposite problem compared to when β was small—now employees are not incentivized to innovate absent intervention, and so promotions and bonuses must now be used to reward employees to innovate. Similar to Proposition 2, this proposition finds that whether the optimal incentive scheme uses bonuses or promotion distortion depends on whether N falls below or above a particular threshold.[9] However, bonuses are now optimally used for small N while promotion distortion is optimally deployed for large N, the reverse of the outcome for small β.

As in the small-β case, the optimal mix of promotions and bonuses is determined by how the organization and marginal employee trade off the two tools. Unlike the small-β case, now the organization increasingly prefers promotion distortion over bonuses as N increases. The incremental cost of a bonus is fixed at 1 for all N; meanwhile the incremental cost of promoting an employee who failed to innovate over one chose a routine project is $R(\pi - \pi_B(N))$, where $\pi_B(N)$ is the organization's posterior inference about employees who failed. This posterior inference is increasing in N, because as more employees innovate the average innovative project is less likely to succeed, even if the employee is high-quality. Hence promotion distortion becomes less costly as N increases, leading the organization to favor this tool all else equal.

Of course, this preference could be overridden if the marginal employee requires increasing boosts to promotion to compensate for decreased bonuses as N increases. The proof proceeds by showing that, under some conditions on model primitives, this does not happen. The logic is similar to that for the small-β case. Suppose that bonuses are paid only for failure and the employee is always promoted following a success. Imagine lowering the marginal employee's bonus payment by an amount ΔT, and raising her probability of being promoted following failure by an amount $\Delta\sigma$ to maintain indifference. The employee's total payoff from choosing her innovative project is

$$U_I = \gamma(N) + (1 - \gamma(N))(T_B - \Delta T + V(\sigma_B + \Delta\sigma)),$$

while the employee's total payoff from choosing a routine project is

$$U_R = V\sigma_\emptyset.$$

[9] Uniqueness holds under the same conditions discussed in footnote 6.

Increasing σ_B requires an offsetting decrease in $\sigma_\emptyset$, and therefore also U_R, in order to maintain the same total number of promotions. To the extent that U_R shrinks as σ_B rises, this force decreases the $\Delta\sigma$ needed to maintain $U_I = U_R$ when the bonus drops. It turns out that $\sigma_\emptyset$ is more responsive to increases in σ_B as N increases. This force favors trading off bonuses for promotions. However, U_I is *also* more responsive to changes in σ_B, since $1 - \gamma(N)$, the likelihood of the marginal innovative project failing, is increasing in N. It is thus not inevitable that a smaller $\Delta\sigma$ will balance incentives as N increases. The requirement that $(1-N)(1-\gamma(N))$ be decreasing ensures that $1-\gamma(N)$ does not increase too quickly. Under that sufficient condition, the employee requires smaller promotion boosts to make up for a reduction in bonuses as N rises.[10]

A variety of forms of γ satisfy the monotonicity requirement of Proposition 4. One class of functions satisfying this property are those for which $\gamma'(N) \geq -(1-\gamma(0))$ for all N. Recall that $\gamma(0) \leq \pi < 1$, so that this condition is satisfied for γ with a sufficiently shallow slope everywhere. In particular, if $\gamma(N) = A - BN$ and $B \leq 1 - A$, the condition is satisfied.

One new feature of the design problem when the organization seeks to incentivize innovative rather than routine projects is that bonuses may be paid conditional on the outcome of the project. Proposition 4 establishes that bonuses are optimally paid after failure, not success. The reason for this is that the average employee who innovates succeeds more often than the marginal innovating employee, meaning that success bonuses are paid more often to the average innovating employee than to the marginal one. Thus bonuses can be shifted away from success and toward failure in a way which leaves the marginal innovating employee's incentives unchanged while reducing the average bonus payment to innovators. It follows that bonuses are paid for failure, not success.

The next proposition jointly characterizes the optimal number N^* of innovative projects along with the optimal incentive scheme.

Proposition 5. Suppose that $\beta > \overline{\beta}$. Then there exist thresholds $0 < R_* \leq R^*$ such that:

1. if $R < R_*$, the organization distorts promotions and pays no bonuses;
2. if $R > R^*$, the organization promotes efficiently and pays bonuses to employees who undertake safe projects.

This proposition establishes that the form of the optimal incentive scheme for small or large R is as in the small-β case. The basic intuition is similar—given a fixed amount of

[10] Absent any conditions on the growth rate of $1 - \gamma(N)$, it is possible to construct examples in which the optimal promotion scheme's dependence on N does not exhibit a threshold structure. Nonetheless, we establish in the proof of Proposition 4 that for any γ, bonuses are optimal for sufficiently small N while promotion distortion is optimal for sufficiently large N.

innovation, bonuses become cheaper as an incentive tool relative to promotion distortion as R rises. This logic drives the result that $\overline{N}$ is increasing in R in Proposition 4. In particular, when R is small promotions are optimal for any target level of innovation, while when R is large bonuses are optimal for all but very extreme amounts of innovation. It follows that an optimal incentive scheme distorts promotions for small R and pays bonuses for large R.

Unlike the small-β case, however, a straightforward threshold dependence on R is no longer ensured. This is because in general increasing R has two countervailing effects on the optimal incentive structure. On the one hand, holding N fixed, the optimal mix of incentives tilts away from promotion distortion and toward bonuses as R increases. But on the other hand, if the amount of innovation increases as well, Proposition 4 indicates that promotions become more favorable. Thus in general whenever N^* is increasing in R, the net effect of a small increase in R on the optimal incentive scheme is ambiguous, and single-crossing is not ensured.

Additionally, when β is large N^* is in general not monotone in R. This is because increasing the amount of induced innovation no longer has a straightforward positive impact on the amount of information gained about employees. To see this, suppose that the organization were able to freely reallocate one more employee from a routine task to innovation. If the organization could track the reassigned employee and personalize her promotion rule, its achievable promotion payoff could only go up from this reassignment. However, promotion is anonymous, and so the practical effect of this reassignment is to shrink the pool of employees engaged on routine tasks available for promotion. If the organization must then promote more employees who failed to innovate to make up its quota, its achievable promotion payoff actually goes down. In particular, instead of being able to promote the reassigned employee even if she had failed, the organization must promote a *random* failed innovator, and the average failed innovator is lower-quality than a failed innovator on the marginal project. So past a certain point, encouraging further innovation actually has a detrimental impact on the highest promotion payoff the organization can achieve.

Because of the ambiguous impact of innovation on information production, the organization may prefer *less* innovation than would maximize its project payoffs. In that case optimal innovation goes down as promotion payoffs become more important. This effect is especially pronounced when the organization uses bonuses and is able to promote efficiently given the amount of innovation it targets. The behavior of N^* as R varies therefore depends heavily on model parameters. Note that over any interval on which N^* is decreasing in R, the complexities regarding single-crossing discussed earlier do not arise, and single-crossing of the optimal incentive scheme in R is assured locally. In particular, if N^* is globally decreasing (which can occur for some model specifications), then single-crossing holds.

6 Conclusion

In this paper we have analyzed how employees' career concerns in an organization may distort their choice of tasks, 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 tasks are overpromoted, while when employees naturally take too few risks, those who take on risky tasks 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 tasks. We therefore abstract from the role of management in directly assigning employees to tasks. 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 tasks, top-down assignment of employees to tasks 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 tasks. 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 task. 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 tasks 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 task. 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 tasks before being evaluated for promotion, in order to understand its implications for optimal design of incentives.

References

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.

Chandrasekhar, Arun G., Benjamin Golub, and He Yang (2019). "Signaling, Shame, and Silence in Social Learning". Unpublished.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999). "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.

Edmondson, Amy (1999). "Psychological safety and learning behavior in work teams". *Administrative Science Quarterly* 44.2, pp. 350–383.

Edmondson, Amy and Zhike Lei (2014). "Psychological safety: The history, renaissance, and future of an interpersonal construct". *Annual Review of Organizational Psychology and Organizational Behavior* 1.1, pp. 23–43.

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.

Google (2014). *Guide: Understand Team Effectiveness*. URL: `https://rework.withgoogle.com/guides/understanding-team-effectiveness/steps/introduction/`.

Green, Jerry R and Nancy L Stokey (1983). "A comparison of tournaments and contracts". *Journal of Political Economy* 91.3, pp. 349–364.

Hill, Linda A and Jennifer Suesse (2003). "Sun Hydraulics: Leading in Tough Times (A)". No. 403-469.

Holmström, Bengt (1999). "Managerial incentive problems: A dynamic perspective". *The Review of Economic studies* 66.1, pp. 169–182.

Holmström, Bengt and Paul Milgrom (1991). "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design". *Journal of Law, Economics, and Organization* 7, p. 24.

Kaarbøe, Oddvar M and Trond E Olsen (2006). "Career concerns, monetary incentives and job design". *Scandinavian Journal of Economics* 108.2, pp. 299–316.

Kahn, William A (1990). "Psychological conditions of personal engagement and disengagement at work". *Academy of Management Journal* 33.4, pp. 692–724.

Kostadinov, Rumen and Aditya Kuvalekar (2018). "Learning in Relational Contracts". *Available at SSRN 2830928*.

Kuvalekar, Aditya and Elliot Lipnowski (2020). "Job Insecurity". *American Economic Journal: Microeconomics* 12.2, pp. 188–229.

Lazear, Edward P and Sherwin Rosen (1981). "Rank-order tournaments as optimum labor contracts". *Journal of Political Economy* 89.5, pp. 841–864.

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.

Appendix

A Proof of Lemma 1

Throughout this proof we use the following notation: for an arbitrary innovation set $\mathcal{N}$ satisfying $|\mathcal{N}| > 0$, we let

$$p_G(\mathcal{N}) \equiv \frac{1}{|\mathcal{N}|} \int_{\mathcal{N}} \gamma(n)\, dn.$$

For any $N > 0$, we let $p_G(N) \equiv p_G([0, N])$.

Fix a scheme satisfying the stated hypotheses. Trivially any such scheme involves $|\mathcal{N}| \in (0, 1)$. Further, such a scheme must involve $V\sigma_B + T_B \geq V\sigma_G + T_G$; for otherwise the payoff to innovating is strictly declining in n, in which case $\mathcal{N} = [0, N]$ is the unique incentive-compatible task allocation. We construct profit-improving incentive schemes separately for the cases $V\sigma_B + T_B = V\sigma_G + T_G$ and $V\sigma_B + T_B > V\sigma_G + T_G$. In the first case we show that the task allocation can be changed in a profit-improving way, while in the second case we show that either bonuses or promotion probabilities can be profitably changed.

Consider first a scheme in which $V\sigma_B + T_B = V\sigma_G + T_G$. In this case all employees receive the same payoff from innovating. Then since $|\mathcal{N}| \in (0, 1)$, employees must also be indifferent between innovating and not, implying

$$V\sigma_G + T_G = V\sigma_\emptyset + T_\emptyset = V\sigma_G + T_B.$$

We will show that there exists another feasible, IC scheme in which $\mathcal{N} = [0, N]$ and profits are strictly higher.

Suppose first that $\sigma_G = \sigma_B$. Then $T_G = T_B$ as well, and the total number of innovating employees promoted, the mix of good and bad outcomes promoted, and total bonus payments to innovating employees, are all independent of $\mathcal{N}$. Thus modifying the original scheme by setting $\mathcal{N} = [0, N]$ preserves feasibility and does not impact bonus or promotion payoffs, but improves the task payoff, yielding strictly higher profits.

Next suppose that $\sigma_G > \sigma_B$. Let $\beta' \leq \beta$ be the total number of employees promoted in the original scheme. In this case changing $\mathcal{N}$ to $[0, N]$ while leaving promotion probabilities unchanged increases the ratio of good to bad outcomes and thus the total number of employees promoted. So consider a new scheme which chooses $\mathcal{N} = [0, N]$ and sets $\boldsymbol{\sigma}' = \boldsymbol{\sigma}/M$ and $\boldsymbol{T}' = \boldsymbol{T}/M$ for the unique $M > 1$ which yields β' total promotions. This new scheme scales down the payoff to innovative and safe projects by the same amount for all employees, preserving indifference and ensuring that $\mathcal{N} = [0, N]$ is incentive-compatible. The change in task allocation under the scheme raises the task payoff. Further, compared to the original

scheme, this new scheme increases the ratio of good to bad outcomes promoted, as well as the ratio of good outcomes to safe projects promoted. Then as the total number of employees promoted is held fixed, the total number of good outcomes promoted must rise while the number of bad outcomes and safe projects promoted both fall, and the total promotion payoff must rise. Finally, bonus payments shrink under the new scheme. To see this, note that $\sigma_G > \sigma_B$ implies $T_G < T_B$ given that $V\sigma_G + T_G = V\sigma_B + T_B$. Thus an increase in the ratio of good to bad outcomes decreases bonus payments to innovating employees. Further, we have shrunk bonus payments to all employees by a factor of M. Both of these forces decrease total bonus payments. So all three components of the profit function rise compared to the original scheme.

Finally, suppose that $\sigma_G < \sigma_B$. Let $\beta' \leq \beta$ be the total number of employees promoted under the original scheme. Now changing $\mathcal{N}$ to $[0, N]$ decreases the total number of employees promoted, reducing the total number of employees promoted. So consider a new scheme with sets $\mathcal{N}' = [0, N]$ and chooses $\sigma'_G = \sigma_G + \Delta/V$ and $T'_G = T_G - \Delta$, where $\Delta > 0$ is chosen so that β' employees are promoted. We first show that this new scheme satisfies the boundary constraints $\sigma'_G \leq 1$ and $T'_G \geq 0$. First observe that if $\sigma'_G \geq \sigma_B$, then the total number of employees promoted under the new scheme would be at least

$$N\sigma_B + (1-N)\sigma_\emptyset > Np_G(\mathcal{N})\sigma_G + N(1-p_G(\mathcal{N}))\sigma_B + (1-N)\sigma_\emptyset = \beta',$$

contradicting the assumption that Δ is chosen to promote exactly β' employees. So $\sigma'_G < \sigma_B \leq 1$, satisfying the boundary constraint on this parameter. Further,

$$V\sigma'_G + T'_G = V\sigma_G + T_G = V\sigma_B + T_B \geq V\sigma_B > V\sigma'_G,$$

meaning $T'_G > 0$. Thus the boundary constraint on T'_G is also satisfied.

We now show that the organization's payoff rises under the new scheme. Certainly the task payoff increases. Meanwhile since the number of safe outcomes promoted is unchanged, this new scheme must also leave the total number of innovating employees promoted unchanged versus the original scheme. Further, both the number of good outcomes and their probability of promotion has gone up, so more good outcomes in absolute terms must be promoted. Thus promotions are transferred from bad to good outcomes, increasing the total promotion payoff. It remains only to show that total bonus payments do not rise under the new scheme. Under the original scheme, total bonuses were

$$B = T_G p_G(\mathcal{N})N + T_B(1-p_G(\mathcal{N}))N + T_\emptyset(1-N),$$

while under the new scheme, bonuses are

$$B' = T'_G p_G(N)N + T_B(1-p_G(N))N + T_\emptyset(1-N).$$

Now, by construction the total number of innovating employees being promoted is the same under the original and the new scheme. Therefore

$$\sigma_G p_G(\mathcal{N}) + \sigma_B(1 - p_G(\mathcal{N})) = \sigma'_G p_G(N) + \sigma_B(1 - p_G(N)).$$

Rearranging this equality yields $\sigma'_G p_G(N) - \sigma_G p_G(\mathcal{N}) = \sigma_B(p_G(N) - p_G(\mathcal{N}))$. Now, by construction $T_G + V\sigma_G = T'_G + V\sigma'_G$. Using this identity to eliminate σ'_G from the previous expression yields

$$(T_G - T'_G)p_G(N) = V(\sigma_B - \sigma_G)(p_G(N) - p_G(\mathcal{N})).$$

Meanwhile the assumption that $V\sigma_G + T_G = V\sigma_B + T_B$ allows $\sigma_B - \sigma_G$ to be eliminated from the previous expression, yielding

$$(T_G - T'_G)p_G(N) = (T_G - T_B)(p_G(N) - p_G(\mathcal{N})),$$

or

$$T_B(p_G(N) - p_G(\mathcal{N})) = T'_G p_G(N) - T_G p_G(\mathcal{N}).$$

Using this identity, we may compute

$$(B' - B)/N = T'_G p_G(N) - T_G p_G(\mathcal{N}) - T_B(p_G(N) - p_G(\mathcal{N})) = 0.$$

So total bonuses are unchanged in the new scheme, implying that overall profits rise.

In the remainder of the proof we consider schemes in which $V\sigma_G + T_G < V\sigma_B + T_B$. In this case the payoff to an innovative project is strictly increasing in n, and so the unique incentive-compatible task allocation is $\mathcal{N} = [1 - N, 1]$, and the IC constraint on the Nth employee must be binding:

$$\gamma(1 - N)(V\sigma_G + T_G) + (1 - \gamma(1 - N))(V\sigma_B + T_B) = V\sigma_\emptyset + T_\emptyset.$$

Note that any alternative scheme which also satisfies $\mathcal{N} = [1-N, 1]$, $V\sigma_G + T_G \leq V\sigma_B + T_B$, and this binding IC constraint is also fully incentive-compatible.

We first show that unless $\min\{T_G, T_\emptyset\} = 0$ and $T_B = 0$, there exists a new scheme with modified bonuses which strictly increase overall profits. Suppose first that $\min\{T_G, T_\emptyset\} > 0$. Then there exists a $\Delta > 0$ sufficiently small such that the new bonus scheme $(T'_G, T'_\emptyset, T'_B) = (T_G - \Delta/\gamma(1 - N), T_\emptyset - \Delta, T_B)$ satisfies all boundary constraints. By construction, this new set of bonuses is fully incentive-compatible for every $\Delta > 0$. Further, this change strictly decreases total bonus payments. So a new incentive scheme with these bonuses is incentive-compatible and strictly increases total probability.

A similar argument yields profitable improvements if $\min\{T_B, T_\emptyset\} > 0$. Thus in particular if $T_B > 0$ and $T_\emptyset > 0$, there exists a profitable improvement. Suppose instead that $T_B > 0$ and $T_\emptyset = 0$. Consider an alternative bonus scheme which sets

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

This new bonus scheme preserves the binding IC constraint for the Nth employee for all Δ. Further, for $\Delta > 0$ sufficiently small, $T'_B(\Delta) > 0$ and additionally $V\sigma_G + T'_G(\Delta) \leq V\sigma_B + T'_B(\Delta)$, so that the scheme satisfies the boundary constraints and is fully IC. Total bonus payments under this new scheme are

$$B(\Delta) = T_\emptyset(1-N) + T'_G(\Delta) \int_{1-N}^{1} \gamma(n)\, dn + T'_B(\Delta) \left(N - \int_{1-N}^{1} \gamma(n)\, dn \right).$$

Differentiating wrt Δ yields

$$\frac{dB}{d\Delta} = \frac{\int_{1-N}^{1} \gamma(n)\, dn}{\gamma(1-N)} - N.$$

Since γ is strictly decreasing, $\int_{1-N}^{1} \gamma(n)\, dn < \gamma(1-N)N$, so that $B'(\Delta) < 0$. Thus bonus payments strictly decrease in Δ, and so for small $\Delta > 0$ this new scheme is feasible, IC and strictly increases profits.

We have so far shown a profitable improvement for any scheme satisfying $V\sigma_B + T_B > V\sigma_G + T_G$ and either $\min\{T_G, T_\emptyset\} > 0$ or $T_B > 0$. It remains only to find a profitable improvement in the remaining case that $\min\{T_G, T_\emptyset\} = 0$ and $T_B = 0$. The binding IC constraint for the Nth employee, combined with $V\sigma_B + T_B > V\sigma_G + T_G$ and $T_B = 0$, implies that

$$V\sigma_B > V\sigma_\emptyset + T_\emptyset > V\sigma_G + T_G.$$

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

Suppose first that the feasibility constraint is slack—that is, fewer than β employees are promoted. We know that at least one of $T_\emptyset$ and T_G is zero. Suppose first that both are zero. Since $\sigma_G, \sigma_\emptyset < 1$, both probabilities may be raised in concert to preserve the Nth employee's binding IC constraint. For a small enough rise in promotion probabilities, this modification also maintains feasibility and preserves global IC. Further, change strictly raises promotion payoffs since more employees are promoted, and task and bonus payments are unchanged, raising total profits. Suppose instead that exactly one of $T_\emptyset$ and T_G is zero. Then raise the promotion payoff of the outcome with no bonus, and lower the bonus on the other outcome by a corresponding amount to preserve the Nth employee's IC constraint. This can be done while respecting boundary constraints, feasibility, and global IC for a small

enough perturbation. This change strictly raises promotion payoffs and strictly lowers bonus payments, again raising total profits.

Suppose instead that the feasibility constraint is binding. Consider a new scheme which leaves $\mathcal{N}$ and bonuses unchanged, and sets promotion probabilities $(\sigma'_G, \sigma'_B, \sigma'_\emptyset)$ satisfying $\sigma'_G = (1-\alpha)\sigma_G + \alpha\overline{\sigma} + \Delta$, $\sigma'_B = (1-\alpha)\sigma_B + \alpha\overline{\sigma} + \Delta$, and $\sigma'_\emptyset = \sigma_\emptyset + \Delta$, where $\overline{\sigma} \equiv \gamma(1-N)\sigma_G + (1-\gamma(1-N))\sigma_B$ and $\alpha \in (0,1)$ and $\Delta > 0$ are constants to be determined momentarily. Since IC binds for the Nth employee in the original scheme, by construction it continues to bind in the new scheme. Further, since $V\sigma_B + T_B > V\sigma_G + T_G$, for α sufficiently close to zero the inequality $V\sigma'_B + T_B > V\sigma_G + T'_G$ holds. Thus this new scheme is incentive-compatible for all employees for α close to zero. Fix any such $\alpha > 0$ going forward.

Under the new incentive scheme, the total number of promoted employees is

$$\begin{aligned}&\left(\int_{1-N}^{1}\gamma(n)\,dn\right)\sigma'_G + \left(N - \int_{1-N}^{1}\gamma(n)\,dn\right)\sigma'_B + (1-N)\sigma'_\emptyset\\ &= (1-\alpha)\beta + \alpha\left(N\overline{\sigma} + (1-N)\sigma_\emptyset\right) + \Delta.\end{aligned}$$

Since $\gamma(1-N) > \frac{1}{N}\int_{1-N}^{1}\gamma(n)\,dn$ and $\sigma_G < \sigma_B$, we have

$$\overline{\sigma} = \gamma(1-N)\sigma_G + (1-\gamma(1-N))\sigma_B < \left(\frac{1}{N}\int_{1-N}^{1}\gamma(n)\,dn\right)\sigma_G + \left(1 - \frac{1}{N}\int_{1-N}^{1}\gamma(n)\,dn\right)\sigma_B,$$

and therefore

$$\begin{aligned}&\left(\int_{1-N}^{1}\gamma(n)\,dn\right)\sigma'_G + \left(N - \int_{1-N}^{1}\gamma(n)\,dn\right)\sigma'_B + (1-N)\sigma'_\emptyset\\ &< (1-\alpha)\beta + \alpha\left(\left(\int_{1-N}^{1}\gamma(n)\,dn\right)\sigma_G + \left(N - \int_{1-N}^{1}\gamma(n)\,dn\right)\sigma_B + (1-N)\sigma_\emptyset\right) + \Delta\\ &= \beta + \Delta.\end{aligned}$$

So the feasibility constraint is slack under the new scheme for $\Delta = 0$, and there exists a unique $\Delta^* > 0$ such that the feasibility constraint just binds. Set $\Delta = \Delta^*$. We must check that for this choice of Δ, the resulting scheme satisfies the boundary constraints $\sigma'_G, \sigma'_\emptyset, \sigma'_B \leq 1$.

Note that Δ^* satisfies

$$\beta = (1-\alpha)\beta + \alpha(N\overline{\sigma} + (1-N)\sigma_\emptyset) + \Delta^*,$$

or $\Delta^* = \alpha(\beta - N\overline{\sigma} - (1-N)\sigma_\emptyset)$. So define $\Delta^0 \equiv \beta - N\overline{\sigma} - (1-N)\sigma_\emptyset$. Then

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

and similarly

$$\sigma'_\emptyset = (1-\alpha)\sigma_\emptyset + \alpha(\sigma_\emptyset + \Delta^0),$$

$$\sigma'_B = (1-\alpha)\sigma_B + \alpha(\overline{\sigma} + \Delta^0).$$

Then as $\sigma_G, \sigma_\emptyset < 1$ and $\overline{\sigma}, \Delta^0$ are independent of α, it must be that $\sigma'_G, \sigma'_\emptyset, \sigma'_B < 1$ for α sufficiently small.

We have shown that for $\alpha > 0$ sufficiently small, and an appropriate choice of Δ, the new scheme is feasible and incentive-compatible. The final step is to show that this new scheme raises profits compared to the original scheme. Note that the new scheme yields identical task and bonus payoffs, since the task allocation and bonuses were not changed. Further, $\sigma'_G > \sigma_G$ and $\sigma'_\emptyset > \sigma_\emptyset$ given that $\Delta > 0$ and $\overline{\sigma} > \min\{\sigma_G, \sigma_B\} = \sigma_G$. Therefore the new scheme promotes more employees with good outcomes and safe projects, while by construction preserving the same total number of promotions. Thus this new scheme strictly raises promotion payoffs and therefore total profits.

B Proof of Proposition 2

We derive the optimal incentive scheme when $V = 1$, and solve the case of general V at the end of the proof. Let

$$\overline{N}(R) \equiv \sup\{N \in [0,1] : p_G(N)N(R(1-\pi) - 1) < \gamma(N)(1-N)\}.$$

We establish that the following is an optimal incentive scheme:

1. If $N \leq \overline{N}(R)$, then the organization distorts promotions but pays no bonuses:

$$\sigma_G = \frac{\beta}{\mu(N)}, \quad \sigma_\emptyset = \frac{\beta\gamma(N)}{\mu(N)}, \quad \sigma_B = 0, \quad T_G = T_B = T_\emptyset = 0,$$

   where $\mu(N) \equiv p_G(N)N + \gamma(N)(1-N)$.

2. If $N > \overline{N}(R)$, then the organization pays bonuses but promotes efficiently. In particular, letting $N^0 \in (0,1)$ be the solution to $p_G(N)N = \beta$,

   (a) If $N \leq N^0$,

$$\sigma_G = 1, \quad \sigma_\emptyset = \frac{\beta - p_G(N)N}{1-N}, \quad \sigma_B = 0, \quad T_\emptyset = \frac{\mu(N) - \beta}{1-N}, \quad T_G = T_B = 0.$$

   (b) If $N > N^0$,

$$\sigma_G = \frac{\beta}{p_G(N)N}, \quad \sigma_\emptyset = 0, \quad \sigma_B = 0, \quad T_\emptyset = \frac{\beta\gamma(N)}{p_G(N)N}, \quad T_G = T_B = 0.$$

We further show that when $N \notin \{0, \overline{N}(R), 1\}$, the optimal incentive scheme is unique.

We first characterize an optimal policy for $N \in (0,1)$, and return to the extremal case afterward. We begin by conjecturing that at the optimum, among all IC constraints only the Nth employee's IC constraint binds, in the direction of choosing a safe project. Thus we solve the relaxed problem in which the only IC constraint is

$$\gamma(N)(\sigma_G + T_G) + (1 - \gamma(N))(\sigma_B + T_B) \leq \sigma_\emptyset + T_\emptyset,$$

and confirm that the resulting optimal scheme involves $\sigma_G + T_G \geq \sigma_B + T_B$ and a binding Nth-employee IC constraint. Thus the solution to the relaxed problem satisfies the full set of IC constraints.

We first argue that both constraints bind at the optimum of the relaxed problem. Suppose first that the feasibility constraint were slack. Then trivially profits are maximized by setting $T_G = T_B = T_\emptyset = 0$ and $\sigma_G = \sigma_\emptyset = \sigma_B = 1$, since this incentive scheme satisfies the Nth-employee IC constraint. But then a measure 1 of employees are promoted, violating feasibility. So the feasibility constraint binds at the optimum. Suppose next that the Nth-employee IC constraint were slack. Then the optimal scheme would pay no bonuses and promote efficiently subject to feasibility, setting $\sigma_G = 1$, $\sigma_\emptyset = \max\{(\beta - p_G(N)N)/(1-N), 0\}$, and $\sigma_B = 0$. But then

$$\gamma(N)(\sigma_G + T_G) + (1 - \gamma(N))(\sigma_B + T_B) = \gamma(N).$$

If $p_G(N)N \geq \beta$ then $\sigma_\emptyset + T_\emptyset = 0$ while $\gamma(N) > 0$ given $N < 1$, violating the IC constraint. On the other hand if $p_G(N)N < \beta$, then $\sigma_\emptyset = (\beta - p_G(N)N)/(1-N)$ and

$$\gamma(N) - (\sigma_\emptyset + T_\emptyset) = \frac{p_G(N)N + \gamma(N)(1-N) - \beta}{1-N} = \frac{\mu(N) - \beta}{1-N},$$

where $\mu(N) \equiv p_G(N)N + \gamma(N)(1-N)$. Note that $\mu'(N) = \gamma'(N)(1-N) < 0$, so μ is uniquely minimized at $N = 1$, where $\mu(1) = p_G(1) > \beta$. Thus $\gamma(N) > \sigma_\emptyset + T_\emptyset$ for all N, again violating the IC constraint. So the IC constraint must bind. (This establishes one of the claims of the previous paragraph about the optimal scheme in the relaxed problem.)

We next argue that $T_G = T_B = 0$ at the optimum. First consider any feasible, IC incentive scheme in which $\max\{T_G, T_B\} > 0$. This scheme may be modified by setting $T_G = T_B = 0$ without violating the IC constraint, and since $N > 0$ this modification strictly increases profits. So any optimal scheme must satisfy $T_G = T_B = 0$.

We now argue that $\sigma_B = 0$ at the optimum. (This combined with $T_G = T_B = 0$ establishes the earlier claim that $\sigma_G + T_G \geq \sigma_B + T_B$ at the optimum of the relaxed problem.) Suppose first that $T_\emptyset > 0$ at the optimum. Then the binding IC constraint

$$\gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B = \sigma_\emptyset + T_\emptyset$$

implies that $\sigma_\emptyset < 1$, for otherwise the rhs would be strictly greater than 1, while the lhs is at most 1. Suppose that $\sigma_B > 0$ at the optimum. Consider a new promotion scheme $(\sigma'_G, \sigma'_\emptyset, \sigma'_B) = (\sigma_G, \sigma_\emptyset + \Delta, \sigma_B - \Delta')$, where Δ, Δ' are sufficiently small that $\sigma_\emptyset + \Delta < 1$ and $\sigma_B - \Delta' > 0$, and are chosen relative to one another to ensure that the feasibility constraint continues to be satisfied. This new scheme lowers the payoff to choosing an innovative project and raises the payoff to choosing a safe project. Therefore the IC constraint may be satisfied by appropriately lowering $T_\emptyset$, which is feasible given $T_\emptyset > 0$ if Δ, Δ' are chosen sufficiently small. This new incentive scheme is feasible and IC, pays strictly lower bonuses, and reallocates promotions from bad outcomes to employees choosing safe projects. Thus total profits must go up, contradicting the presumed optimality of the original scheme. Thus $\sigma_B = 0$ if $T_\emptyset > 0$ at the optimum.

Suppose instead that $T_\emptyset = 0$ at the optimum. Then the optimal promotion probabilities must solve the reduced optimization problem in which $T_\emptyset = 0$. With the bonus eliminated, the firm's problem is to maximize the value of promoted employees

$$N(p_G(N)\pi_G\sigma_G + (1 - p_G(N))\pi_B(N)\sigma_B) + (1 - N)\pi\sigma_\emptyset$$

subject to the constraints

$$\begin{cases} \gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B \leq \sigma_\emptyset, \\ \beta \geq N(p_G(N)\sigma_G + (1 - p_G(N))\sigma_B) + (1 - N)\sigma_\emptyset. \end{cases}$$

The arguments used previously for the full problem with bonuses continue to imply that both constraints must bind in the problem without bonuses. Use the binding IC constraint to eliminate $\sigma_\emptyset$ from the problem, leaving the objective

$$(Np_G(N)\pi_G + (1 - N)\gamma(N)\pi)\sigma_G + (N(1 - p_G(N))\pi_B(N) + (1 - N)(1 - \gamma(N))\pi)\sigma_B,$$

subject to the reduced feasibility constraint

$$\beta \geq (Np_G(N) + (1 - N)\gamma(N))\sigma_G + (N(1 - p_G(N)) + (1 - N)(1 - \gamma(N))\sigma_B,$$

which must bind at the optimum. The derivatives of the corresponding Lagrangian are

$$\frac{\partial \mathscr{L}}{\partial \sigma_G} = p_G(N)N\left(\pi_G - \lambda\left(1 + \frac{\gamma(N)(1 - N)}{p_G(N)N}\right)\right)$$

and

$$\frac{\partial \mathscr{L}}{\partial \sigma_B} = (1 - p_G(N))N\left(\pi_B(N) - \lambda\left(1 + \frac{(1 - \gamma(N))(1 - N)}{(1 - p_G(N))N}\right)\right),$$

where $\lambda > 0$ at the optimum given that the feasibility constraint must bind. Given that $N > 0$, we have $\gamma(N) < p_G(N)$ and thus $1 - \gamma(N) > 1 - p_G(N)$. It follows that

$$\pi_G - \lambda\left(1 + \frac{\gamma(N)(1-N)}{p_G(N)N}\right) > \pi_B(N) - \lambda\left(1 + \frac{(1-\gamma(N))(1-N)}{(1-p_G(N))N}\right).$$

Hence at the optimum, at most one of σ_G and σ_B may be interior, and $\sigma_G \geq \sigma_B$. In particular, if $\sigma_B > 0$ then $\sigma_G = 1$. But then the constraint reduces to

$$\beta = (Np_G(N) + (1-N)\gamma(N)) + (N(1-p_G(N)) + (1-N)(1-\gamma(N))\sigma_B.$$

Note that $Np_G(N) + (1-N)\gamma(N)$ has derivative $(1-N)\gamma'(N) < 0$, so that the expression is minimized at $N = 1$, where it equals $p_G(1)$. Thus in particular the rhs of the constraint is strictly greater than $p_G(1)$, which by assumption is strictly greater than β, a contradiction. So it must be that $\sigma_B = 0$ if $T_\emptyset = 0$ at the optimum.

Going forward, we restrict attention to the problem with both constraints enforced with equality and $\sigma_B = T_G = T_B = 0$:

$$\gamma(N)\sigma_G = \sigma_\emptyset + T_\emptyset. \tag{B.1}$$

$$\beta = N\sigma_G p_G(N) + (1-N)\sigma_\emptyset, \tag{B.2}$$

We may use these constraints to eliminate $\sigma_\emptyset$ and $T_\emptyset$ from the firm's problem, yielding a maximization problem wrt σ_G:

$$\max_{\sigma_G \in [0,1]} -\Big((1-N)\gamma(N)\sigma_G - (\beta - \sigma_G p_G(N)N)\Big) + R\Big(\beta\pi + N\sigma_G p_G(N)(1-\pi)\Big), \tag{B.3}$$

subject only to the boundary constraints that $\sigma_G, \sigma_\emptyset \in [0,1]$ and $T_\emptyset \geq 0$, with $T_\emptyset$ and $\sigma_\emptyset$ characterized in terms of σ_G by equations (B.1) and (B.2).

The boundary constraints on $\sigma_G, \sigma_\emptyset$, and $T_\emptyset$ collectively imply that $\sigma_G \in [\underline{\sigma}_G, \overline{\sigma}_G]$, where

$$\overline{\sigma}_G \equiv \min\left\{\frac{\beta}{p_G(N)N}, 1\right\}$$
$$\underline{\sigma}_G \equiv \max\left\{0, \frac{\beta - (1-N)}{p_G(N)N}, \frac{\beta}{\mu(N)}\right\}.$$

We now show that $\underline{\sigma}_G$ must equal the third argument of the max operator. Note that the third argument is strictly greater than zero, so it remains to show that the third argument exceeds the second argument.

As previously noted, $\mu(N)$ is uniquely minimized at $N = 1$, and so

$$\mu(N) > p_G(1) > \beta > \gamma(N)\beta,$$

and $\mu(N) > \gamma(N)\beta$ may be shown by some algebra to be equivalent to the inequality

$$\frac{\beta}{\mu(N)} > \frac{\beta - (1-N)}{p_G(N)N}.$$

Thus

$$\underline{\sigma}_G = \frac{\beta}{\gamma(N)(1-N) + p_G(N)N}.$$

Now note that the derivative of the objective (B.3) with respect to σ_G is constant and equal to

$$\xi(N) \equiv p_G(N)N(R\Delta\pi - 1) - \gamma(N)(1-N),$$

which implies

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

Recall the definition of $\overline{N}(R) \equiv \sup\{N \in [0,1] : \xi(N) < 0\}$, and observe that $\xi(0) = -\gamma(0) < 0$. If $R\Delta\pi > 1$, then $\xi(N)$ is strictly increasing in N, with $\xi(1) = p_G(1)(R\Delta\pi - 1) > 0$, so $\overline{N}(R) \in (0,1)$. On the other hand, if $R\Delta\pi \leq 1$, then $\xi(N)$ is weakly decreases from $\xi(0) < 0$, so $\xi(N) < 0$ for all N and $\overline{N}(R) = 1$. In either case, $N < \overline{N}(R)$ implies $\xi(N) < 0$ and $N > \overline{N}(R)$ implies $\xi(N) > 0$.

So if $N < \overline{N}(R)$, then $\xi(N) < 0$ and therefore $\sigma_G = \underline{\sigma}_G$. If $N > \overline{N}(R)$, then $\xi(N) > 0$ and therefore $\sigma_G = \overline{\sigma}_G$, which equals 1 whenever $N \leq N^0$ and equals $\beta/[p_G(N)N]$ otherwise. Finally, if $N = \overline{N}(R)$ then either scheme is optimal, and by convention we choose the first scheme. For each scheme, the corresponding values of $\sigma_\emptyset$ and $T_\emptyset$ may then be computed from equations (B.1) and (B.2), and are reported in the lemma statement.

Finally, consider the extremal cases $N = 0, 1$. Note that the organization's objective function is continuous in $(\boldsymbol{\sigma}, \boldsymbol{T}, N)$, and the set of feasible, IC incentive schemes is characterized by a set of weak inequalities which are each continuous in N. Thus the constraint correspondence is continuous in N. This correspondence is not compact, 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. (Indeed, it is never necessary to offer a transfer larger than 1 to any employee to implement any desired N and promotion probabilities.) 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 schemes for $N \in (0,1)$ remain optimal in the limits $N = 0, 1$.

To complete the proof, we consider the case when $V \neq 1$. Note that the set of incentive-compatible schemes for a given V is the same as the set of schemes $(\boldsymbol{\sigma}, V\boldsymbol{T})$, ranging over

all schemes $(\boldsymbol{\sigma}, \boldsymbol{T})$ which are IC when $V = 1$. Let $\Pi^{IC}(\boldsymbol{\sigma}, \boldsymbol{T}, V', R')$ be the principal's payoff from promotions and bonus payments (excluding project payoffs, which are the same for all IC incentive schemes implementing N innovative projects) from the scheme $(\boldsymbol{\sigma}, V'\boldsymbol{T})$ when V' is the value of promotion to employees and R' is the value to the organization of promoting a high-quality employee. Since this profit function is a weighted sum of bonus payments and promotion payoffs, with weights V' and R', we may write

$$\Pi(\boldsymbol{\sigma}, \boldsymbol{T}, V', R') = V'\Pi(\boldsymbol{\sigma}, \boldsymbol{T}, 1, R'/V').$$

Hence an optimal incentive scheme for arbitrary V can be derived by solving for an optimal scheme when $V' = 1$ and $R' = R/V$, and then scaling up the bonus payments by a factor of V.

C Proof of Proposition 3

We first define three notions of set-ordered monotonicity which will be invoked in the proof.

Definition C.1. Let $F : X \rightrightarrows Y$ be a correspondence with $X, Y \subset \mathbb{R}$ and Y compact. Then F is *increasing* if $x > x'$ and $y \in F(x), y' \in F(x')$ implies that either $y, y' \in F(x) \cap F(x')$ or else $y > y'$. It is *strongly increasing* if it is increasing and additionally $y \in F(x) \cap \operatorname{Int} Y$ implies that $y > y'$. It is *strictly increasing* if it is increasing and additionally $y' \in F(x')$ implies $y' \notin F(x)$.

Note that these definitions are successively stronger notions of monotonicity—all strictly increasing functions are strongly increasing, and all strongly increasing functions are increasing, but the reverse implications do not hold. Also notice that the property of being strongly increasing is defined relative to the codomain of the correspondence.

By Proposition 2, an optimal incentive scheme either offers a bonus for choosing a safe project, or distorts promotion decisions, but not both. When inducing a given fraction N of employees to innovate, the optimal promotion-distortion scheme yields the organization a payoff of

$$\Pi^{Pr}(N; R) = f(N) + R\beta(\pi + \omega(N)(1 - \pi)),$$

where

$$\omega(N) \equiv \frac{p_G(N)N}{p_G(N)N + \gamma(N)(1 - N)}.$$

Meanwhile the optimal bonus scheme yields the organization a payoff of

$$\Pi^{B}(N; R) = \begin{cases} \Pi^{B}_{-}(N; R), & N \leq N^0 \\ \Pi^{B}_{+}(N; R), & N > N^0 \end{cases}$$

where

$$\Pi^B_-(N;R) = f(N) - V(\gamma(N)(1-N) + p_G(N)N - \beta) + R\beta\left(\pi + \frac{p_G(N)N}{\beta}(1-\pi)\right)$$

and

$$\Pi^B_+(N;R) = f(N) - V\frac{\gamma(N)(1-N)}{p_G(N)N}\beta + R\beta\pi_G.$$

Note that $\Pi^B_-(N^0;R) = \Pi^B_+(N^0;R)$, so that $\Pi^B(N;R)$ is continuous in N.

Let $N^{*,Pr}(R)$ be the set of optimizers of $\Pi^{Pr}(N;R)$ wrt N for a given R, and let $N^{*,B}_-(R)$ and $N^{*,B}_+(R)$ be the optimizers of $\Pi^B(N;R)$ wrt N for a given R subject to the constraints $N \leq N^0$ and $N \geq N^0$, respectively. Let $\Pi^{*,Pr}(R)$, $\Pi^{*,B}_-(R)$, and $\Pi^{*,B}_+(R)$ be the corresponding optimal profit functions. By the maximum theorem, each optimizer correspondence is compact- and non-empty-valued and upper hemicontinuous in R, and each optimal profit function is continuous in R. Further, note that R enters as an additive shift to $\Pi^B_+(N;R)$, so that $N^{*,B}_+$ is independent of R.

We next establish that $N^{*,Pr}(R)$ and $N^{*,B}_-(R)$ are strongly increasing in R, relative to their respective codomains $[0,1]$ and $[0,N^0]$, in the set-valued sense of Definition C.1. This follows from a strict version of Topkis's theorem (see Theorem 1 of [CITE]) provided $\partial^2\Pi/\partial R\partial N > 0$ for every $N < 1$.[11] Note that

$$\frac{\partial^2\Pi^{Pr}}{\partial R\partial N} = \beta\omega'(N)(1-\pi),$$

and

$$\omega'(N) = \omega(N) \times \left\{\frac{\gamma(N)}{p_G(N)N} - \frac{\gamma'(N)(1-N)}{p_G(N)N + \gamma(N)(1-N)}\right\}.$$

When $N = 0$ this simplifies to $\omega'(0) = 1$, while for every $N \in (0,1)$ this expression is strictly positive. Meanwhile,

$$\frac{\partial^2\Pi^B_-}{\partial R\partial N} = \gamma(N)(1-\pi),$$

which is strictly positive for $N < 1$ given that γ satisfies the same property.

Let $N^{*,B}(R)$ be the unconstrained maximizer of $\Pi^B(N;R)$ wrt N. This correspondence is also upper hemicontinuous by the maximum theorem, since $\Pi^B(N;R)$ is continuous in N everywhere.

[11] Formally, the hypotheses of Theorem 1 of [CITE] require that $\partial^2\Pi/\partial R\partial N > 0$ even at $N = 1$. However, the proof of that theorem requires only that the inequality hold strictly for interior values of N, and we rely on that slight generalization of the result.

Lemma C.1. Either $N^{*,B}(R) = N_+^{*,B}$ for all $R > 0$, or else there exists an $R^0 > 0$ such that

$$N^{*,B}(R) = \begin{cases} N_-^{*,B}(R), & R < R^0 \\ N_-^{*,B}(R) \cup N_+^{*,B}, & R = R^0 \\ N_+^{*,B}, & R > R^0. \end{cases}$$

In either case $N^{*,B}(R)$ is increasing in R everywhere, and in the latter case $N^{*,B}(R)$ is strictly increasing in R for $R \leq R^0$.

Proof. We first show that $N_-^{*,B}(R) = \{N^0\}$ for R sufficiently large. Note that

$$\frac{\partial \Pi_-^B}{\partial N} = \gamma(N)\left(1 - V\frac{\gamma'(N)(1-N)}{\gamma(N)} + R(1-\pi)\right).$$

Since $1 - V\frac{\gamma'(N)(1-N)}{\gamma(N)}$ is bounded below on $[0, N^0]$, there exists an R sufficiently large that $1 - V\frac{\gamma'(N)(1-N)}{\gamma(N)} + R(1-\pi) > 0$ for all $N \in [0, N^0]$. Thus for sufficiently large R, we have $\partial \Pi_-^B / \partial N > 0$ for all $N \leq N^0$, implying $N_-^{*,B}(R) = \{N^0\}$ for R large.

We further show that there exists an $R^\# \geq 0$ such that $N_-^{*,B}(R) = \{N^0\}$ for $R > R^\#$ and $N^0 \notin N_-^{*,B}(R)$ for $R^\#$. Let $R^\# = \inf\{R > 0 \ : \ N^0 \in N_-^{*,B}(R)\}$. By the work of the previous paragraph, we know that $R^\# < \infty$. If $R^\# = 0$, then the result is immediate. Otherwise, by upper hemicontinuity of $N_-^{*,B}$, we know that $N^0 \in N_-^{*,B}(R^\#)$. So

$$\Pi_-^B(N^0; R^\#) \geq \Pi_-^B(N; R^\#)$$

for all $N < N^0$. Then for any $R > R^\#$, we may write

$$\begin{aligned} \Pi_-^B(N^0; R) &= \Pi_-^B(N^0; R^\#) + (R - R^\#)\beta\left(\pi + \frac{p_G(N^0)N^0}{\beta}(1-\pi)\right) \\ &> \Pi_-^B(N; R^\#) + (R - R^\#)\beta\left(\pi + \frac{p_G(N)N}{\beta}(1-\pi)\right) \\ &= \Pi_-^B(N; R) \end{aligned}$$

for every $N < N^0$. So $N_-^{*,B}(R) = \{N^0\}$ for all $R > R^\#$. And by definition $N^0 \notin N_-^{*,B}(R)$ for any $R < R^\#$. This establishes the desired result.

We next argue that $\Delta\Pi^{*,B} \equiv \Pi_+^{*,B}(R) - \Pi_-^{*,B}(R)$ satisfies single-crossing in R. Suppose that there exists an $\widehat{R} > 0$ such that $\Delta\Pi^{*,B}(\widehat{R}) = 0$. Then it follows that

$$\Pi_-^B(N; \widehat{R}) \leq \Pi_+^{*,B}(\widehat{R})$$

for all $N \leq N^0$, with equality for at least one N. Now consider any $R > \widehat{R}$. Then we may write

$$\Pi_-^B(N; R) = \Pi_-^B(N; \widehat{R}) + (R - \widehat{R})\beta(1 + \omega(N)(1-\pi)).$$

Recall that $\omega(N) < 1$ for every $N < 1$, so that

$$\Pi_-^B(N;R) < \Pi_-^B(N;\widehat{R}) + (R-\widehat{R})\beta \leq \Pi_+^{*,B}(\widehat{R}) + (R-\widehat{R})\beta = \Pi_+^{*,B}(R)$$

for every $N \leq N^0$. It follows that $\Pi_+^{*,B}(R) > \Pi_-^{*,B}(R)$ for every $R > \widehat{R}$. Now consider any $R < \widehat{R}$. Then similar reasoning yields

$$\Pi_-^B(N;R) > \Pi_-^B(N;\widehat{R}) + (R-\widehat{R})\beta$$

for every $N \leq N^0$. In particular, for N satisfying $\Pi_-^B(N;\widehat{R}) = \Pi_+^{*,B}(\widehat{R})$, we have

$$\Pi_-^B(N;R) > \Pi_+^{*,B}(\widehat{R}) + (R-\widehat{R})\beta = \Pi_+^{*,B}(R).$$

Hence $\Pi_-^{*,B}(R) > \Pi_+^{*,B}(R)$ for every $R < \widehat{R}$.

Now suppose that $\min N_+^{*,B} = N^0$. We showed earlier that there exists an $R^\# \geq 0$ such that $N_-^{*,B}(R) = \{N^0\}$ for $R > R^\#$ and $N^0 \notin N_-^{*,B}(R)$ for $R < R^\#$. Hence whenever $R > R^\#$, $N_-^{*,B}(R) \subset N_+^{*,B}$ and so immediately $N^{*,B}(R) = N_+^{*,B}$. If $R^\# = 0$ we're done, so suppose $R^\# > 0$. Then whenever $R < R^\#$, $\Pi_-^{*,B}(R) > \Pi_-^B(N^0;R) = \Pi_+^{*,B}(R)$. So $N^{*,B}(R) = N_-^{*,B}(R)$ in this case. As for $R = R^\#$, by upper hemicontinuity of $N^{*,B}(R)$ at $R = R^\#$, $N_+^{*,B} \subset N^{*,B}(R^\#)$, and by upper hemicontinuity of $N^{*,B}(R)$ and $N_-^{*,B}(R)$ at $R = R^\#$, $N_-^{*,B}(R^\#) \subset N^{*,B}(R^\#)$. But also by definition of $N^{*,B}(R)$, $N^{*,B}(R^\#) \subset N_-^{*,B}(R^\#) \cup N_+^{*,B}$, meaning $N^{*,B}(R^\#) = N_-^{*,B}(R^\#) \cup N_+^{*,B}$. Thus $N^{*,B}(R)$ has the desired form for $R^0 = R^\#$.

Suppose instead that $\min N_+^{*,B} > N^0$. Recall that $N^0 \in N_-^{*,B}(R) = \{N^0\}$ for $R \geq R^\#$. Thus for $R \geq R^\#$ we must have $\Pi_+^{*,B}(R) > \Pi_-^{*,B}(R)$ given that $N^0 \notin N_+^{*,B}$. We further showed that $\Delta\Pi^{*,B}(R)$ satisfies single crossing in R. If $\Delta\Pi^{*,B}(0) \geq 0$, then $\Pi_+^{*,B}(R) > \Pi_-^{*,B}(R)$ for every $R > 0$, and so $N^{*,B}(R) = N_+^{*,B}$ for all $R > 0$. On the other hand, if $\Delta\Pi^{*,B}(0) < 0$, then by single crossing there exists a unique $R^0 \in (0, R^\#)$ such that $\Delta\Pi^{*,B}(R) < 0$ for $R < R^0$ while $\Delta\Pi^{*,B}(R) > 0$ for $R > R^0$. So $N^{*,B}(R) = N_-^{*,B}(R)$ for $R < R^0$ and $N^{*,B}(R) = N_+^{*,B}$ for $R > R^0$, and by upper hemicontinuity argument of the previous paragraph we may conclude that $N^{*,B}(R^0) = N_-^{*,B}(R) \cup N_+^{*,B}$. This immediately implies the desired form of $N^{*,B}(R)$ for this choice of R^0.

We complete the proof by establishing the claimed monotonicity properties. In the first case in the lemma statement monotonicity is trivial. So suppose we are in the latter case. We first prove strict monotonicity for $R \leq R^0$. Note that the arguments of the previous two paragraphs imply that $R^0 \leq R^\#$, so that $N^0 \notin N^{*,B}(R)$ for $R < R^0$. Fix $R \leq R^0$ and $R' < R^0$, and choose $n \in N^{*,B}(R)$ and $n' \in N^{*,B}(R')$. Since $N^0 \notin N^{*,B}(R') = N_-^{*,B}(R')$ and $\max N_-^{*,B}(R') \leq N^0$, it follows that $n' < N^0$. If $R < R^0$ then by the same argument $n < N^0$. On the other hand if $R = R^0$ and $n \geq N^0$, then immediately $n > n'$. So without loss we

may assume that $n < N^0$, in which case $n \in N_-^{*,B}(R)$. The desired result that $n > n'$ then follows from strong monotonicity of $N_-^{*,B}(R)$ provided that $0 \notin N_-^{*,B}(R)$. But note that

$$\frac{\partial \Pi_-^B}{\partial N}(0) = f'(0) - V\gamma'(0) + R\gamma(0)(1-\pi),$$

and since $f'(0), \gamma(0) > 0$ while $\gamma'(0) < 0$, it follows that $\partial \Pi_-^B / \partial N > 0$ at $N = 0$, so $0 \notin N_-^{*,B}(R)$, as desired.

Finally, consider global weak monotonicity. Fix any $R > R'$. If $R \leq R^0$ then monotonicity follows from strict monotonicity, while if $R' > R^0$ then monotonicity follows from the fact that $N^{*,B}(R) = N^{*,B}(R')$. So suppose that $R' = R^0$. Then for any $n' \in N^{*,B}(R')$, either $n' \in N_+^{*,B}$, in which case $n' \in N^{*,B}(R) = N_+^{*,B}$, or else $n' \in N_-^{*,B} \setminus N_+^{*,B}$, in which case $n' < n$ for every $n \in N_+^{*,B} = N^{*,B}(R)$. So monotonicity holds in this case. Finally, suppose that $R > R^0 > R'$. Then for any $n \in N^{*,B}(R) = N_+^{*,B}$, also $n \in N^{*,B}(R^0)$, meaning by strict monotonicity that $n > n'$ for every $n' \in N^{*,B}(R')$. So weak monotonicity holds in all cases. □

We now derive the threshold $R^* > 0$ claimed in the proposition statement. Specifically, we will establish existence of an $R^* > 0$ such that $\Delta\Pi^*(R) \equiv \Pi^{*,B}(R) - \Pi^{*,Pr}(R)$ satisfies $\Delta\Pi^*(R) < 0$ for $R < R^*$ while $\Delta\Pi^*(R) > 0$ for $R > R^*$.

The proof of Proposition 2 established single-crossing of $\Pi^{Pr}(N;R) - \Pi^B(N;R)$ in N for fixed R, with $\Pi^{Pr}(N;R) > \Pi^B(N;R)$ for $N < \overline{N}(R)$ and $\Pi^{Pr}(N;R) < \Pi^B(N;R)$ for $N > \overline{N}(R)$. The threshold value $\overline{N}(R)$ is equal to 1 for $R \leq V/(1-\pi)$, and is interior and strictly decreasing in R for $R > V/(1-\pi)$, with $\lim_{R\to\infty} \overline{N}(R) = 0$. Also, if $R < V/(1-\pi)$ then $\Pi^{Pr}(N;R) > \Pi^B(N;R)$ for all N, while if $R \geq V/(1-\pi)$ then $\Pi^{Pr}(\overline{N}(R);R) = \Pi^B(\overline{N}(R);R)$

We first establish that a promotion-distortion scheme is uniquely optimal for sufficiently small R, while a bonus scheme is uniquely optimal for sufficiently large R. The small-R result is immediate from the fact that $\Pi^{Pr}(N;R) > \Pi^B(N;R)$ for all N whenever $R < V/(1-\pi)$. So consider the limit of large R. Note that for any $R > 0$ and $N \leq N^\dagger$,

$$\frac{\partial \Pi^{Pr}}{\partial N} = f'(N) + R\beta\omega'(N)(1-\pi) > 0$$

given that $f'(N) \geq 0$ for $N \leq N^\dagger$ and $\omega'(N) > 0$ for $N < 1$. Thus $\min N^{*,Pr}(R) > N^\dagger$ for every R. But also for sufficiently large R, $\overline{N}(R) < N^\dagger$. For any such R, choose $N' \in N^{*,Pr}(R)$. Then $N' > N^\dagger > \overline{N}(R)$ and so we must have

$$\Pi^{*,B}(R) \geq \Pi^B(N';R) > \Pi^{Pr}(N';R) = \Pi^{*,Pr}(R).$$

Hence a bonus scheme is uniquely optimal for large R.

Suppose that $1 \in N^{*,Pr}(V/\Delta\pi)$. Since $\Pi^B(N; V/\Delta\pi) < \Pi^{Pr}(N; V/(1-\pi))$ for $N < 1$ while $\Pi^B(1; V/(1-\pi)) = \Pi^{Pr}(1; V/(1-\pi))$, it follows that $N^{*,B}(V/(1-\pi)) = \{1\}$ and $\Pi^{*,Pr}(V/(1-\pi)) = \Pi^{*,B}(V/(1-\pi))$. Further, for every $R > V/(1-\pi)$ the strong monotonicity of $N^{*,Pr}(R)$ in R implies that $N^{*,Pr}(R) = \{1\}$. Since $1 > \overline{N}(R)$ for $R > V/(1-\pi)$, it follows that

$$\Pi^{*,B}(R) \geq \Pi^B(1; R) > \Pi^{Pr}(1; R) = \Pi^{*,Pr}(R),$$

and thus a bonus scheme is uniquely optimal for every $R > V/(1-\pi)$. Finally, recall that $\Delta\Pi^*(R) < 0$ for $R < V/(1-\pi)$. The desired result therefore follows for $R^* = V/(1-\pi)$.

Suppose instead that $1 \notin N^{*,Pr}(V/(1-\pi))$. Then for all N,

$$\Pi^{*,Pr}(V/(1-\pi)) \geq \Pi^{Pr}(N; V/(1-\pi)) \geq \Pi^B(N; V/(1-\pi)),$$

with the first inequality strict when $N = 1$ and the second inequality strict when $N < 1$. Thus $\Delta\Pi^*(V/(1-\pi)) < 0$. And we established above that $\Delta\Pi^*(R) < 0$ for $R < V/(1-\pi)$, while $\Delta\Pi^*(R) > 0$ for R sufficiently large. We will show that $\Delta\Pi^*(R)$ satisfies single-crossing for $R > V/(1-\pi)$, which immediately implies existence of an $R^* > V/(1-\pi)$ with the desired properties.

Lemma C.2. For every N,

$$\omega(N) \leq \min\left\{\frac{p_G(N)N}{\beta}, 1\right\},$$

with the inequality strict whenever $N < 1$.

Proof. First note that straightforwardly $\omega(1) = 1$, while by assumption $p_G(1) > \beta$. So the inequality in the lemma statement holds with equality for $N = 1$. Suppose instead that $N < 1$. We previously established that $\omega'(N) > 0$ for all $N < 1$. Then since $\omega(1) = 1$, we must have $\omega(N) < 1$ for all $N < 1$. Further, the derivative of $p_G(N)N + \gamma(N)(1-N)$ is $\gamma'(N)(1-N) < 0$, so this expression is strictly decreasing in N and equal to $p_G(1)$ at $N = 1$. Therefore for all $N < 1$,

$$\omega(N) < \frac{p_G(N)N}{p_G(1)} < \frac{p_G(N)N}{\beta},$$

with the final inequality following from the assumption that $\beta < p_G(1)$. □

Consider any $\widehat{R} > V/(1-\pi)$ at which $\Delta\Pi^*(\widehat{R}) = 0$. Select any $n' \in N^{*,Pr}(\widehat{R})$ and $n'' \in N^{*,B}(\widehat{R})$. The fact that $\Pi^B(n''; \widehat{R}) \geq \Pi^{Pr}(N; \widehat{R})$ for all N implies that $n'' \geq \overline{N}(\widehat{R})$, while the fact that $\Pi^{Pr}(n'; \widehat{R}) \geq \Pi^B(N; \widehat{R})$ for all N implies that $n' \leq \overline{N}(\widehat{R})$.

Now fix any $R > \widehat{R}$. Since $\overline{N}(R)$ is strictly decreasing in R, we know that $\Pi^{Pr}(N;R) < \Pi^{B}(N;R)$ for every $N \geq \overline{N}(\widehat{R})$. Hence $\Pi^{*,B}(R) > \Pi^{Pr}(N;R)$ for all $N \geq \overline{N}(\widehat{R})$. So fix any $N < \overline{N}(\widehat{R})$. Note that $\overline{N}(\widehat{R}) < 1$, so we have

$$\begin{aligned}
\Pi^{Pr}(N;R) &= \Pi^{Pr}(N;\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \omega(N)(1-\pi)\right) \\
&< \Pi^{Pr}(N;\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \min\left\{\frac{p_G(N)N}{\beta}, 1\right\}(1-\pi)\right) \\
&\leq \Pi^{*,Pr}(\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \min\left\{\frac{p_G(N)N}{\beta}, 1\right\}(1-\pi)\right) \\
&= \Pi^{*,B}(\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \min\left\{\frac{p_G(N)N}{\beta}, 1\right\}(1-\pi)\right).
\end{aligned}$$

Now, note that $N < \overline{N}(\widehat{R}) \leq n''$ and $\min\{p_G(n)n/\beta, 1\}$ is weakly increasing in n. Therefore

$$\Pi^{Pr}(N;R) < \Pi^{*,B}(\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \min\left\{\frac{p_G(n'')n''}{\beta}, 1\right\}(1-\pi)\right) = \Pi^{B}(n'';R).$$

Hence $\Pi^{*,B}(R) > \Pi^{Pr}(N;R)$ for all N, meaning $\Delta\Pi^*(R) > 0$.

Finally, fix any $R < \widehat{R}$. Since $\overline{N}(R)$ is strictly decreasing in R, we know that $\Pi^{Pr}(N;R) > \Pi^{B}(N;R)$ for every $N \leq \overline{N}(\widehat{R})$. Hence $\Pi^{*,Pr}(R) > \Pi^{B}(N;R)$ for all $N \leq \overline{N}(\widehat{R})$. So fix any $N > \overline{N}(\widehat{R})$. Then we have

$$\begin{aligned}
\Pi^{B}(N;R) &= \Pi^{B}(N;\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \min\left\{\frac{p_G(N)N}{\beta}, 1\right\}(1-\pi)\right) \\
&\leq \Pi^{B}(N;\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \omega(N)(1-\pi)\right) \\
&\leq \Pi^{*,B}(\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \omega(N)(1-\pi)\right) \\
&= \Pi^{*,Pr}(\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \omega(N)(1-\pi)\right).
\end{aligned}$$

Further, $\omega(N)$ is a strictly increasing function, and as $n' \leq \overline{N}(\widehat{R}) < N$, we therefore have

$$\Pi^{B}(N;R) < \Pi^{*,Pr}(\widehat{R}) + (R-\widehat{R})\beta\left(\pi + \omega(n')(1-\pi)\right) = \Pi^{Pr}(n';R).$$

So $\Pi^{*,Pr}(R) > \Pi^{B}(N;R)$ for all N, meaning $\Delta\Pi^*(R) < 0$. Thus $\Delta\Pi^*(R)$ satisfies single-crossing, as desired.

We now turn to monotonicity properties of $N^*(R)$. First note that for $R < R^*$, $N^*(R) = N^{*,Pr}(R)$. Let $R^{\#} \equiv \inf\{R > 0 \ : \ 1 \in N^{*,Pr}(R)\}$. For every R,

$$\frac{\partial \Pi^{Pr}}{\partial N}(1;R) = f'(1) + R\beta\omega'(1)(1-\pi),$$

and as $\omega'(1)$ is finite while $f'(1) < 0$, for sufficiently small R we have $\partial\Pi^{Pr}/\partial N < 0$ at $N = 1$, so that $1 \notin N^{*,Pr}(R)$ for R sufficiently small. Thus $R^{\#} > 0$. At the other extreme,

$$\frac{\partial \Pi^{Pr}}{\partial N}(0;R) = f'(0) + R\beta(1-\pi).$$

Since $f'(0) > 0$, it follows that $\partial \Pi^{Pr}/\partial N > 0$ at $N = 0$, so that $0 \notin N^{*,Pr}(R)$ for any R. These results, combined with strong monotonicity of $N^{*,Pr}(R)$, collectively ensure that $N^{*,Pr}(R)$ satisfies strict monotonicity for $R \leq R^{\#}$. Meanwhile, strong monotonicity also implies that $N^{*,Pr}(R) = \{1\}$ for $R > R^{\#}$. So $N^{*,Pr}(R)$ is weakly monotone everywhere.

Now, recall that either $R^* = V/(1-\pi)$ and $1 \in N^{*,Pr}(R^*)$ while $N^{*,B}(R^*) = \{1\}$, or else $R^* > V/(1-\pi)$ and $\max N^{*,Pr}(R^*) \leq \overline{N}(R^*) < 1$ while $\min N^{*,B}(R^*) \geq \overline{N}(R^*)$. In both cases

$$N^*(R) = \begin{cases} N^{*,Pr}(R), & R < R^* \\ N^{*,Pr}(R) \cup N^{*,B}(R), & R = R^* \\ N^{*,B}(R), & R > R^*. \end{cases}$$

In the former case, Lemma C.1 ensures that $N^{*,B}(R^*) = \{1\}$ for all $R > R^*$, in which case the monotonicity properties of $N^{*,Pr}(R)$ ensure that $N^*(R)$ must be monotone everywhere, strictly monotone for $R \leq R^{\#}$, and constant for $R > R^{\#}$. Thus the desired monotonicity properties hold for $\overline{R} = R^{\#} > 0$. So consider the latter case. Recall the monotonicity properties of $N^{*,B}(R)$ established in Lemma C.1, and let $R^0 = 0$ if $N^{*,B}(R) = N_+^{*,B}$ for all $R > 0$. These properties, combined with the monotonicity properties of $N^{*,Pr}(R)$, ensure that $N^*(R)$ is monotone everywhere, strictly monotone for $R \leq \max\{R^*, R^0\}$, and constant for $R > \max\{R^*, R^0\}$. Thus the desired monotonicity properties hold for $\overline{R} = \max\{R^*, R^0\} > 0$.

D Proof of Proposition 4

We derive the optimal incentive scheme when $V = 1$, with the optimal incentive for general V obtainable by beginning with the optimal incentive scheme for $R' = R/V$ and $V' = 1$ and then scaling up transfers by a factor of V. See the proof of Proposition 2 for details on this procedure.

Let

$$\xi(N) \equiv (1-\gamma(N))(1-N)(1-R(\pi - \pi_B(N))) + (1-p_G(N))N.$$

We establish that the following is an optimal incentive scheme:

1. If $\xi(N) \leq 0$, the organization pays bonuses but promotes efficiently. In particular, letting $N^0 \in (0,1)$ be the solution to $p_G(N)N + (1-N) = \beta$,

   (a) If $N \leq N^0$,

$$\sigma_G = 1, \quad \sigma_\emptyset = \frac{\beta - p_G(N)N}{1-N}, \quad \sigma_B = 0, \quad T_B = \frac{\beta - \mu(N)}{(1-\gamma(N))(1-N)}, \quad T_G = T_\emptyset = 0.$$

(b) If $N > N^0$,

$$\sigma_G = \sigma_\emptyset = 1, \quad \sigma_B = \frac{\beta - Np_G(N) - (1-N)}{(1-p_G(N))N}, \quad T_B = \frac{1-\beta}{(1-p_G(N))N}, \quad T_G = T_\emptyset = 0.$$

2. If $\xi(N) > 0$, the organization distorts promotions but pays no bonuses:

$$\sigma_G = 1, \quad \sigma_\emptyset = \gamma(N) + (1-\gamma(N))\frac{\beta - \mu(N)}{1-\mu(N)}, \quad \sigma_B = \frac{\beta - \mu(N)}{1-\mu(N)}, \quad T_G = T_B = T_\emptyset = 0,$$

where $\mu(N) \equiv p_G(N)N + \gamma(N)(1-N)$.

We further show that when $N \notin \{0,1\} \cup \xi^{-1}(0)$, the optimal incentive scheme is unique. Finally, we establish that when $(1-N)(1-\gamma(N))$ is monotone, ξ satisfies single-crossing, so that defining

$$\overline{N}(R) \equiv \sup\{N \in [0,1] \ : \ \xi(N) < 0\},$$

the organization optimally pays bonuses when $N < \overline{N}(R)$ and optimally distorts promotions when $N > \overline{N}(R)$.

We first characterize an optimal policy for $N \in (0,1)$, and return to the extremal case afterward. We conjecture that the binding constraint involves the Nth employee, in the direction of choosing an innovative project. Thus we solve the relaxed problem in which the only IC constraint is

$$\gamma(N)(\sigma_G + T_G) + (1-\gamma(N))(\sigma_B + T_B) \geq \sigma_\emptyset + T_\emptyset,$$

and confirm that the resulting optimal scheme involves $\sigma_G + T_G \geq \sigma_B + T_B$ and a binding Nth-employee IC constraint. Thus the solution to the relaxed problem satisfies the full set of IC constraints.

We first argue that both constraints bind at the optimum of the relaxed problem. Suppose first that the feasibility constraint were slack. Then trivially profits are maximized by setting $\boldsymbol{T} = (0,0,0)$ and $\boldsymbol{\sigma} = (1,1,1)$, since this scheme satisfies the Nth-employee IC constraint. But then a measure 1 of employees are promoted, violating feasibility. So the feasibility constraint binds at the optimum. Suppose next that the Nth-employee IC constraint were slack. Then the optimal scheme would pay no bonuses and promote efficiently subject to feasibility, setting

$$\sigma_G = 1, \quad \sigma_\emptyset = \min\left\{\frac{\beta - p_G(N)N}{1-N}, 1\right\}, \quad \sigma_B = \max\left\{\frac{\beta - p_G(N)N - (1-N)}{(1-p_G(N))N}, 0\right\}.$$

Note that $\sigma_B < 1$, and so if $\sigma_\emptyset = 1$ then

$$\sigma_\emptyset + T_\emptyset = 1 > \gamma(N)(\sigma_G + T_G) + (1-\gamma(N))(\sigma_B + T_B),$$

violating IC. On the other hand if $\sigma_\emptyset < 1$ then $\sigma_B = 0$, in which case

$$\gamma(N)(\sigma_G + T_G) + (1 - \gamma(N))(\sigma_B + T_B) = \gamma(N).$$

Thus

$$\sigma_\emptyset + T_\emptyset - (\gamma(N)(\sigma_G + T_G) + (1 - \gamma(N))(\sigma_B + T_B)) = \frac{\beta - \mu(N)}{1 - N}.$$

Recall that $\mu(N)$ is strictly decreasing in N, and so is maximized at $N = 0$, where $\mu(0) = p_G(0) < \beta$. Thus the rhs of the previous expression is strictly positive, again violating IC. So the IC constraint must also bind. (This establishes one of the claims of the previous paragraph.)

We now argue that $T_G = T_\emptyset = 0$ at the optimum of the relaxed problem. Given that the IC constraint binds at the optimum, the Lagrange multiplier λ on that constraint must be strictly positive. The derivative of the Lagrangian wrt $T_\emptyset$ at the optimum is therefore

$$\frac{\partial \mathscr{L}}{\partial T_\emptyset} = -(1 - N) - \lambda < 0,$$

implying $T_\emptyset = 0$ at the optimum. Now, suppose that (T_G, T_B) are both perturbed by an amount Δ. The derivative of the Lagrangian wrt Δ is

$$\frac{\partial \mathscr{L}}{\partial \Delta} = -N + \lambda,$$

which much be weakly negative at the optimum in order for there not to be an improvement involving an increase of (T_G, T_B). Hence $\lambda \leq N$. Additionally, the derivative of the Lagrangian wrt T_G is

$$\frac{\partial \mathscr{L}}{\partial T_G} = -p_G(N)N + \lambda\gamma(N) \leq N(\gamma(N) - p_G(N)),$$

which is strictly negative given that $\gamma(N) < p_G(N)$ for all $N > 0$. So it must be that $T_G = 0$ at the optimum.

We now argue that $\sigma_G = 1$ at the optimum. Suppose first that $T_B > 0$ at the optimum. Then the IC constraint implies that $\sigma_\emptyset > 0$, for otherwise the rhs would be zero while the lhs is strictly greater than 0. Suppose additionally that $\sigma_G < 1$ at the optimum. Consider a new promotion scheme $(\sigma'_G, \sigma'_\emptyset, \sigma'_B) = (\sigma_G + \Delta', \sigma_\emptyset - \Delta)$, where Δ, Δ' are sufficiently small that $\sigma_\emptyset - \Delta > 0$ and $\sigma_G + \Delta' < 1$, and are chosen relative to one another to ensure that the feasibility constraint continues to be satisfied. This new scheme raises the payoff to choosing an innovative project and lowers the payoff to choosing a safe project. Therefore the IC constraint may be satisfied by appropriately lowering T_B, which is feasible given $T_B > 0$ if Δ, Δ' are chosen sufficiently small. This new incentive scheme is feasible and IC, pays

strictly lower bonuses, and reallocates promotions from employees choosing safe projects to employees achieving good outcomes. Thus total profits must go up, contradicting the presumed optimality of the original scheme. Thus $\sigma_G = 1$ if $T_B > 0$ at the optimum.

Suppose instead that $T_B = 0$ at the optimum. Then the optimal promotion probabilities must solve the reduced optimization problem with no bonuses. With bonuses eliminated, the firm's problem is to maximize the value of promoted employees

$$N(p_G(N)\pi_G\sigma_G + (1 - p_G(N))\pi_B(N)\sigma_B) + (1 - N)\pi\sigma_\emptyset$$

subject to the constraints

$$\begin{cases} \gamma(N)\sigma_G + (1 - \gamma(N))\sigma_B \geq \sigma_\emptyset, \\ \beta \geq N(p_G(N)\sigma_G + (1 - p_G(N))\sigma_B) + (1 - N)\sigma_\emptyset. \end{cases}$$

The arguments used to establish that both constraints bind in the problem with bonuses continue to hold here, and so both constraints must bind in the reduced problem without bonuses. Use the binding IC constraint to eliminate $\sigma_\emptyset$ from the problem, leaving the objective

$$(Np_G(N)\pi_G + (1 - N)\gamma(N)\pi)\sigma_G + (N(1 - p_G(N))\pi_B(N) + (1 - N)(1 - \gamma(N))\pi)\sigma_B,$$

subject to the reduced feasibility constraint

$$\beta \geq (Np_G(N) + (1 - N)\gamma(N))\sigma_G + (N(1 - p_G(N)) + (1 - N)(1 - \gamma(N))\sigma_B,$$

which must bind at the optimum. The derivatives of the corresponding Lagrangian are

$$\frac{\partial \mathscr{L}}{\partial \sigma_G} = p_G(N)N\left(\pi_G - \lambda\left(1 + \frac{\gamma(N)(1 - N)}{p_G(N)N}\right)\right)$$

and

$$\frac{\partial \mathscr{L}}{\partial \sigma_B} = (1 - p_G(N))N\left(\pi_B(N) - \lambda\left(1 + \frac{(1 - \gamma(N))(1 - N)}{(1 - p_G(N))N}\right)\right),$$

where $\lambda > 0$ at the optimum given that the feasibility constraint must bind. Given that $N > 0$, we have $\gamma(N) < p_G(N)$ and thus $1 - \gamma(N) > 1 - p_G(N)$. It follows that

$$\pi_G - \lambda\left(1 + \frac{\gamma(N)(1 - N)}{p_G(N)N}\right) > \pi_B(N) - \lambda\left(1 + \frac{(1 - \gamma(N))(1 - N)}{(1 - p_G(N))N}\right).$$

Hence at the optimum, at most one of σ_G and σ_B may be interior, and $\sigma_G \geq \sigma_B$. In particular, if $\sigma_G < 1$ then $\sigma_B = 0$. But then the constraint reduces to

$$\beta = (Np_G(N) + (1 - N)\gamma(N))\sigma_G.$$

Note that $Np_G(N) + (1-N)\gamma(N)$ has derivative $(1-N)\gamma'(N) < 0$, so that the expression is maximized at $N = 0$, where it equals $p_G(0)$. Thus in particular the rhs of the constraint is strictly less than $p_G(0)$, which by assumption is strictly less than β, a contradiction. So it must be that $\sigma_G = 1$ if $T_B = 0$ at the optimum.

Going forward, we restrict attention to the problem with $T_G = T_\emptyset = 0$, $\sigma_G = 1$, and both constraints enforced with equality:

$$\gamma(N) + (1-\gamma(N))(\sigma_B + T_B) = \sigma_\emptyset, \tag{D.1}$$

$$\beta = N(p_G(N) + \sigma_B(1-p_G(N))) + (1-N)\sigma_\emptyset. \tag{D.2}$$

Use equation (D.1) to eliminate $\sigma_\emptyset$ from the problem, yielding the reduced constraint

$$\beta = p_G(N)N + \gamma(N)(1-N) + ((1-p_G(N))N + (1-\gamma(N))(1-N))\sigma_B + (1-\gamma(N))(1-N)T_B.$$

This constraint may then be used to further eliminate T_B from the problem. After dropping an additive constant, this reduction yields the following optimization problem for σ_B:

$$\max_{\sigma_B \in [0,1]} \frac{(1-p_G(N))N}{(1-\gamma(N))(1-N)}\xi(N)\sigma_B,$$

subject to the boundary constraints that $\sigma_\emptyset \in [0,1]$ and $T_B \geq 0$. This objective is strictly increasing in σ_B if $\xi(N) > 0$, and is strictly decreasing in σ_B if $\xi(N) < 0$. (When $\xi(N) = 0$, any feasible scheme in this reduced class is optimal.)

Consider first the case $\xi(N) > 0$. Recall that the reduced constraint implies a linear relationship between σ_B and T_B. Thus if $\xi(N)$ is positive, σ_B is optimally set as high as possible subject to $T_B \geq 0$, with T_B specified by the reduced constraint, and the boundary constraints on σ_B and $\sigma_\emptyset$, with $\sigma_\emptyset$ specified by the IC constraint. The reduced constraint combined with $T_B \geq 0$ imply that

$$\sigma_B \leq \overline{\sigma}_B \equiv \frac{\beta - \mu(N)}{1-\mu(N)},$$

where $\mu(N) \equiv p_G(N)N + \gamma(N)(1-N)$. Note that $\mu'(N) = \gamma'(N)(1-N) < 0$, so μ is maximized at $N = 0$, where $\mu(0) = p_G(0) < \beta$. Thus $\mu(N) < \beta$ for every N, implying $\overline{\sigma}_B \in (0,1)$, so this value of σ_B satisfies with its boundary constraint. Further, the IC constraint with $\sigma_B = \overline{\sigma}_B$ and $T_B = 0$ implies

$$\sigma_\emptyset = \gamma(N) + (1-\gamma(N))\overline{\sigma}_B,$$

which is a weighted average of 1 and a quantity in $(0,1)$, with a weight of $\gamma(N) < 1$ on 1. Thus $\sigma_\emptyset \in (0,1)$ for these choices of σ_B and T_B. It follows that $\sigma_B = \overline{\sigma}_B$ and $T_B = 0$ are the unique optimal choices of (σ_B, T_B) when $\xi(N) > 0$.

Now consider the case $\xi(N) < 0$. Now σ_B is optimally set as small as possible subject to the boundary constraints on T_B, σ_B, and $\sigma_\emptyset$, with T_B specified by the reduced constraint and $\sigma_\emptyset$ specified by the IC constraint. We first check whether $\sigma_B = 0$ satisfies all boundary constraints. In this case the reduced constraint yields

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

which satisfies the boundary constraint $T_B \geq 0$ given that $\mu(N) < \beta$. Substituting into the IC constraint then yields

$$\sigma_\emptyset = \gamma(N) + \frac{\beta - \mu(N)}{1-N} = \frac{\beta - p_G(N)N}{1-N}.$$

This expression is always strictly positive, but is less than 1 only for $N \leq N^0$, where N^0 is the unique solution to $p_G(N)N + (1-N) = \beta$. Thus whenever $N \leq N^0$, the optimal incentive scheme sets $\sigma_B = 0$ and $T_B = (\beta - \mu(N))/((1-\gamma(N))(1-N))$. On the other hand, when $N > N^0$ the boundary constraint $\sigma_\emptyset \leq 1$ must bind at the optimum. In this case the feasibility constraint implies that

$$\sigma_B = \frac{\beta - Np_G(N) - (1-N)}{(1-p_G(N))N}.$$

The IC constraint may then be used to recover T_B, yielding

$$T_B = 1 - \sigma_B = \frac{1-\beta}{(1-p_G(N))N}.$$

We now characterize the sign of ξ. Recall that $\pi_B(N)$ is increasing in N, and so $\Delta\pi(N) \equiv \pi - \pi_B(N)$ is decreasing in N. So suppose first that $R \leq 1/\Delta\pi(0)$. Then $\xi(N) > 0$ for all $N \in (0,1)$. Suppose instead that $R > 1/\Delta\pi(0)$. Then $\xi(1) = 1 - p_G(1) > 0$, while $\xi(0) = (1-\gamma(0))(1 - R\Delta\pi(0)) < 0$ given that $\gamma(0) < 1$. So ξ is negative for N sufficiently close to 0, and positive for N sufficiently close to 1. We further show that ξ satisfies single-crossing whenever $(1-N)(1-\gamma(N))$ is weakly decreasing. First note that whenever $N > 0$ and $R\Delta\pi(N) \geq 1$, $\xi(N) > 0$. So consider N such that $R\Delta\pi(N) < 0$. In that case the fact that $\Delta\pi(N)$ and $(1-\gamma(N))(1-N)$ are both decreasing in N implies that

$$(1-\gamma(N))(1-N)(1-R\Delta\pi(N))$$

is increasing in N. Meanwhile $(1-p_G(N))N$ has derivative $1-\gamma(N)$, and so is strictly increasing in N. Thus $\xi(N)$ is strictly increasing in N whenever $R\Delta\pi(N) < 0$. It follows that ξ satisfies single-crossing.

Whenever ξ satisfies single-crossing, the form of an optimal incentive scheme depends on how N compares to the threshold $\overline{N}(R)$. The arguments of the previous paragraph imply that $\overline{N}(R) = 0$ for $R \leq 1/\Delta\pi(0)$, and $\overline{N}(R) < 1$ for all R. Further,

$$\frac{\partial \xi}{\partial R} = (1 - \gamma(N))(1 - N)\pi_B(N),$$

which is strictly decreasing in N whenever $(1 - \gamma(N))(1 - N)$ is decreasing. So $\overline{N}(R)$ must be strictly increasing in R whenever it is nonzero.

Finally, consider the extremal cases $N = 0, 1$. Note that the organization's objective function is continuous in $(\boldsymbol{\sigma}, \boldsymbol{T}, N)$, and the set of feasible, IC incentive schemes is characterized by a set of weak inequalities which are each continuous in N. Thus the constraint correspondence is continuous in N. This correspondence is not compact, 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. (Indeed, it is never necessary to offer a transfer larger than 1 to any employee to implement any desired N and promotion probabilities.) 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 schemes for $N \in (0, 1)$ remain optimal in the limits $N = 0, 1$.

E Proof of Proposition 5

In light of Proposition 4, an optimal incentive scheme takes one of two forms: it either distorts promotions, or offers a bonus in case of a bad outcome on an innovative project, but not both. We proceed by computing total profits under each scheme as a function of N, and comparing profits from each scheme under their respective optimal choices of N.

Under the promotion-distortion scheme, total payoffs from implementing N innovative projects are

$$\begin{aligned}\Pi^{Pr}(N; R) = f(N) + R \Bigg\{ & p_G(N)N + (1 - p_G(N))N\frac{\beta - \mu(N)}{1 - \mu(N)}\pi_B(N) \\ & + (1 - N)\left(\gamma(N) + (1 - \gamma(N))\frac{\beta - \mu(N)}{1 - \mu(N)}\right)\pi \Bigg\}.\end{aligned}$$

By Bayes' rule,

$$\pi_B(N) = \frac{\left(1 - \frac{1}{N}\int_0^N q(n)\, dn\right)\pi}{1 - p_G(N)} = \frac{\pi - p_G(N)}{1 - p_G(N)}.$$

Inserting this expression into $\Pi^{Pr}(N;R)$ and simplifying yields

$$\Pi^{Pr}(N;R) = f(N) + R\beta\left(\pi + (1-\pi)\frac{1-\beta}{\beta}\frac{p_G(N)N}{1-\mu(N)}\right).$$

Meanwhile under the bonus scheme, total payoffs from implementing $N \leq N^0$ innovative projects are

$$\begin{aligned}\Pi^B_-(N;R) = f(N) &- \frac{\beta-\mu(N)}{(1-\gamma(N))(1-N)}(1-p_G(N))N \\ &+ R\left(p_G(N)N + \frac{\beta - p_G(N)N}{1-N}(1-N)\pi\right),\end{aligned}$$

or equivalently

$$\Pi^B_-(N;R) = f(N) - \frac{\beta-\mu(N)}{(1-\gamma(N))(1-N)}(1-p_G(N))N + R\beta\left(\pi + (1-\pi)\frac{p_G(N)N}{\beta}\right).$$

And when $N > N^0$, total payoffs are

$$\Pi^B_+(N;R) = f(N) - (1-\beta) + R\left(p_G(N)N + (1-N)\pi + (\beta - p_G(N)N - (1-N))\pi_B(N)\right).$$

Inserting the expression for $\pi_B(N)$ above and simplifying yields

$$\Pi^B_+(N;R) = f(N) - (1-\beta) + R\beta\left(\pi + (1-\pi)\frac{1-\beta}{\beta}\frac{p_G(N)}{1-p_G(N)}\right).$$

We define $\Pi^B(N;R)$ for $N \in [0,1]$ by letting $\Pi^B(N;R) = \Pi^B_-(N;R)$ for $N \leq N^0$ and $\Pi^B(N;R) = \Pi^B_+(N;R)$ for $N > N^0$. Note that $\Pi^B(N;R)$ is a continuous function of N given that $\Pi^B_-(N^0;R) = \Pi^B_+(N^0;R)$.

Define $\Pi^{*,Pr}(R) = \sup_{N\in[0,1]}\Pi^{Pr}(N;R)$, and similarly $\Pi^{*,B}_-(R) = \sup_{N\in[0,N^0]}\Pi^B(N;R)$, $\Pi^{*,B}_+(R) = \sup_{N\in[N^0,1]}\Pi^B(N;R)$, and $\Pi^{*,B}(R) = \sup_{N\in[0,1]}\Pi^B(N;R)$. Each of these functions is the maximized value of a continuous objective over a compact objective set, and so each is a continuous function of R.

We first show that $\Pi^{*,Pr}(R) > \Pi^{*,B}(R)$ for sufficiently small $R > 0$. The proof of Proposition 4 established that for $R \leq V/(\pi - \pi_B(0))$, that $\Pi^{Pr}(N;R) \geq \Pi^B(N;R)$ for all N, with the inequality strict when $N \in (0,1)$. Further, when $R = 0$, $\Pi^{Pr}(N;0)$ is uniquely maximized at $N = N^\dagger \in (0,1)$. Thus in particular $\Pi^{*,Pr}(0) > \Pi^{Pr}(0;0), \Pi^{Pr}(1;0)$. Since $\Pi^{*,Pr}(R)$ is continuous in R, it follows that for $R > 0$ sufficiently small, we also have $\Pi^{*,Pr}(R) > \Pi^{Pr}(0;R), \Pi^{Pr}(1;R)$. Hence for $R > 0$ sufficiently small, $\Pi^{*,Pr}(R) \geq \Pi^{Pr}(N;R) > \Pi^B(N;R)$ for all $N \in (0,1)$, and further $\Pi^{*,Pr}(R) > \Pi^{Pr}(N;R) \geq \Pi^B(N;R)$ for $N = 0,1$. Thus $\Pi^{*,Pr}(R) > \Pi^{*,B}(R)$ for R sufficiently small.

We complete the proof by showing that $\Pi^{*,B}(R) > \Pi^{*,Pr}(R)$ for sufficiently large R. Define $\Gamma(N) \equiv p_G(N)N/(1-\mu(N))$, and let $\widehat{N}$ be any maximizer of $\Gamma(N)$ on $[0,1]$. (Since $\Gamma(N)$ is continuous, at least one such maximizer must exist.) For any R, it must be the case that the payoff from the optimal promotion-distortion scheme is bounded above by

$$\Pi^{*,Pr}(R) \leq f(N^\dagger) + R\beta\left(\pi + (1-\pi)\frac{1-\beta}{\beta}\Gamma(\widehat{N})\right).$$

Meanwhile, for any R the payoff from the optimal bonus scheme is bounded below by

$$\Pi^{*,B}(R) \geq \Pi^B_-(N^0;R) = f(N^0) - (1-\beta) + R\beta\left(\pi + (1-\pi)\frac{p_G(N^0)N^0}{\beta}\right).$$

Therefore

$$\Pi^{*,B}(R) - \Pi^{*,Pr}(R) \geq f(N^0) - f(N^\dagger) - (1-\beta) + R(1-\pi)\left(p_G(N^0)N^0 - (1-\beta)\Gamma(\widehat{N})\right).$$

We now show that $\Gamma(N) < p_G(N^0)N^0/(1-\beta)$ for all N, so that in particular $(1-\beta)\Gamma(\widehat{N}) < p_G(N^0)N^0$. Note that since $\mu(N)$ is strictly decreasing in N, we have $\mu(N) \leq \mu(0) = \gamma(0) = p_G(0) < \beta$ for all N, and hence $1-\mu(N) > 1-\beta$ for all N. We may therefore bound $\Gamma(N)$ above as $\Gamma(N) < p_G(N)N/(1-\beta)$ for all N. Since $p_G(N)N$ is strictly increasing in N, it follows that $\Gamma(N) < p_G(N^0)N^0/(1-\beta)$ for $N \leq N^0$. We may derive another upper bound on $\Gamma(N)$ by noting that

$$1-\mu(N) = (1-p_G(N))N + (1-\gamma(N))(1-N) \geq (1-p_G(N))N.$$

Hence

$$\Gamma(N) \leq \frac{p_G(N)N}{(1-p_G(N))N} = \frac{p_G(N)}{1-p_G(N)}$$

for all N. Since $p_G(N)$ is strictly decreasing in N, this upper bound is strictly decreasing in N, and so in particular for all $N > N^0$ we have

$$\Gamma(N) < \frac{p_G(N^0)}{1-p_G(N^0)}.$$

Now, by definition N^0 satisfies $p_G(N^0)N^0 + (1-N^0) = \beta$, which is equivalently $1-p_G(N^0) = (1-\beta)/N^0$. Thus

$$\Gamma(N) < \frac{p_G(N^0)N^0}{1-\beta}$$

for all $N > N^0$, as desired.

Given that $(1-\beta)\Gamma(\widehat{N}) < p_G(N^0)N^0$ and N^0 and $\widehat{N}$ are both independent of R, we have that $\Pi^{*,B}(R) - \Pi^{*,Pr}(R)$ is bounded below by an expression which is an affine, strictly increasing function of R. Thus $\Pi^{*,B}(R) > \Pi^{*,Pr}(R)$ for sufficiently large R, as desired.