Incentivizing Autonomous Workers[*]

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

June 20, 2023

Abstract

Workers in an important category of jobs select tasks autonomously. We study the tradeoff between monetary bonuses and non-monetary prizes as tools for guiding their choices. An optimal incentive scheme prioritizes workers for prizes in return for taking on underserved tasks, and this prioritization increases as incentives power up. Bonuses may additionally be used when incentives are sufficiently high-powered, but the optimal bonus is often non-monotone in the strength of incentives. Our results have important implications for the design of worker reward programs on freelancing platforms such as Uber and Airbnb.

Keywords: Incentive pay, non-monetary rewards, worker autonomy, task selection, platform economics

JEL Classification: D82, D86, M51, M52

[*]An earlier version of this paper circulated under the title "Incentive Design for Talent Discovery". We thank Frank Yang and Zi Yang Kang for exemplary research assistance; Heski Bar-Isaac, Karen Bernhardt-Walther, Ben Brooks, Guido Friebel, Bob Gibbons, Boyan Jovanovic, Anton Kolotilin, Fei Li, George Mailath, Debraj Ray, Marzena Rostek, Kathryn Shaw, and Michael Waldman for valuable discussions; and seminar participants and our colleagues and students for helpful comments and suggestions.

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

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

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

1 Introduction

An important category of jobs afford workers autonomy regarding the tasks or projects that they take on. This flexibility is a defining feature of freelance work, where workers choose projects or clients rather than being assigned them. Freelance roles have become increasingly important due to the growth of online platforms which match freelancers and clients in the markets for ride-hailing (Uber, Lyft, Curb), lodging (Airbnb, VRBO), food delivery (Grubhub, Doordash, UberEats), coding and design work (upwork), and other services (Taskrabbit, Mechanical Turk).[1] Autonomy is similarly a central aspect of basic research, for instance as conducted by university faculty, who are free to allocate their time across research projects as well as between research and non-research tasks such as teaching. In addition, autonomy is a well-documented feature of jobs at firms utilizing "flat" organizational structures[2] or which reserve time for employees to pursue side projects.[3]

Whenever workers enjoy autonomy, organizations may need to reward them for undertaking particular tasks to ensure that all necessary work is completed. For instance, ride-hailing platforms may need to reward drivers for working in high-demand locations or during rush hour. Similarly, universities may need to reward faculty for covering teaching shortfalls or undertaking burdensome administrative work such as department chairmanships. And software companies may need to reward engineers for prioritizing stability and usability enhancements over more glamorous projects such as launching a new product.[4]

[1] The growth of freelance work on these platforms is a phenomenon which has been variously described as the "gig economy" or "platform economy" in the media.

[2] Such structures are perhaps most commonly associated with tech firms. For instance, the game publisher Valve emphasizes in their employee handbook that employees choose their own projects under a "Flatland" organizational structure (https://cdn.cloudflare.steamstatic.com/apps/valve/Valve_NewEmployeeHandbook.pdf). The software developer Github relied on a similar organizational structure described as "open allocation" for a number of years (Burton et al. 2017) Some firms in other industries have been described as following similar organizing principles, for instance W. L. Gore & Associates's "lattice organization" (Grønning 2016) and Sun Hydraulics's "horizontal management" system (Hill and Suesse 2003).

[3] Prominent examples of programs offering dedicated time for side projects include 3M's "15% rule" (https://www.3m.co.uk/3M/en_GB/careers/culture/15-percent-culture/) and Google's "20% time" (https://abc.xyz/investor/founders-letters/2004-ipo-letter/), both of which refer to a percentage of an employee's time which they are free to allocate.

[4] A former tech lead of the Google sheets team has described how a preference for product launches over bug fixes distorted the projects that software engineers spent time on (Lloyd 2022).

In general, such rewards can take either monetary or non-monetary forms. In the ride-hailing market, drivers could be rewarded through cash bonuses or valuable non-cash prizes such as priority matching with high-value trips.[5] Similarly, hosts on Airbnb could be rewarded with cash rebates or increased visibility to guests.[6] And engineers at software companies could be rewarded with bonuses or an increased chance of promotion. In this paper, we study the optimal use of both types of rewards to incentivize autonomous workers.

Both monetary and non-monetary rewards are costly to deploy. While the financial cost of a monetary reward is straightforward, preferentially awarding prizes like the ones discussed above distorts their allocation. For instance, if Google promotes engineers who have completed underserved projects into management roles over more talented peers, it may reduce the productivity of its engineering teams. Similarly, if Uber drivers who work during high-demand periods are prioritized for high-value trips, riders may end up waiting longer to be picked up. It is therefore not clear ex ante how an organization should balance the costs of these two forms of incentives.

We build a simple, flexible model of autonomous work to answer this question. In our model, a large population of workers within an organization choose freely between two alternative tasks. Workers have heterogeneous intrinsic preferences over tasks, and the organization's goal is to incentivize a specified proportion of workers to choose each task. It can reward particular task choices with a monetary *bonus*, and it can additionally commit to preferentially distribute a valuable non-monetary *prize* which is in limited supply.[7] The organization has distributional preferences over prizes, summarized by a worker-specific match value that varies across the population and is realized only after tasks are chosen.

We characterize the optimal incentive scheme, which involves setting a prize priority and monetary bonus for choosing an underserved task. The organization commits in advance to the match standard necessary to earn a prize, conditional on a worker's task choice, and

[5]Uber prominently uses monetary rewards to reallocate drivers toward high-demand areas via its surge-pricing program (Lu, Frazier, and Kislev 2018). Its "Uber Pro" program rewards drivers with both money and non-cash prizes: a status badge visible to riders, priority matching for airport pickups, and flexibility to filter trips by origin and destination (https://www.uber.com/us/en/drive/uber-pro/rewards/).

[6]Airbnb's "Superhost" program rewards hosts with cash rebates as well as non-cash prizes: featured placement in promotional emails and a status badge visible to guests, who can filter for Superhosts during searches (https://www.airbnb.com/d/superhost).

[7]We assume that the supply of prizes is fixed throughout our analysis. We discuss this assumption further in Section 9.

it optimally employs a (strictly) lower standard for workers selecting an underserved task. Depending on the environment, the optimal scheme may couple this priority with a bonus. The optimal priority is determined by a first-order condition equating the (rising) marginal cost of the priority with the (constant) marginal cost of bonuses.

Our main results concern how the optimal use of each incentive tool varies with desired incentive power, as measured by the fraction of workers the organization wishes to undertake an underserved task. In general, the optimal prize priority increases as incentives grow stronger (Theorem 1). In contrast, the optimal bonus may vary non-monotonically with incentive power (Theorem 2). In a benchmark class of models in which worker match values are uniformly-distributed, this non-monotonicity manifests as a hump-shaped response to rising incentive power (Proposition 2).

These results are shaped by two novel effects which illuminate the distinction between monetary and prize rewards. First, a *group-size effect* reduces the marginal cost of prize incentives relative to bonuses as more workers choose the underserved task. Intuitively, prizes are in fixed supply and are reallocated across workers, while bonuses are in variable supply and are awarded to a single group of workers. This distinction means that only prize incentives directly penalize workers choosing the overserved task, and this impact is larger the fewer workers are in this group. As a result, the optimal prize priority rises as incentives power up. This effect also favors a reduction in the use of bonuses.

The optimal use of bonuses is additionally shaped by a *gap-size effect* stemming from the organization's distributional preferences over prizes. These preferences mean that the marginal cost of prize incentives rises as prizes are reallocated to increasingly mismatched workers, favoring the use of money on the margin to increase the gap between the rewards for each task. The gap-size effect tends to dominate when desired incentives are weak, while the group-size effect wins out when they are strong, generating a potentially non-monotone response of the optimal bonus.

We additionally explore how the size of the prize endowment affects the shape of an optimal scheme. We show that the bonus and prize priority are largest when prizes are particularly scarce or plentiful, and both shrink for intermediate prize endowments (Theorem 3). This effect is driven by non-uniformity of the match distribution: When the marginal prize-winner is in the tails of the match distribution, significant incentives require a large prize priority; on the other hand, near the center of the distribution incentives can be cheaply provisioned with a small priority.[8] When match value is uniformly distributed, this force is absent and we show that the optimal prize priority and bonus are independent of the prize endowment (Proposition 3).[9]

Our results provide a number of testable predictions for incentive design in real-world organizations. In Section 7 we translate our abstract results into concrete predictions about incentive schemes in several contexts. We then propose hypothetical datasets involving plausibly measurable covariates which could be used to evaluate these predictions.

The remainder of the paper is structured as follows. Section 1.1 discusses related literature. Section 2 describes our model, and Section 3 formalizes the incentive schemes we study. We derive an optimal incentive scheme in Section 4. We characterize how the optimal scheme changes with the magnitude of required incentives in Section 5 and with the size of the prize endowment in section 6. We discuss testable predictions of our model in Section 7. Section 8 provides further details regarding the optimal provision of incentives when prize or monetary incentives become scarce. We offer concluding remarks and directions for future research in Section 9. All proofs are relegated to the Appendix.

1.1 Related literature

Our paper contributes to a large body of work studying the use of incentive pay and non-monetary prizes such as career advancement[10] as motivational tools within organizations. In particular, existing work emphasizes how linking rewards to performance can encourage workers to exert unobserved effort. For surveys of theoretical and empirical work making this point, see Milgrom and Roberts (1992), Prendergast (1999), and Gibbons and Roberts (2013). This literature focuses on the incentive effects of rewards and has typically abstracted from matching effects, especially in formal models.[11] For instance, the literature

[8]Formally, the relevant measure of central tendency in our context is the distribution's mode, and we impose a unimodality assumption to ensure that the distribution's center is unambiguously identified.

[9]More precisely, this result holds for a range of prize endowments under which prize incentives have not been exhausted. Additional forces shape the optimal scheme outside this range, which we explore further in Section 8.1.

[10]Career advancement can take the form of promotion within an organizational hierarchy, as in the tournaments literature, or external rewards such as visibility to the labor market (Waldman 1984; Bar-Isaac and Lévy 2022).

[11]One exception is Schöttner and Thiele (2010), who analyze how performance pay to overcome moral hazard impacts the quality of promoted candidates in tournaments.

on tournaments (Lazear and Rosen 1981; Rosen 1986; Green and Stokey 1983; Nalebuff and Stiglitz 1983) interprets prizes flexibly as money or promotion but assumes that the cost of awarding a prize is independent of the recipient.

Our work differs from these studies in two ways. First, in our setting organizations wish to influence (observable) task choices rather than (unobservable) effort. Second, matching is a first-order concern when awarding non-monetary prizes. The combination of these two features generates a novel tradeoff between different types of rewards. In particular, linking prizes to task choices degrades matching and naturally limits their use as rewards, whereas linking prizes to performance *facilitates* matching whenever performance is a signal of match value as well as effort.[12]

Our interest in comparing monetary and non-monetary prizes connects our work to a literature studying the design of status hierarchies (Auriol and Renault 2008; Besley and Ghatak 2008; Moldovanu, Sela, and Shi 2007; Dubey and Geanakoplos 2020). These papers consider how conferral of status can be used alongside, or in place of, money as a motivational tool. Status is an intangible reward which, unlike the tangible prizes we consider, has no intrinsic value to the organization. Its use is instead constrained by an inverse relationship between the value of status to workers and the extent to which it is awarded. Designing status rewards therefore involves very different considerations from distributing tangible prizes. Additionally, status incentives are likely muted in many settings involving autonomous workers, notably on freelancing platforms.

Our paper is also related to a literature studying the design of pricing systems in ride-hailing markets (Bimpikis, Candogan, and Saban 2019; Guda and Subramanian 2019; Buchholz 2022; Cachon, Daniels, and Lobel 2017). These papers highlight how non-uniform pricing can improve the allocation of drivers when market conditions vary across time or space. In all of these papers, drivers are incentivized exclusively through monetary payments. We complement their analysis by considering how alternative non-monetary prizes, such as priority matching with high-value trips (as discussed in footnote 5), can be used in conjunction with money to reduce incentive costs.

[12] Lazear (2004) makes this point theoretically in a model without incentive concerns. Benson, Li, and Shue (2019) empirically examine promotion-for-performance policies for salespeople and find that performance is an important predictor of managerial ability.

2 Model

A continuum of workers of mass 1 complete a mix of two tasks A and B for an organization. Each worker chooses exactly one task to complete, and workers have heterogeneous intrinsic preferences over tasks, summarized by a dollar value ν capturing the strength of a worker's preference for task B over A. Intrinsic values are independently and identically distributed across the population of workers, with $\nu \sim F$. We assume that $\mathbb{R}_+ \subset \text{supp}(F)$ and that F has a continuous, positive density function f on $\mathbb{R}_+$. (We allow for $F(0) > 0$ but require no assumptions on the smoothness of the value distribution for negative values.) Worker preferences are not observed by the organization.

The organization's goal is to ensure that a fraction N of workers choose task A, where $N \geq F(0)$. Task choices are observable, but workers are autonomous—the organization may not assign tasks. The organization can motivate workers to choose task A by paying monetary *bonuses* and allocating non-monetary *prizes* which are valued by workers. Prizes are indivisible and at most one prize can be allocated to a given worker. Workers assign a common dollar value $V > 0$ to earning a prize, which is known to the organization. A worker with intrinsic preference ν who chooses task i and is promised a bonus T and a probability $\sigma \in [0, 1]$ of earning a prize enjoys total payoff[13]

$$U = \nu \cdot \mathbf{1}\{i = B\} + V \cdot \sigma + T.$$

The organization has unlimited financial resources and is additionally endowed with a fixed mass $\beta \in (0, 1)$ of prizes to allocate. It earns a positive worker-specific profit from allocating a prize which varies across workers. We assume that these profits are determined and observed by the organization only after workers choose tasks, and that the organization's distributional preferences are drawn independently of workers' intrinsic preferences or task choices.

We summarize each worker's suitability for a prize by their *match quantile* $q \sim U([0, 1])$. The organization earns profits $\rho(q)$ from allocating a prize to a worker with match quantile q, where ρ is non-negative, continuous, and increasing in q. We will refer to $\rho(q)$ as the *match value* of a worker of match quantile q. It will be convenient to decompose ρ as $\rho(q) = R \cdot \rho_0(q)$,

[13] For concreteness, we normalize payoffs so that each worker earns a payoff of 0 from choosing task A and v from choosing task B. Since we abstract from workers' participation decision, this normalization is innocuous. See Section 9 for further discussion of the participation margin.

where ρ_0 is a baseline match value function and $R > 0$ is a parameter that we allow to vary across environments. If the organization promises a worker with match quantile q a bonus T and probability σ of winning a prize, it generates expected profits

$$\Pi = \rho(q) \cdot \sigma - T$$

from that worker.

3 Incentive schemes

The organization may commit to a scheme for allocating money and prizes as a function of workers' task choices and match quantiles. We assume that workers cannot be charged a fee at the time of task choice or when receiving a prize, and so all transfers must flow from the organization to workers.[14]

All payoffs are additive in allocation decisions, and it is therefore without loss to restrict attention to schemes in which bonuses are not contingent on a worker's match quantile or prize allocation. Additionally, since workers do not observe their match quantile before choosing a task, they care only about the ex ante probability of winning a prize and not the correlation between prizes and match values. As a result, given a pool of prizes promised to workers choosing a given task, the organization optimally prioritizes the best-matched workers within the group for prizes.

An incentive scheme may therefore be summarized by a triplet $\mathcal{C} = (q_A, q_B, T_A, T_B)$, where $q_i \in [0, 1]$ is the *match standard* that must be met to earn a prize after choosing task $i \in \{A, B\}$, and $T_i \geq 0$ is the bonus paid for choosing that task.[15] We will refer to $\Delta q \equiv q_B - q_A$ as the *relative prize priority* implemented by a given incentive scheme, and $\Delta\rho \equiv \rho(q_B) - \rho(q_A)$ as the *absolute prize priority.*

[14] Our assumption that workers are not charged for prizes is consistent with observed practice on freelancing platforms and in promotion decisions. More broadly, a floor on permissible transfers, which we normalize to zero for convenience, captures a commitment to a minimum base wage. In Section 9 we discuss the possibility of designing this wage to control the supply of workers.

[15] Without loss, we assume that match standards and transfers are non-random. Since all players are risk-neutral over transfers, random transfers can be replaced by their means without disturbing payoffs or incentives. Meanwhile, workers are risk-neutral over match standards, while the organization's profits are strictly concave. As a result, replacing match standards with their means preserves incentives and raises profits.

Given a scheme $\mathcal{C}$, the gap in task payoffs $v(\mathcal{C}) \equiv T_A - T_B + V \cdot \Delta q$ is also the highest intrinsic preference ν for which a worker chooses task A. A scheme $\mathcal{C}$ is *feasible* if it allocates no more prizes than are available and achieves the required target:

$$\begin{cases} \beta \geq F(v(\mathcal{C}))(1-q_A) + (1 - F(v(\mathcal{C})))(1-q_B) \\ N = F(v(\mathcal{C})) \end{cases}$$

The organization's payoff from a feasible incentive scheme is

$$\Pi(\mathcal{C}) = N \cdot \left[\int_{q_A}^{1} \rho(q)\, dq - T_A\right] + (1-N) \cdot \left[\int_{q_B}^{1} \rho(q)\, dq - T_B\right].$$

It will be convenient to describe the organization's task-mix goal via a threshold preference target $v \geq 0$ rather than a group-size target N. Given a preference target v, the implied group-size target is $N = F(v)$. Going forward, we will refer to a desired preference target v as simply *the target*. With this convention, the feasibility constraints may be written

$$\begin{cases} \beta \geq F(v)(1-q_A) + (1 - F(v))(1-q_B) \\ v = T_A - T_B + V \cdot \Delta q \end{cases}$$

The following result establishes that, without loss, we may focus on incentive schemes which pay no bonus for choosing task B.

Lemma 1. *Suppose that a feasible incentive scheme* $\mathcal{C} = (q_A, q_B, T_A, T_B)$ *satisfies* $T_B > 0$. *Then there exists another feasible scheme* $\mathcal{C}' = (q'_A, q'_B, T'_A, T'_B)$ *such that* $\Pi(\mathcal{C}') > \Pi(\mathcal{C})$ *and* $T'_B = 0$.

To streamline notation, we will write T without subscript to refer to the bonus paid to workers choosing task A, and we will describe an incentive scheme via a triple $\mathcal{C} = (q_A, q_B, T)$. The feasibility constraints for an incentive scheme implementing a given target v are

$$\begin{cases} \beta \geq F(v)(1-q_A) + (1 - F(v))(1-q_B) \\ v = T + V \cdot \Delta q \end{cases}$$

The following result establishes existence of a unique optimal scheme as well as some basic properties of this scheme.

Lemma 2. *The set of feasible incentive schemes is non-empty, and there exists a unique optimal scheme* $\mathcal{C}^* = (q_A^*, q_B^*, T^*)$, *which satisfies* $T^* \in [0, v)$ *and* $q_B^* > 1 - \beta > q_A^*$ *whenever* $v > 0$.

Notably, an optimal scheme always distorts the allocation of prizes: $\Delta q^* \equiv q_B^* - q_A^* > 0$. This result arises because the first misallocated prize is free to first order. On the other hand, the first dollar of bonuses incurs a first-order cost, and so it may be that $T^* = 0$.

4 Characterizing an optimal scheme

In the absence of a task-mix goal, the organization optimally awards prizes strictly in order of match value, ignoring each worker's task choice. The match standards under this policy are $q_A^{FB} = q_B^{FB} = 1 - \beta$, yielding total profits

$$\Pi^{FB} \equiv F(v) \int_{1-\beta}^{1} \rho(q)\, dq + (1 - F(v)) \int_{1-\beta}^{1} \rho(q)\, dq.$$

Any incentive scheme $\mathcal{C} = (q_A, q_B, T)$ which chooses $(q_A, q_B) \neq (1 - \beta, 1 - \beta)$ distorts the allocation of prizes and reduces the organization's profits. Define

$$\Delta R(q_A, q_B) \equiv (1 - F(v)) \int_{1-\beta}^{q_B} \rho(q)\, dq - F(v) \int_{q_A}^{1-\beta} \rho(q)\, dq$$

to be this allocational loss. The first term represents the match value lost from withholding prizes from workers choosing task B, while the second term represents the value gained from allocating extra prizes to workers choosing task A. The first term is always larger than the second when $(q_A, q_B) \neq (1 - \beta, 1 - \beta)$, resulting in a positive loss.

The problem of designing an optimal incentive scheme can be divided into two steps. First, given a specified bonus T, the organization minimizes the loss $\Delta R(q_A, q_B)$ among all feasible schemes $\mathcal{C}$ paying this bonus. Second, it maximizes total profits over T, using the loss-minimizing scheme found in the first part.

For a given bonus $T \geq 0$, let $\mathbb{C}(T)$ be the set of feasible incentive schemes paying T. Define

$$\Delta R^*(T) \equiv \min_{\mathcal{C} \in \mathbb{C}(T)} \Delta R(q_A, q_B)$$

to be the minimized loss calculated in the first step. Then the profit function maximized in the second step is

$$\Pi^*(T) \equiv \Pi^{FB} - \Delta R^*(T) - F(v) \cdot T.$$

The following lemma establishes that the loss function resulting from the first step is a decreasing, convex function of the bonus.

Lemma 3. *Whenever $v > 0$, there exists a bonus $\underline{T} \in [0, v)$ such that:*

- $\Delta R^*(T) = \infty$ *for all* $T < \underline{T}$,
- $\Delta R^*(T)$ *is finite, decreasing, and strictly convex on* $[\underline{T}, v]$.

The minimized loss $\Delta R^*(T)$ is finite only if there exists a feasible incentive scheme delivering bonus T. When v is relatively small, there exists such a scheme for any bonus $T \geq 0$, in which case $\underline{T} = 0$. However, for large v sufficient incentives cannot be provisioned by prize incentives alone, and a minimal bonus $\underline{T} > 0$ is required, a result formally established in the following lemma:

Lemma 4. *There exists a target $\overline{v}_0 \in (0, V]$ such that $\underline{T} = 0$ when $v \in [0, \overline{v}_0]$ while $\underline{T}$ is positive and increasing in v on $(\overline{v}_0, \infty)$.*

For bonuses in the feasible range $[\underline{T}, v]$, there exist many feasible incentive schemes.[16] These schemes all involve some *reallocation* of prizes from workers choosing task B to task A; some may, in addition, allocate only a fraction of the prize endowment, a feature we refer to as prize *burning*. The unique loss-minimizing incentive scheme is characterized by the property that it minimizes prize burning.

As the specified bonus increases, fewer prizes need be misallocated through reallocation or burning, leading to a downward-sloping loss ΔR^*. Further, the marginal misallocated prize becomes less costly the fewer prizes have already been misallocated, since the gap between the match value of the marginal prize-winners shrinks and the lost match value from a burned prize falls. As a result, ΔR^* is strictly convex. This convexity implies that the second-stage profit function Π^* is strictly concave in T on $[\underline{T}, v]$. The maximum of Π^* can therefore be characterized by the property that 0 is a superderivative of Π^* at the maximum.

This derivation of an optimal scheme utilizes the bonus T as the primary design variable and is useful for proving our main results. A complementary characterization focuses on the optimal magnitude of prize incentives. This approach is more complex to formalize, because prize incentives might be furnished either through reallocation or burning. However, it can be simply described in environments where no prizes are burned, and we will rely on this heuristic when developing intuition for our main results.[17]

[16] There also exist feasible incentive schemes for a range of bonuses above v. However, as established in Lemma 2, the optimal bonus never lies in this range, and so we ignore such bonuses.

[17] See the proof of Theorem 1 for a general derivation which accounts for the possibility of prize-burning.

Intuitively, an optimal scheme can be designed by reallocating prizes across tasks until either the reward gap between tasks equals the incentive target v, or else the marginal cost of further reallocation equals the (constant) marginal cost of bonuses. These marginal costs can be explicitly calculated to obtain a useful first-order condition for the optimal magnitude of prize incentives.

Using money to raise the reward gap $T + V \cdot \Delta q$ by a dollar costs the organization $F(v)$ dollars, since the extra bonus must be paid to every worker choosing task A. The marginal cost of bonuses is therefore $MC^B = F(v)$. Meanwhile, reallocating a mass σ of prizes from group B to group A lowers q_A by σ/N and raises q_B by $\sigma/(1 - F(v))$. Hence

$$\Delta q = \frac{\sigma}{F(v) \cdot (1 - F(v))},$$

meaning that the organization must reallocate a mass $\Delta\sigma = F(v) \cdot (1 - F(v))/V$ of the prize to increase the reward gap by a dollar. The cost of this reallocation is $\Delta\rho(\sigma) \cdot \Delta\sigma$ to first order, where $\Delta\rho(\sigma)$ is the gap between the marginal prize-winner's match value in each group. The marginal cost of prize incentives is therefore $MC^P(\sigma) = \Delta\rho(\sigma) \cdot F(v) \cdot (1 - F(v))/V$.

The marginal cost of prizes is rising in σ, reflecting the convexity of prize incentive costs noted earlier. The optimal mass of reallocated prizes σ^*, and therefore the optimal prize priority $\Delta\rho^* = \Delta\rho(\sigma^*)$, is then uniquely determined by the optimality condition

$$\frac{MC^P(\sigma^*)}{MC^B} = \Delta\rho^* \cdot (1 - N) \leq 1, \tag{FOC}$$

with equality if the optimal bonus is non-negative.

5 Role of the target

In this section we present our main results, which describe how an optimal scheme is shaped by the desired magnitude of incentives. We find that raising the incentive target increases the optimal magnitude of prize incentives but has an ambiguous effect on the size of the optimal bonus. Specifically, the prize priority increases with the target, while the bonus tends to first increase and later decrease as the target increases. The increasing prize priority and eventually declining bonus reflect a decline in the relative cost of prize incentives as workers move between tasks, a force we designate the *group-size effect*. Meanwhile, the initially increasing bonus reflects convexity in the cost of reallocating successive prizes, a force we label the *gap-size effect*.

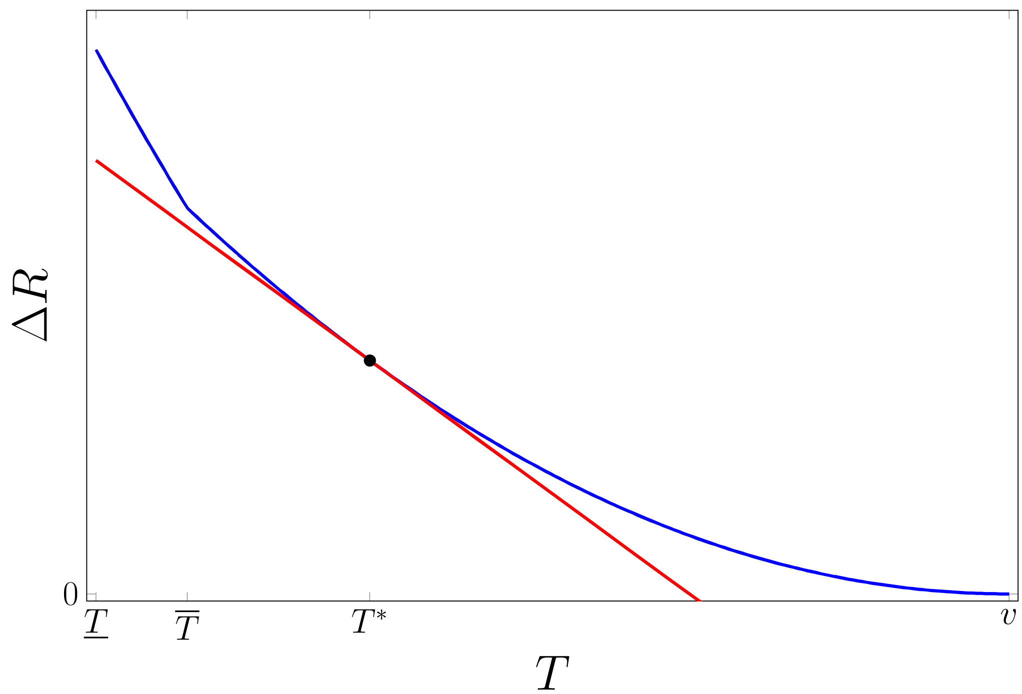

Figure 1: The feasible frontier ΔR^* (in blue) lies tangent to an isoprofit line (in red) at the optimal bonus T^*.

In section 5.2 we develop these results in a baseline setting in which match values are uniformly distributed across the population of workers. We then generalize the results to arbitrary match distributions in section 5.3.

5.1 Preliminaries

The optimal tradeoff between bonus and prize incentives depends on whether prizes are available to be reallocated (or burned) on the margin. The following lemma connects this availability to the size of the incentive target.

Lemma 5. *There exists a target $\overline{v} \in [\overline{v}_0, \infty)$ such that:*

- *If $v < \overline{v}$, then either $\underline{T} = 0$ or $T^* > \underline{T}$.*
- *If $v \geq \overline{v}$, then $\mathcal{C}^* = (\max\{0, 1 - \beta/F(v)\}, 1, \underline{T})$.*

When $v < \overline{v}$, any lower bound $\underline{T} > 0$ on feasible bonuses imposed by limited prize incentives is non-binding. As a result, both tools are available on the margin under an optimal contract, and the optimal tradeoff between them is determined by balancing their marginal costs. Conversely, when $v \geq \bar{v}$ prize incentives are exhausted and the marginal cost of bonuses rises above the marginal cost of the last unit of prize incentives. Our main results characterize how the optimal scheme varies with v when prize incentives have not been exhausted.

5.2 Uniform match values

We now present results under the assumption that match values are uniformly distributed across the population of workers. Since the match-value function ρ is the inverse distribution function of match values, uniform match values correspond to a linear match-value function:

Assumption 1. $\rho(q) = R \cdot (q + b)$ *for some* $b \geq 0$.

We first characterize how optimal prize incentives vary with the target. We establish that both the absolute and relative prize priorities $\Delta\rho^* = \rho(q_B^*) - \rho(q_A^*)$ and $\Delta q^* = q_B^* - q_A^*$ are increasing in v, as is the match standard q_B^* for workers choosing task B.

Proposition 1. *Suppose that Assumption 1 holds. Optimal prize incentives vary with* v *as follows:*

- Δq^* *and* $\Delta\rho^*$ *are increasing in* v *on* $[0, \bar{v}]$.
- q_B^* *is increasing in* v *on* $[0, \bar{v}]$.

This monotonicity stems from a *group-size effect* affecting incentive costs as workers shift between tasks. Recall the optimality condition (FOC), which balanced the marginal costs of monetary and prize incentives and set the optimal magnitude of prize incentives. When the optimal bonus is positive, this condition may be written

$$\frac{\Delta\rho^*}{V} \cdot \underbrace{(1 - N)}_{\text{group-size effect}} = 1 \qquad \text{(FOC')}$$

where we have set $F(v) = N$ to emphasize the fact that the term $1 - F(v)$ captures the number of workers choosing task B. Since the right-hand side of (FOC') is declining in the group size N, the absolute prize priority $\Delta\rho^*$ equalizing the marginal cost of the two tools is therefore increasing in N, or equivalently in the target $v = F^{-1}(N)$. We refer to this phenomenon as the group-size effect.

Intuitively, monetary incentives operate by raising the payoff of choosing task A, while prize incentives operate through a combination of an increased payoff to task A and a decreased payoff to task B. As more workers choose task A, more money and/or prizes must be allocated to the group to raise each worker's payoff by a given amount. This force raises the marginal cost of manipulating that group's payoff symmetrically for each tool.[18] However, a corresponding reduction in the number of workers choosing task B reduces the cost of raising their match standard (i.e., withholding prizes from well-matched workers), an effect which benefits only prize incentives. The group-size effect therefore lowers the marginal cost of prize incentives relative to bonuses, encouraging the organization to provision more incentives through prizes.

Under uniform match values, the absolute and relative prize priorities are related via the identity $\Delta\rho = R \cdot \Delta q$. Thus a rise in $\Delta\rho^*$ induces a corresponding rise in Δq^*. The behavior of q_B^* is closely linked to the behavior of Δq^*. The overall number of prizes allocated must satisfy the resource constraint, under which the number of available prizes is a weighted average of the fraction of prize-winners in each group, with weights equal to the group sizes.[19] Mechanically, as the gap between the fraction of prize-winners in each group rises while group sizes are held fixed, q_A^* must fall while q_B^* rises. Hence a rising Δq^* implies a rising q_B^*. A reallocation of workers toward task A amplifies this effect by increasing the weight on the larger prize allocation in the resource constraint, requiring a compensating rise in q_B^* for a given Δq^* as v rises.

This same logic implies that the behavior of q_A^* as v rises is ambiguous. On the one hand, a rising Δq^* implies a smaller q_A^*, holding group sizes fixed. But on the other hand a shift in the group sizes requires a larger q_A^* to balance the resource constraint, holding Δq^* fixed. These two forces compete, and either may dominate. This possibility is formally demonstrated below in Theorem 2.

We next characterize how optimal monetary incentives vary with the target v. For this

[18]If the marginal prize is burned rather than reallocated, then this force affects only monetary incentives, further strengthening the group-size effect.

[19]For simplicity, this exposition focuses on environments in which no prizes are burned. Otherwise, incentive power from reallocating prizes must be exhausted, in which case $q_A^* = 0$ and $q_B^* = \Delta q^*$. A similar logic therefore links the behavior of q_B^* and Δq^*.

result we impose a monotonicity condition on the *virtual value* function $\varphi(v) \equiv v - (1 - F(v))/f(v)$. Monotonicity of the virtual value function is a familiar assumption in the mechanism design literature and is satisfied by a variety of common distributional families.

Assumption 2. *φ is increasing for $v \geq 0$ and $\varphi(\infty) > 0$.*

We establish that T^* is quasiconcave in v under Assumption 2. Quasiconcavity implies that T^* is either nondecreasing or hump-shaped,[20] and we further establish that the latter behavior arises under some model parameterizations.

Proposition 2. *Suppose that Assumptions 1-2 hold. Then optimal bonus incentives vary with v as follows:*

- *T^* is quasiconcave in v on $[0, \bar{v}]$.*
- *$T^* = 0$ for $v > 0$ sufficiently small.*
- *If R is sufficiently large, then $T^* > 0$ for some $v \in [0, \bar{v}]$.*
- *If V is sufficiently large, there exists a non-degenerate interval of parameters R for which T^* is non-monotone in v on $[0, \bar{v}]$.*

The behavior of T^* as a function of v is determined by the interplay of two forces: the group-size effect just discussed, which tends to depress the optimal bonus as the target rises; and a *gap-size* effect, which tends to increase it. To understand these effects, decompose the total required reward gap v into

$$v = T + V \cdot \Delta q.$$

As v rises, the reward gap must grow through a combination of a larger bonus and/or increased reallocation of prizes. Consider the exercise of increasing this gap without allowing employees to actually switch tasks. In this benchmark, the group size N remains fixed, in which case $\Delta q^* = \Delta \rho^*/R$ is determined by (FOC') independently of the reward gap. Essentially, prize incentives are naturally limited by the convexity of prize reallocation costs, and all marginal incentives are optimally provisioned through bonuses. T^* therefore grows with the reward gap, a phenomenon we call the gap-size effect.

[20] Because the optimal bonus must be non-negative and vanishes at $v = 0$, the remaining possibility that T^* is decreasing is ruled out.

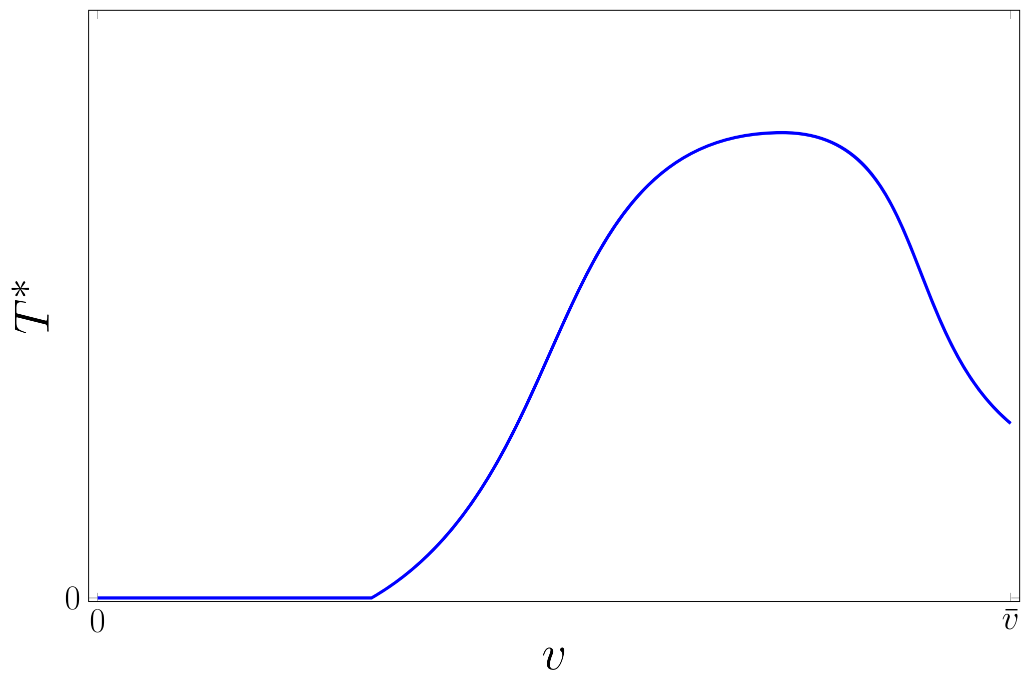

Figure 2: An optimal bonus T^* which is both quasiconcave and nonmonotone in v.

The group-size effect and gap-size effect oppose one another, leading to an ambiguous total effect of increased incentives on the optimal bonus. Proposition 2 establishes that the gap-size effect dominates when v is small, while the group-size effect dominates when v is large. This behavior can be understood by expanding the optimality condition (FOC') which pins down the optimal bonus whenever $T* > 0$. Using the identities $\Delta\rho = R \cdot \Delta q$ and $v = T + V \cdot \Delta q$, this optimality condition may be written in terms of T^* as

$$1 = \frac{R}{V^2} \cdot \underbrace{(v - T^*)}_{\text{gap-size effect}} \cdot \underbrace{(1 - F(v))}_{\text{group-size effect}} .$$

Under Assumption 2, the function $(v - T)(1 - F(v))$ is single-peaked in v for any choice of T. This behavior ensures that the optimal bonus is quasiconcave in v. In particular, the marginal cost of prize incentives (as captured by the gap-size effect) is sufficiently modest for small targets that the optimal bonus is 0 when v is close to zero. Additionally, T^* is single-peaked provided that the first-order condition stated above holds with equality over a sufficiently wide range of v. This outcome is ensured when V is sufficiently large (implying that $\bar{v}$ is large) and R is neither too small (which would yield $T^* = 0$ everywhere) nor too large (which would entail very small prize incentives everywhere and global dominance of the gap-size effect). Figure 2 provides an illustration of a non-monotone, quasiconcave T^*.

5.3 General match values

We continue our analysis by relaxing the uniform match value assumption. Instead, we impose only the mild condition that the (inverse) match distribution function be smooth at two key quantiles.

Assumption 3. *ρ is continuously differentiable near $q = 1 - \beta$ and $q = 1$.*

We also dispense with the monotonicity condition previously imposed on the virtual value function. We now require only that the hazard rate of the value distribution not vanish too rapidly.[21]

Assumption 4. $\lim_{v\to\infty} v\frac{f(v)}{1-F(v)} = \infty$.

We first generalize the results of Proposition 1 describing how optimal prize incentives change with the target. We additionally provide a formal statement of the possibility that q_A^* is non-monotone in the target.

Theorem 1. *Optimal prize incentives vary with v as follows:*

- *$\Delta\rho^*$ is increasing in v on $[0, \bar{v}]$.*
- *q_B^* is increasing in v on $[0, \bar{v}]$.*
- *If Assumption 4 holds and V is sufficiently large, then q_A^* is non-monotone in v on $[0, \bar{v}]$.*

Unlike in the uniform match value case, we can no longer guarantee that the optimal *relative* prize priority Δq^* varies monotonically with v. However, it remains true that the *absolute* prize priority $\Delta\rho^*$ varies monotonically, for the same reasons as in the uniform match value case. The logic for monotonicity of q_B^* is also similar: Ignoring the group-size effect, increasing v would lead to an increase in Δq^* and therefore a mechanical increase in q_B^*. While the group-size effect has an ambiguous effect on Δq^* for general match distributions, it continues to unambiguously amplify the effect of v on q_B^*. Additionally, as noted

[21] The boundary condition $\varphi(\infty) > 0$ imposed in Assumption 2 implies that $\liminf_{v\to\infty} v\frac{f(v)}{1-F(v)} \geq 1$. Assumption 4 slightly strengthens this requirement to accommodate non-uniform match distributions.

in the discussion following Proposition 1, the effect of a changing target on q_A^* is in general ambiguous regardless of the match distribution.

We next prove an analogue of Proposition 2 describing how the optimal bonus changes with the target. Under non-uniform match values, quasiconcavity of T^* is no longer assured. Nonetheless, the general phenomenon of non-monotonicity continues to hold for appropriate choices of V and R. (Recall that we have assumed that $\rho(q) = R \cdot \rho_0(q)$, where ρ_0 is a fixed reference match-value function and R is a scale factor. As a result, when we allow the parameter R to vary, we hold fixed the *shape* of the match distribution and vary only its *scale*.)

Theorem 2. *optimal bonus incentives vary with v as follows:*

- $T^* = 0$ *for sufficiently small* $v > 0$.
- *If R is sufficiently large, then $T^* > 0$ for some $v \in [0, \overline{v}]$.*
- *If Assumptions 3-4 hold and V is sufficiently large, there exists a non-degenerate interval of parameters R for which T^* is non-monotone in v on $[0, \overline{v}]$.*

This result relies on the same basic interplay between the group- and gap-size effects that drove Proposition 2, although this interplay is more complex when ρ is nonlinear. In general, the optimality condition (FOC') characterizing a nonzero T^* can be written

$$1 = \left[\rho\left(1 - \beta + \underbrace{\frac{F(v)}{V}}_{\text{group}}\underbrace{(v - T^*)}_{\text{gap}}\right) - \rho\left(1 - \beta - \underbrace{\frac{1 - F(v)}{V}}_{\text{group}}\underbrace{(v - T^*)}_{\text{gap}}\right)\right] \cdot \underbrace{(1 - F(v))}_{\text{group}}.$$

The appearance of v in the terms $v - T^*$ reflects the gap-size effect, which tends to increase the optimal usage of bonuses as v rises. Meanwhile, terms involving $F(v)$ reflect changing group sizes. As in the uniform match value case, the final term $1 - F(v)$ captures increased incentive power in group B as the size of that group shrinks.

The additional appearances of $F(v)$ reflect the fact that as group sizes change, match standards must shift systematically across both groups to maintain a fixed relative prize priority. When match values are uniform, level shifts of the match standards q_A and q_B do not affect the marginal cost of prize incentives (which is controlled by $\Delta\rho$). However, for more general distributions this shift can either increase or decrease the $\Delta\rho$ corresponding to a given Δq. As a result, the direction of the group-size effect is ambiguous, and the

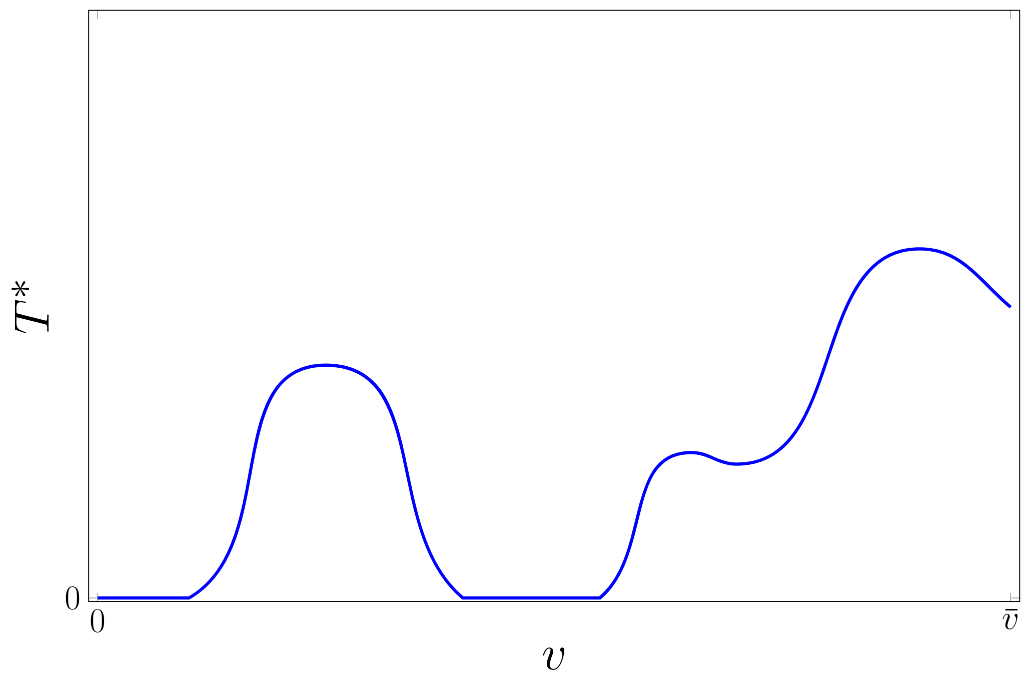

Figure 3: An optimal bonus T^* which is nonmonotone but not quasiconcave in v.

interplay between the gap- and group-size effects is complex. In particular, quasiconcavity of T^* in v can no longer be ensured. (Figure 3 illustrates the possibilities in this more general environment.) Nonetheless, the proof of the theorem establishes that the gap-size effect continues to dominate for small v while the group-size effect both dominates and tends to reduce the relative cost of prizes for large v.

Given the weaker guarantees on the shape of T^* promised by Theorem 2, we do not require the monotone virtual cost assumption required to prove Proposition 2. However, we do still need a condition on the limiting behavior of the value distribution F in order to ensure that the group-size effect dominates for large v. Intuitively, the term $(v - T)(1 - F(v))$ appearing in the first-order condition for the uniform match value case is decreasing for large v provided that the value hazard rate $f(v)/(1 - F(v))$ vanishes no faster than $1/v$ asymptotically. Assumption 4 imposes this asymptotic condition, which the proof of the Proposition shows is sufficient to ensure dominance of the group-size effect for large v under general match distributions.

6 Role of the resource constraint

We now complement our main results by studying how an optimal incentive scheme depends on the number of prizes available to the organization. Our primary finding is that, when match values are uniformly distributed, β has no impact on the optimal magnitude of monetary or prize incentives. We establish this result in Section 6.2. We then extend our analysis to general single-peaked match distributions in 6.3, where we show that variation in β induces a U-shaped response of both bonuses and the (absolute) prize priority.

6.1 Preliminaries

Our main results focus on the structure of the optimal scheme in environments where prize incentives have not been exhausted for either group. As the following lemma establishes, in such environments prizes are neither too scarce nor too plentiful.[22]

Lemma 6. *Suppose that $v \in (0, V)$. Then there exist prize endowments $\underline{\beta}$ and $\overline{\beta}$ satisfying $0 < \underline{\beta} < \overline{\beta} < 1$ such that:*

- $q_B^* < 1$ *if and only if* $\beta > \underline{\beta}$.
- $q_A^* > 0$ *if and only if* $\beta < \overline{\beta}$.

We will show that as β varies within $[\underline{\beta}, \overline{\beta}]$, all variation in the optimal incentive scheme is driven by non-uniformity of the match distribution. Outside this range, the optimal incentive scheme may additionally vary with β in order to avoid burning prizes (if $\beta > \overline{\beta}$) or because prize incentives are exhausted (if $\beta < \underline{\beta}$). These possibilities arise only when $q_A^* = 0$ or $q_B^* = 1$, i.e., when prizes are awarded unselectively to workers choosing task A or are withheld entirely from workers choosing task B. We defer further discussion of the optimal incentive scheme in these regimes until Section 8.1.

6.2 Uniform match values

We now establish that, when match values are uniformly distributed, the number of available prizes has *no* impact on the optimal magnitude of monetary or prize incentives within the range $[\underline{\beta}, \overline{\beta}]$.

[22] The upper bound $v < V$ on the target ensures that $\underline{\beta} < \overline{\beta}$. For larger targets, this region is degenerate or empty when R is sufficiently small.

Proposition 3. *Suppose that $v \in (0, V)$ and Assumption 1 holds. Then T^*, Δq^*, and $\Delta\rho^*$ are independent of β on $[\underline{\beta}, \overline{\beta}]$.*

In general, the prize endowment β enters the tradeoff between prizes and bonuses only insofar as it affects how a given relative priority Δq, which controls the strength of prize incentives, translates into an absolute priority $\Delta\rho = \rho(q_B) - \rho(q_A)$, which controls the marginal cost of prize incentives. In general, given a desired relative priority Δq, a change in the prize endowment shifts both q_A and q_B uniformly to satisfy the resource constraint. Under uniform match values, $\Delta\rho = R \cdot \Delta q$, implying that uniform shifts to the match standards do not affect the relationship between Δq and $\Delta\rho$. As a result, β drops out of the tradeoff between bonuses and prizes completely.

Proposition 3 establishes the perhaps surprising result that β plays no role in the optimal mix of prize and monetary incentives. However, β does enter the form of an optimal scheme via the individual match standards for each group. Specifically, when more prizes are available, the match standards shift downward to reflect the larger number of workers who earn prizes.

6.3 General match values

We now study the role of the resource constraint when match values are not uniformly distributed across the population of workers. As noted in the discussion following Proposition 3, the irrelevance of β for optimal incentives in that result relied crucially on a direct proportionality between the relative and absolute prize priorities. When ρ is nonlinear, $\Delta\rho$ and Δq are no longer proportional, opening a new channel for β to affect the shape of an optimal scheme.

Our main result in this subsection characterizes the effect of this channel under a mild regularity condition on the match distribution. We require that the match value density function be single-peaked, which corresonds to an inverse-sigmoidal shape for ρ. To streamline exposition, we also enforce a symmetric-density assumption at the edges of the distribution.

Definition 1. *A function $f : [0, 1] \to \mathbb{R}$ is* inverse-sigmoidal *if it is continuously differentiable and increasing and f' is single-troughed.*

Assumption 5. *ρ is inverse-sigmoidal and $\rho'(0) = \rho'(1)$.*

When the match distribution has a single-peaked density, we establish that absolute incentives, as measured by both T^* and $\Delta\rho^*$, are weakest for intermediate β and grow stronger when β is small or large.

Theorem 3. *Suppose that $v \in (0, V)$ and Assumption 5 holds. Then there exist prize endowments $\beta_L \in [\underline{\beta}, \overline{\beta})$ and $\beta_H \in (\underline{\beta}, \overline{\beta}]$, satisfying $\beta_L \leq \beta_H$, such that:*

- *T^* is decreasing, Δq^* is increasing, and $\Delta\rho^*$ is constant in β on $[\underline{\beta}, \beta_L]$,*
- *$T^* = 0$, Δq^* is constant, and $\Delta\rho^*$ is single-troughed in β on (β_L, β_H),*
- *T^* is increasing, Δq^* is decreasing, and $\Delta\rho^*$ is constant in β on $[\beta_H, \overline{\beta}]$.*

Further, there exists a non-degenerate interval of parameters R for which $\underline{\beta} < \beta_L < \beta_H < \overline{\beta}$.

An optimal scheme is designed by reallocating prizes across groups until either the target has been achieved, or else the marginal cost of further reallocation equals the (constant) marginal cost of bonuses. In the first case, no bonuses are paid; otherwise, bonuses are used to provision all residual incentives. The marginal cost of reallocation is directly proportional to the (absolute) prize priority $\Delta\rho$, since this gap records the cost of reallocating one more prize. Meanwhile, the contribution of prize incentives to the target is $V \cdot \Delta q$. The optimal choice of $\Delta\rho$ (and hence also T) therefore depends on the linkage between $\Delta\rho$ and Δq.

When the match distribution is single-peaked, there are many workers with similar match values in the middle of the match distribution. As a result, in this portion of the distribution the organization can generate significant prize incentives through reallocation without distorting total match value very much. Conversely, in the tails of the distribution workers have relatively differentiated match values, and so redistribution incurs a large allocative cost here. The prize endowment β acts as a shifter on the match standards q_A and q_B and determines where in the distribution they lie. When β is intermediate, these standards correspond to match values in the middle of the distribution, while for extreme values of β they correspond to match values in the tails.

If R is sufficiently small, then for intermediate β the incentive cost of reallocating prizes is small enough that bonuses are not used. In that case, $\Delta q^* = v/V$ is independent of β, inducing variation in the absolute prize priority $\Delta\rho^*$ to maintain a fixed relative priority Δq^*. Given the linkage between Δq^* and $\Delta\rho^*$ described above, $\Delta\rho^*$ is therefore single-dipped in β.

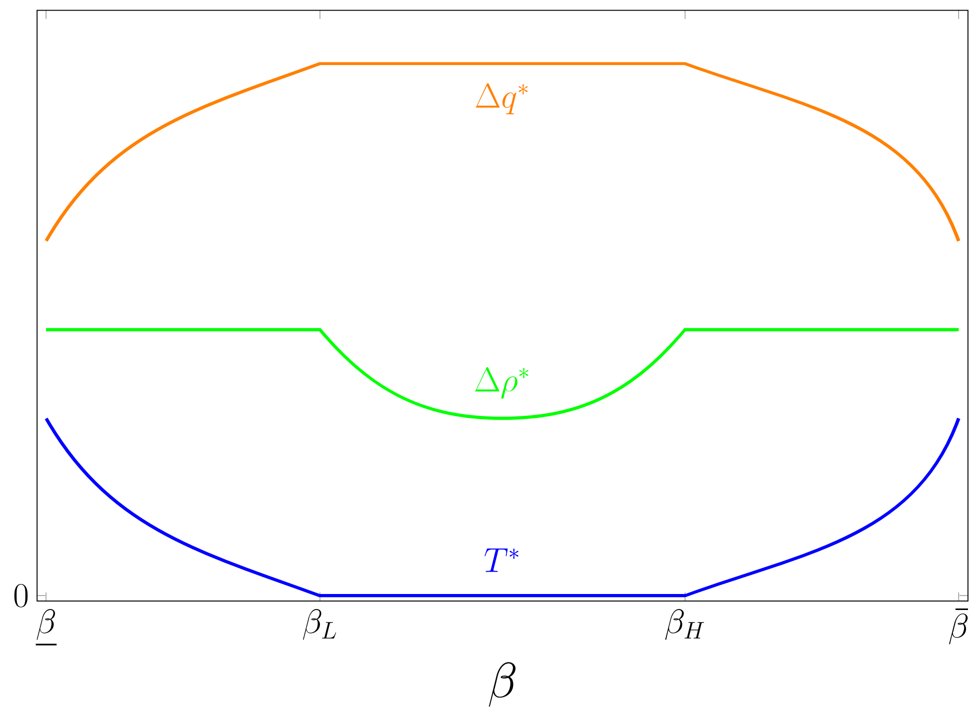

Figure 4: Optimal bonus T^*, relative prize priority Δq^*, and absolute prize priority $\Delta \rho^*$ as functions of the prize endowment β.

Meanwhile, for extreme β, achieving the target by reallocating prizes incurs a large incentive cost; if R is sufficiently large, then bonuses are optimally used. In this regime, $\Delta \rho^*$ remains fixed at the level equalizing the marginal costs of the two tools. Since in this case $\Delta \rho^*$ does not respond to variation in β, the linkage between Δq and $\Delta \rho$ described above implies that Δq^* decreases as β becomes more extreme, with a corresponding increase in T^*.

7 Testable predictions

Our results offer testable predictions about observable aspects of incentive schemes in real markets. As one example, consider Uber's problem of incentivizing drivers to pick up riders in underserved locations. Theorem 2 predicts that monetary incentives may be used most intensively when misallocation of drivers is moderate, such as during a daily rush hour. Conversely, when drivers are overwhelmingly needed in a particular location, such as during exceptionally heavy demand driven by a concert or sporting event, Uber might prefer not to pay bonuses to every driver serving the event. Instead, Theorem 1 predicts that it might boost supply mostly by withholding prizes such as priority ride-matching from drivers shunning the event. Predictions such as this one could be evaluated given sufficiently rich data on worker choices and rewards.

As another example, consider Google's problem of incentivizing software engineers to spend time on side projects that create significant value for the company. Theorem 3 predicts that the optimal bonus and prize priority are roughly U-shaped in the number of available promotions, with both incentives weakening as the marginal promoted employee tends toward mediocrity. Supposing that Google is a relatively hierarchical company with few available promotions, a drop in the company's growth rate should move the marginal promoted employee deeper into the right tail of the talent distribution. In that case, Theorem 3 predicts that the company may reward high-value side projects with both more money and a larger weight in promotion decisions when the company is growing slowly than when it is growing quickly.

The data requirements for evaluating our model's predictions are reasonably modest in many applications. To bring our model to the data, an analyst would need to observe the distribution of worker task choices; any incentive payments made to them; the allocation of non-monetary prizes; and a measure of worker match values for these prizes. In the platform markets discussed above, these variables may either be directly observed by the platform operator or could be plausibly proxied with variables which are observed.

For instance, consider our ride-hailing example above. Task choices might correspond to the number of times a driver has picked up fares in areas with many unmatched riders, or the fraction of days that the driver has worked during rush hour; match value could be proxied by metrics like the time to pickup; incentive payments could be estimated by payments from surge pricing and other driver incentive programs; and the strength of prize incentives could be estimated by regressing proxies for prioritization—such as frequency of airport pickups, rating of matched riders, and wait time between matches—on past task choices, controlling for match value and other pertinent factors such as driver rating.

As an alternative route to evaluating our model predictions, an analyst could survey management practices using the techniques developed in Bloom and Reenen (2007). While those techniques were developed to evaluate the connection between performance incentives and worker productivity, they could be fruitfully deployed to understand how autonomous workers are managed. In particular, managers could be surveyed on the contexts in which tasks go underserved; the tools used to encourage workers to complete these tasks; and in particular the extent to which managers view non-cash prizes as effective rewards for completing particular tasks. Managers could also be surveyed on the extent to which task allocation concerns guide decisions about the quantity and value of prizes or recruitment of new workers, topics we discuss further in Section 9.

8 The limits of incentives

Our main results focus on environments in which money and prize incentives are plentiful. In particular, we place no limits on the size of the bonus pool, and we place bounds on v and β ensuring that prize incentives are not exhausted in an optimal incentive scheme. These assumptions allow us to cleanly characterize the tradeoff between the two tools when both are available. However, in some organizations one or the other tool might be in short supply. We now discuss how an optimal incentive scheme is affected by such constraints.

8.1 Scarce prize incentives

Scarce prize incentives can be formally captured in our model by either a large v (relative to the value V of a prize) or a prize endowment β which is close to 0 or 1. In such environments, prize incentives may be exhausted:

Definition 2. *A feasible incentive scheme* $\mathcal{C} = (q_A, q_B, T)$ exhausts prize incentives *if* $q_B - q_A = \min\{1, \beta/F(v)\}$.

An incentive scheme exhausts prize incentives by maximizing $\Delta q = q_B - q_A$ among all feasible schemes. If $\beta \leq F(v)$, then this maximum is achieved by allocating all prize to workers choosing task A and none to workers choosing task B. Otherwise, it is achieved by guaranteeing all workers a prize in return for choosing task A and burning all excess prizes to ensure that no workers choosing task B receive one. In either case, $q_B = 1$ under any scheme which exhausts prize incentives. The following lemma demonstrates that $q_B^* = 1$ is in fact a necessary and sufficient condition for exhausting prize incentives under an optimal scheme:[23]

[23] There exist additional suboptimal schemes which set $q_B = 1$ but do not exhaust prize incentives. These schemes unnecessarily burn prize by setting $q_A > \max\{1 - \beta/F(v), 0\}$.

Lemma 7. *$\mathcal{C}^*$ exhausts prize incentives if and only if $q_B^* = 1$.*

When the target v grows large, Lemma 5 demonstrated that eventually $q_B^* = 1$ and therefore an optimal incentive scheme exhausts prize incentives. Meanwhile, Lemma 6 established that prize incentives are not exhausted under an optimal scheme when $v < V$ and $\beta > \underline{\beta}$. On the other hand, since $q_B^* = 1$ for $\beta \leq \underline{\beta}$, prize incentives are optimally exhausted for small β.[24]

For small targets ($v < V$), it is not necessary to exhaust prize incentives to achieve the target when β is close to 1. Essentially, in this limit prize incentives do not become scarce because prizes can be burned. When $v > V$, however, the situation changes. In this regime, burning prizes is not sufficient to provision incentives. It is then possible for prize incentives to be exhausted for large β, provided that R is small enough that the organization prefers to burn all excess prizes before resorting to bonuses:

Proposition 4. *Suppose that $v > V$. If β is sufficiently close to 1, then $\mathcal{C}^*$ exhausts prize incentives when R is sufficiently small.*

8.2 Scarce monetary incentives

Scarce monetary incentives can be captured by augmenting our model with a financial constraint reflecting total money available to the organization to pay bonuses. Formally, a financial constraint requires that $F(v) \cdot T \leq M$, where M captures available financial resources. Since this constraint tightens as v grows, the set of feasible incentive schemes is non-empty only if the target is not too large, a result formalized by the following result.

Lemma 8. *Given available financial resources $M > 0$, there exists a target $v^* \in (\overline{v}_0, \infty)$ such that the set of feasible incentive schemes is non-empty if and only if $v \leq v^*$. The target v^* is increasing in M and satisfies $v^* = \overline{v}_0$ when $M = 0$ and $v^* \to \infty$ as $M \to \infty$.*

The set of targets $v \in [0, v^*]$ for which the financial constraint binds may have a complex structure. However, the constraint can be shown to never bind when M is sufficiently large:

Lemma 9. *If M is sufficiently large, then the optimal bonus without a financial constraint is no larger than $M/F(v)$ for all $v \in [0, v^*]$.*

[24]This result continues to hold when $v \geq V$.

Further, no matter the size of M, the financial constraint does not bind for sufficiently small targets. Indeed, Lemma 2 established the bound $T^* < v$ on the (unconstrained) optimal bonus. Hence whenever v is small enough that $vF(v) \leq M$, the constraint is guaranteed to lie slack. Whether the constraint binds for larger targets depends on the cost of misallocating prizes, as measured by R. The following result establishes that the financial constraint never binds for feasible targets if R is small, while conversely the constraint must bind for v sufficiently close to v^* provided that R is sufficiently large:

Proposition 5. *If R is sufficiently small, then the optimal bonus without a financial constraint is no larger than $M/F(v)$ for all $v \in [0, v^*]$.*

If R is sufficiently large, then for v sufficiently close to v^*, the optimal bonus under a financial constraint equals $M/F(v)$ and is smaller than the optimal bonus with no constraint.

Whenever the financial constraint does bind, strict concavity of $\Pi^*(T)$ in T implies that $T^* = M/F(v)$. The optimal bonus is therefore mechanically decreasing in the target when the constraint binds. In that case, the targeting constraint implies that the optimal relative prize priority $\Delta q^* = (v - T^*)/V$ is increasing in the target, and it can be shown that whenever Δq^* increases so does the optimal absolute priority $\Delta\rho^*$.[25] A financial constraint therefore exerts a force complementary to the group-size effect which simultaneously boosts the optimal strength of prizes incentives and diminishes the strength of monetary incentives as v rises.

9 Discussion and conclusion

In this paper we identify a fundamental economic tradeoff between monetary and non-monetary incentives for autonomous workers. We formally illustrate this tradeoff in a stylized setting where prizes and workers are in fixed supply, prizes have a fixed value, and the incentive target is exogenous. Generalizing these model elements would shed light on additional dimensions of the incentive design problem.

A central element of our model is a fixed endowment of prizes which are profitable for the organization to allocate. This endowment may arise naturally from high-level organizational design decisions. For instance, Uber may determine how much to prioritize drivers with high

[25]This result is immediate under uniform match values. In the proof of Theorem 1, we show that under general match distributions, a rising Δq^* continues to imply a rising $\Delta\rho^*$. See Lemma A.2.

ratings and low cancellation rates in order to maximize rider satisfaction. Similarly, Airbnb may choose the number of Superhost badges to award in order to maximize their impact on guest booking decisions. And Google may identify a set of profitable projects to assign to engineering teams. In all of these cases, the number of workers who qualify these prizes determines a natural prize endowment.

Organizations may benefit from boosting the supply of prizes beyond this natural level to strengthen worker incentives.[26] Of course, this tactic is costly—too much priority matching for drivers leads to long wait times for riders, an abundance of Superhosts dilutes the signaling value of the badge to guests, and excess engineering teams waste resources on unprofitable projects. Our model could be used to weigh the incentive benefits of generating extra prizes against the costs of these activities in order to determine the optimal supply of prizes.

A closely related possibility is that organizations might be able to influence the value V of individual prizes. In some cases, this value derives directly from the scarcity of the prize—for instance, Superhost badges may be valued by hosts to the extent that they distinguish their bearers from the broader population. But in other cases, the prize itself may be subject to design—for instance, as in Bar-Isaac and Lévy (2022), where a job can be made more attractive to workers by increasing its visibility to an external labor market. Our model could similarly be used to shed light on the tradeoffs involved in this design decision.

Organizations may also be able to achieve task goals not just by moving workers between tasks, but also by attracting new workers. Freelancing platforms, for instance, sometimes offer sign-up bonuses or raise base wages to attract additional workers. Our model could be enhanced with an extensive margin of participation to study how these tools should be used in conjunction with rewards for task-switching to meet overall task goals.

Finally, organizations may have some flexibility to adjust their task allocation targets in order to economize on incentive costs. Our model provides a cost function for achieving particular goals, which could be coupled with an organizational production function to optimize over task allocations. Enhancing the model in this way would permit a fuller assessment of the costs of autonomy, which could then be compared against alternative institutional arrangements that provide more top-down direction to workers.

[26] See Ke, Li, and Powell (2018) for a related example of an organizational setting in which high-level positions are optimally oversupplied in order to provide prizes for workers in low-level positions.

References

Auriol, Emmanuelle and Régis Renault (2008). "Status and incentives". *The RAND Journal of Economics* 39 (1), pp. 305–326.

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

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

Besley, Timothy and Maitreesh Ghatak (2008). "Status Incentives". *The American Economic Review* 82 (2). Proceedings of the One Hundred Twentieth Annual Meeting of the American Economic Association (May, 2008), pp. 206–211.

Bimpikis, Kostas, Ozan Candogan, and Daniela Saban (2019). "Spatial Pricing in Ride-Sharing Networks". *Operations Research* 67 (3), pp. 744–769.

Bloom, Nicholas and John Van Reenen (2007). "Measuring and Explaining Management Practices Across Firms and Countries". *The Quarterly Journal of Economics* 122 (4), pp. 1351–1408.

Buchholz, Nicholas (2022). "Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry". *The Review of Economic Studies* 89 (2), pp. 556–591.

Burton, Richard M., Dorthe Døjbak Håkonsson, Jackson Nickerson, Phanish Puranam, Maciej Workiewicz, and Todd Zenger (2017). "GitHub: exploring the space between boss-less and hierarchical forms of organizing". *Journal of Organization Design* 6, p. 10.

Cachon, Gérard P., Kaitlin M. Daniels, and Ruben Lobel (2017). "The Role of Surge Pricing on a Service Platform with Self-Scheduling Capacity". *Manufacturing & Service Operations Management* 19 (3), pp. 368–384.

Dubey, Pradeep and John Geanakoplos (2020). "Money and Status in a Meritocracy". Unpublished.

Gibbons, Robert and John Roberts (2013). "2. Economic Theories of Incentives in Organizations". In: *The Handbook of Organizational Economics*. Ed. by Robert Gibbons and John Roberts. Princeton: Princeton University Press, pp. 56–99.

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

Grønning, Terje (2016). "Working Without a Boss: Lattice Organization With Direct Person-to-Person Communication at W. L. Gore & Associates, Inc." In: *Sage Business Cases*. London: Sage Publications, Inc.

Guda, Harish and Upender Subramanian (2019). "Your Uber Is Arriving: Managing On-Demand Workers Through Surge Pricing, Forecast Communication, and Worker Incentives". *Management Science* 65 (5), pp. 1995–2014.

Hill, Linda A. and Jennifer M. Suesse (Jan. 2003). "Sun Hydraulics: Leading in Tough Times (A)". Harvard Business School Case 403-069. (Revised April 2003).

Ke, Rongzhu, Jin Li, and Michael Powell (2018). "Managing Careers in Organizations". *Journal of Labor Economics* 36 (1), pp. 197–252.

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

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

Lloyd, Zach (May 2022). *The Google Incentive Mismatch: Problems with Promotion-Oriented Cultures*. Accessed on June 19, 2023. URL: https://www.warp.dev/blog/problems-with-promotion-oriented-cultures.

Lu, Alice, Peter I. Frazier, and Oren Kislev (2018). "Surge Pricing Moves Uber's Driver-Partners". In: *Proceedings of the 2018 ACM Conference on Economics and Computation*, p. 3.

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

Moldovanu, Benny, Aner Sela, and Xianwen Shi (2007). "Contests for Status". *Journal of Political Economy* 115 (2), pp. 338–363.

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.

Prendergast, Canice (1999). "The Provision of Incentives in Firms". *Journal of Economic Literature* 37 (1), pp. 7–63.

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

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

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

A Proofs

A.1 Proof of Lemma 1

Fix such a scheme $\mathcal{C} = (q_A, q_B, T_A, T_B)$. If $T_A \geq T_B$, then let $T'_A = T_A - T_B$ and $\mathcal{C}' = (q_A, q_B, T'_A, 0)$. This modification preserves the total number of prizes awarded and induces the same fraction of workers to choose task A. Hence $\mathcal{C}'$ is feasible. Since a lower bonus is paid to workers choosing each task under $\mathcal{C}'$ than under $\mathcal{C}$, it must therefore be that $\Pi(\mathcal{C}') > \Pi(\mathcal{C})$.

Suppose instead that $T_A < T_B$. Then set

$$q'_A = q_A + (1 - F(v))\frac{T_B - T_A}{V}, \quad q'_B = q_B - F(v)\frac{T_B - T_A}{V}$$

and $\mathcal{C}' = (q'_A, q'_B, 0, 0)$. Note that

$$V \cdot (q'_B - q'_A) = V \cdot (q_B - q_A) + T_A - T_B = v,$$

and so $\mathcal{C}'$ induces the required fraction of workers to choose each task. Since $v > 0$, it must be that $q_B > q'_B > q'_A > q_A$ and so q'_A and q'_B lie in $[0, 1]$. Additionally,

$$F(v)(1 - q'_A) + (1 - F(v))(1 - q'_B) = F(v)(1 - q_A) + (1 - F(v))(1 - q_B),$$

and so $\mathcal{C}'$ allocates the same number of prizes as $\mathcal{C}$. Therefore $\mathcal{C}'$ is feasible.

Let $n \equiv \Delta q \cdot F(v) \cdot (1 - F(v))$ and $n' \equiv \Delta q' \cdot F(v) \cdot (1 - F(v))$ be the number of prizes reallocated between groups under the two incentive schemes. The difference in profits between the two schemes can be written

$$\Pi(\mathcal{C}') - \Pi(\mathcal{C}) = F(v) \cdot T_A + (1 - F(v)) \cdot T_B + \Pi_P(n') - \Pi_P(n),$$

where

$$\Pi_P(m) \equiv F(v) \int_{1-B-m/F(v)}^{1} \rho(q)\, dq + (1 - F(v)) \int_{1-B+m/(1-F(v))}^{1} \rho(q)\, dq$$

is the match value generated by prizes under a scheme which reallocates m prizes and $B \equiv F(v)(1 - q_A) + (1 - F(v))(1 - q_B)$ is the total number of prizes allocated across both groups. Note that

$$\Pi'_P(m) = \rho\left(1 - B - \frac{m}{F(v)}\right) - \rho\left(1 - B + \frac{m}{1 - F(v)}\right) < 0,$$

and so $\Pi_P(n') > \Pi_P(n)$ given that $\Delta q' < \Delta q$. Therefore $\Pi(\mathcal{C}') > \Pi(\mathcal{C})$, as claimed.

A.2 Proof of Lemma 2

The scheme $\mathcal{C} = (1 - \beta, 1 - \beta, v)$ is feasible, and so the set of feasible schemes is non-empty. The bounds $q_A, q_B \in [0, 1]$ and $T_A \geq 0$ along with the targeting constraint $v = T + V \cdot \Delta q$ imply that the set of feasible schemes is bounded. Further, the feasibility constraints are weak inequalities which are continuous in (q_A, q_B, T). Hence the set of feasible schemes is compact. Meanwhile, Π is continuous in (q_A, q_B, T). The extreme value theorem therefore implies that an optimal scheme $\mathcal{C}^* = (q_A^*, q_B^*, T^*)$ exists.

We next prove that this optimum is unique. Note that Π is additively separable in (q_A, q_B, T), linear in T, and strictly concave in q_A and q_B given that ρ is an increasing function. Meanwhile, all feasibility constraints are linear in in (q_A, q_B, T). So fix any optimal scheme $\mathcal{C}^{**} = (q_A^{**}, q_B^{**}, T^{**})$, and suppose by way of contradiction that $\mathcal{C}^{**} \neq \mathcal{C}^*$. At least one of $q_A^* \neq q_A^{**}$ or $q_B^* \neq q_B^{**}$ must hold, since otherwise the targeting constraint implies that also $T^* = T^{**}$ and therefore $\mathcal{C}^* = \mathcal{C}^{**}$. But in that case, any convex combination of $\mathcal{C}^*$ and $\mathcal{C}^{**}$ preserves feasibility given the linearity of each constraint in all variables. Further, this mixture must strictly increase profits given linearity of Π in T and strict concavity in q_A and q_B. This result contradicts the presumed optimality of each scheme.

Now, fix $v > 0$ and any feasible scheme $\mathcal{C} = (q_A, q_B, T)$ satisfying $T \geq v$. We next establish that $\mathcal{C}$ is not an optimal scheme and therefore that $T^* < v$. The bound $T \geq 0$ combined with the targeting constraint implies that $q_A \geq q_B$, in which case the resource constraint implies that $q_A \geq 1 - \beta$ since the resource constraint must hold. If $T^* > v$, then $q_A > q_B$ so that $q_B < 1$. If $T^* = v$ and $q_A = q_B = 1$, then $\mathcal{C}' = (1 - \Delta, 1 - \Delta, T)$ is feasible for sufficiently small Δ and yields a strictly higher profit for the organization than $\mathcal{C}$. So assume that $q_B < 1$. Define the family of schemes

$$\mathcal{C}'(\Delta) = (q_A - \Delta \cdot (1 - F(v))/V, q_B + \Delta \cdot F(v)/V, T - \Delta).$$

The scheme $\mathcal{C}'(\Delta)$ satisfies the targeting and resource constraint for all Δ, and for sufficiently small $\Delta > 0$ it is feasible given that $q_A \geq 1 - \beta > 0$ and $q_B > 1$ and $T \geq v > 0$. Profits under this scheme are

$$\begin{aligned}&\Pi(\mathcal{C}'(\Delta))\\&=\Pi(\mathcal{C})+F(v)\left(\Delta+\int_{q_A-\Delta\cdot(1-F(v))/V}^{q_A}\rho(q)\,dq\right)-(1-F(v))\int_{q_B}^{q_B+\Delta\cdot F(v))/V}\rho(q)\,dq.\end{aligned}$$

Differentiating wrt Δ yields

$$\frac{d}{d\Delta}\Pi(\mathcal{C}'(\Delta))=F(v)+\frac{F(v)(1-F(v))}{V}\left[\rho\left(q_A-\Delta\frac{1-F(v)}{V}\right)-\rho\left(q_B+\Delta\frac{F(v)}{V}\right)\right].$$

At $\Delta=0$ this expression becomes

$$\frac{d}{d\Delta}\Pi(\mathcal{C}'(\Delta))\bigg|_{\Delta=0}=F(v)+\frac{F(v)(1-F(v))}{V}\left[\rho(q_A)-\rho(q_B)\right],$$

which is strictly positive given that $q_A\geq q_B$. Hence $\Pi(\mathcal{C}'(\Delta))>\Pi(\mathcal{C})$ for sufficiently small $\Delta>0$.

Given that $T^*<v$, the targeting constraint requires that $q_B^*>q_A^*$. The resource constraint therefore requires that $q_B^*>1-\beta$. Now, if $q_A^*\geq 1-\beta$, then the resource constraint is slack under $\mathcal{C}^*$. In this case, for sufficiently small $\Delta>0$ the scheme $\mathcal{C}'=(q_A^*-\Delta,q_B^*-\Delta,T^*)$ is feasible and strictly increases the match value from prizes, so that $\Pi(\mathcal{C}')>\Pi(\mathcal{C}^*)$, contradicting the optimality of $\mathcal{C}^*$. Therefore $q_A^*<1-\beta$.

A.3 Proof of Lemma 3

For $T\leq v$, define $Q(T)\equiv(v-T)/V\in[0,1]$. The set $\mathbb{C}(T)$ consists of triples (q_A,q_B,T) such that $(q_A,q_B)\in[0,1]^2$ and the system

$$\begin{cases}q_B-q_A=Q(T)\\F(v)q_A+(1-F(v))q_B\geq 1-\beta\end{cases}$$

is satisfied. Fix $T\in[0,v]$. The first constraint implies that the set of feasible quantiles (q_A,q_B) is totally ordered using the usual vector order. Since ΔR is increasing in both q_A and q_B, it is therefore minimized by the smallest feasible (q_A,q_B). Using the targeting constraint to eliminate q_B from the resource constraint yields

$$q_A\geq 1-\beta-(1-F(v))Q(T).$$

Hence, provided that the feasible set is non-empty, the quantiles $(q_A^{**}(T), q_B^{**}(T))$ which minimize $\Delta R(q_A, q_B)$ satisfy

$$q_A^{**}(T) = \max\{0, 1-\beta-(1-F(v))Q(T)\}, \quad q_B^{**}(T) = \max\{Q(T), 1-\beta+F(v)Q(T)\}.$$

This set of quantiles is feasible if and only if $q_B^{**}(T) \leq 1$. Otherwise, no (q_A, q_B) satisfying the targeting and resource constraints satisfies $q_B \leq 1$, in which case $\mathbb{C}(T)$ is empty. Note that $q_B^{**}(T)$ is continuous and decreasing in T, and $q_B^{**}(v) = 1-\beta < 1$. Hence $\underline{T} \equiv \inf\{T \in [0, v] \ : \ q_B^{**}(T) \leq 1\} \in [0, v)$, and $q_B^{**}(T) \leq 1$ iff $T \geq \underline{T}$. In particular, $\Delta R^*(T)$ is finite iff $T \geq \underline{T}$.

Going forward, we restrict attention to $T \in [\underline{T}, v]$. Let $\overline{Q} \equiv Q(\underline{T})$. Since Q is linear and decreasing in T, the function $\Delta R^*(T)$ is decreasing and strictly convex in T on $[\underline{T}, v]$ if and only if the function

$$\Delta R^{**}(Q) \equiv (1-F(v)) \int_{1-\beta}^{\max\{Q, 1-\beta+F(v)Q\}} \rho(q)\, dq - F(v) \int_{\max\{0, 1-\beta-(1-F(v))Q\}}^{1-\beta} \rho(q)\, dq$$

it is increasing and strictly convex in Q on $[0, \overline{Q}]$. Let $\underline{Q} \equiv \sup\{Q \in [0, \overline{Q}] \ : \ 1-\beta-(1-F(v))Q > 0\} \in (0, \overline{Q}]$. Then on $[0, \underline{Q}]$, we have

$$\Delta R^{**}(Q) = (1-F(v)) \int_{1-\beta}^{1-\beta+F(v)Q} \rho(q)\, dq - F(v) \int_{1-\beta-(1-F(v))Q}^{1-\beta} \rho(q)\, dq.$$

Differentiating this expression wrt Q yields

$$\frac{d\Delta R^{**}}{dQ} = F(v)(1-F(v)) \left[\rho(1-\beta+F(v)Q) - \rho(1-\beta-(1-F(v))Q)\right].$$

Since ρ is increasing, this derivative is nonnegative and increasing in Q on $[0, \underline{Q}]$. Hence ΔR^{**} is increasing and convex in Q on this interval. Meanwhile, on $(\underline{Q}, \overline{Q}]$ we have

$$\Delta R^{**}(Q) = (1-F(v)) \int_{1-\beta}^{Q} \rho(q)\, dq - F(v) \int_{0}^{1-\beta} \rho(q)\, dq,$$

which has derivative

$$\frac{d\Delta R^{**}}{dQ} = (1-F(v))\rho(Q).$$

Since ρ is nonnegative and increasing, so is this derivative. Hence ΔR^{**} is increasing and convex in Q on $(\underline{Q}, \overline{Q}]$. Since ΔR^{**} is additionally continuous on $[0, \overline{Q}]$, it follows that it is increasing on $[0, \overline{Q}]$. Finally, when $\underline{Q} < \overline{Q}$, we have

$$\frac{d\Delta R^{**}}{dQ^-}(\underline{Q}) = F(v)(1-F(v)) \left[\rho(\underline{Q}) - \rho(0)\right] < (1-F(v))\rho(\underline{Q}) = \frac{d\Delta R^{**}}{dQ^+}(\underline{Q}),$$

and so ΔR^{**} has a convex kink at $Q = \underline{Q}$. Hence ΔR^{**} is strictly convex on $[0, \overline{Q}]$.

A.4 Proof of Lemma 4

Recall from the proof of Lemma 3 that $\underline{T} = \inf\{T \in [0, v] \ : \ q_B^{**}(T) \leq 1\}$. Since $q_B^{**}(T) = \max\{Q(T), 1 - \beta + F(v)Q(T)\}$, this expression may be equivalently written $\underline{T} = \inf\{T \in [0, v] \ : \ Q(T) \leq 1 \text{ and } 1 - \beta + F(v)Q(T) \leq 1\}$, where $Q(T) = (v - T)/V$.

The inequality $Q(T) \leq 1$ is equivalently $T \geq v - V$, while the inequality $1 - \beta + F(v)Q(T) \leq 1$ is equivalently $T \geq v - V \cdot \beta/F(V)$. Hence

$$\underline{T} = \max\left\{0, v - V, v - \frac{V \cdot \beta}{F(v)}\right\}.$$

This expression is clearly nondecreasing in v and increasing whenever it is positive. Additionally, it vanishes if and only if $v \leq V$ and $vF(v) \leq V \cdot \beta$. Since $vF(v)$ is increasing in v, these two inequalities are satisfied if and only if $v \leq \overline{v}_0$, where

$$\overline{v}_0 \equiv \max\{v \in [0, V] \ : \ vF(v) \leq \beta \cdot V\}.$$

Clearly $\overline{v}_0 \in (0, V]$, proving the result.

A.5 Proof of Lemma 5

For $v \geq \overline{v}_0$, define

$$\Pi_0'(v) \equiv \begin{cases} \frac{1-F(v)}{F(v)} \cdot \frac{\rho(1)}{V} - 1, & F(v) < \beta \\ (1 - F(v)) \cdot \frac{\rho(1) - \rho(1-\beta/F(v))}{V} - 1, & F(v) \geq \beta \end{cases}$$

Note that Π_0' is decreasing and right-continuous in v and $\Pi_0'(v) \to -1$ as $v \to \infty$. We complete the proof by demonstrating that $\overline{v} \equiv \min\{v \geq \overline{v}_0 \ : \ \Pi_0'(v) \leq 0\}$ has the claimed properties. Since $\underline{T} = 0$ for $v < \overline{v}_0$, there is nothing to prove on this regime, so we focus on $v \geq \overline{v}_0$.

The proof of Lemma 3 established that

$$\Pi^*(T) \equiv \Pi^{FB} - F(v) \cdot T - (1 - F(v)) \int_{1-\beta}^{\max\{(v-T)/V, 1-\beta+F(v)(v-T)/V\}} \rho(q)\, dq$$
$$+ F(v) \int_{\max\{0, 1-\beta-(1-F(v))(v-T)/V\}}^{1-\beta} \rho(q)\, dq$$

for all $T \in [\underline{T}, v]$. Define

$$\overline{T} \equiv v - \frac{(1-\beta)V}{1 - F(v)}.$$

For all $T \neq \overline{T}$, the derivative $d\Pi^*/dT$ exists and equals

$$\frac{d\Pi^*}{dT}(T) = F(v) \cdot \left(\frac{1 - F(v)}{V} \cdot \Delta\bar{\rho}(T) - 1 \right),$$

where

$$\Delta\bar{\rho}(T) \equiv \begin{cases} \rho\left(1 - \beta + F(v)\frac{v-T}{V}\right) - \rho\left(1 - \beta - (1 - F(v))\frac{v-T}{V}\right), & T > \overline{T}(v) \\ \rho\left(\frac{v-T}{V}\right) / F(v), & T < \overline{T}(v) \end{cases}$$

Additionally, at $T = \overline{T}$ one-sided derivatives of Π^* exist and equal

$$\frac{d\Pi^*}{dT+}(\overline{T}) = \frac{d\Pi^*}{dT}(\overline{T}+), \quad \frac{d\Pi^*}{dT-}(\overline{T}) = \frac{d\Pi^*}{dT}(\overline{T}-).$$

Suppose first that $F(V) \geq \beta$. Then $\overline{v}_0 \leq V$ satisfies $\overline{v}_0 F(\overline{v}_0) = V \cdot \beta$, implying $F(\overline{v}_0) = \beta \cdot \frac{V}{\overline{v}_0} \geq \beta$. Hence $F(v) \geq \beta$ for all $v \geq \overline{v}_0$. In this case $\underline{T} = v - V \cdot \beta / F(v)$ and $\underline{T} \geq \overline{T}$ for all $v \geq \overline{v}_0$. Hence

$$\frac{d\Pi^*}{dT}(\underline{T}) = F(v) \cdot \Pi_0'(v)$$

for all $v \geq \overline{v}_0$. Since $T^* = \underline{T}$ if and only if this derivative is non-positive, and since $\Pi_0'(v) \leq 0$ if and only if $v \geq \overline{v}$, it follows that $T^* = \underline{T}$ on $[\overline{v}_0, \infty)$ if and only if $v \geq \overline{v}$. When $v \geq \overline{v}$, we have $q_A^{**}(T^*) = 1 - \beta/F(v)$ and $q_B^{**}(T^*) = 1$, and so $\mathcal{C}^* = (1 - \beta/F(v), 1, \underline{T})$ for all $v \geq \overline{v}$, as claimed. Meanwhile for $v \in (\overline{v}_0, \overline{v})$ (supposing this interval is non-empty) we have $T^* > \underline{T}$, while for $v \in [0, \overline{v}_0]$ we have $\underline{T} = 0$. So $\overline{v}$ satisfied the stated properties when $F(V) \geq \beta$.

Now suppose that $F(V) < \beta$. Then $\overline{v}_0 = V$, and for $v \in [V, F^{-1}(\beta))$ we have $\underline{T} = v - V$ and $\overline{T} > \underline{T}$. Hence

$$\frac{d\Pi^*}{dT}(\underline{T}) = F(v) \cdot \Pi_0'(v)$$

for all $v \in [V, F^{-1}(\beta))$. Meanwhile for $v \geq F^{-1}(\beta)$ we have $\underline{T} = v - V \cdot \beta/F(v)$ and $\underline{T} \geq \overline{T}$, in which case again

$$\frac{d\Pi^*}{dT}(\underline{T}) = F(v) \cdot \Pi_0'(v).$$

It follows that $T^* = \underline{T}$ on $[\overline{v}_0, \infty)$ if and only if $v \geq \overline{v}$. When $v \geq \overline{v}$, we have $q_A^{**}(T^*) = \max\{0, 1 - \beta/F(v)\}$ and $q_B^{**}(T^*) = 1$, and so $\mathcal{C}^* = (\max\{0, 1 - \beta/F(v)\}, 1, \underline{T})$ for all $v \geq \overline{v}$, as claimed. Meanwhile for $v \in (\overline{v}_0, \overline{v})$ (supposing this interval is non-empty) we have $T^* > \underline{T}$, while for $v \in [0, \overline{v}_0]$ we have $\underline{T} = 0$. So $\overline{v}$ satisfied the stated properties when $F(V) < \beta$.

A.6 Proof of Proposition 1

The comparative statics involving $\Delta\rho^*$ and q_B^* are a special case of the corresponding results established in Theorem 1. The comparative static involving Δq^* follows from the fact that $\Delta\rho^* = R \cdot \Delta q^*$ under Assumption 1.

A.7 Proof of Proposition 2

Throughout this proof, we will write v explicitly as an argument of quantities which depend on it. We will additionally make free use of concepts and notation developed in the proof of Lemma 5. We begin with a technical lemma useful for establishing quasiconcavity:

Lemma A.1. *Let*

$$\alpha(v) \equiv v - \frac{K}{1 - F(v)}$$

for some $K \in \mathbb{R}$ *on a non-empty interval* $[v_0, v_1] \subset \mathbb{R}_+$*. If Assumption 2 holds, then* α *is strictly quasiconcave on* $[v_0, v_1]$*.*

Proof. α is continuously differentiable everywhere and satisfies the ordinary differential equation

$$\alpha'(v) = \frac{\alpha(v) - \varphi(v)}{v - \varphi(v)}.$$

Any critical point v^* of α must therefore satisfy $\alpha(v^*) = \varphi(v^*)$. Under Assumption 2, the virtual cost φ is increasing, implying that for any two critical points v^* and $v^{**} > v^*$ we have $\alpha(v^{**}) > \alpha(v^*)$.

Fix w_0 and $w_1 > w_0$ in $[v_0, v_1]$. Strict quasiconcavity requires that $\alpha(v) > \min\{\alpha(w_0), \alpha(w_1)\}$ for all $v \in (w_0, w_1)$. Suppose by way of contradiction that this condition were violated. Then there exists a minimizer $w^* \in (w_0, w_1)$ of α on $[w_0, w_1]$. This minimizer must be a critical point. If there existed another critical point $w^{**} \in [w_0, w^*)$, then $\alpha(w^*) > \alpha(w^{**})$, contradicting the minimality of w^* on $[w_0, w_1]$. Hence $\alpha'(v) \neq 0$ for every $v \in [w_0, w^*)$. This fact, combined with the minimality of w^* and the continuity of α', implies that $\alpha'(v) < 0$ for every $v \in [w_0, v^*)$. In particular, $\alpha'(w_0) < 0$.

Now, $\alpha'(w^*) = 0$ implies that $\alpha(w^*) = \varphi(w^*)$. Meanwhile, $\alpha'(w_0) < 0$ implies that $\alpha(w_0) < \varphi(w_0)$ given that $v > \varphi(v)$ for all v. But then monotonicity of φ implies that $\alpha(w_0) < \alpha(w^*)$, a contradiction of the minimality of w^*. So strict quasiconcavity must hold. □

Note that the objective $\Pi^*(T; v)$ and the constraint set $[\underline{T}, v]$ are both continuous in v, and therefore the optimal bonus $T^*(v)$ is continuous in v as well. Under Assumption 1, we have

$$\Delta\bar{\rho}(T; v) = \begin{cases} R \cdot \frac{v-T}{V}, & T > \overline{T}(v) \\ R \cdot \left(\frac{v-T}{V} + b\right) / F(v), & T < \overline{T}(v) \end{cases}$$

The facts that $T^*(v) = 0$ for v sufficiently small and $\max_{v \in [0,\overline{v}]} T^*(v) > 0$ for R sufficiently large are special cases of the corresponding results in Theorem 2. We next prove quasiconcavity. Let $\Phi(v) \equiv v(1 - F(v))$. Note that $\overline{T}(v) > 0$ if and only if $\Phi(v) > (1 - \beta)V$. Since $\Phi'(v) = -\varphi(v) f(v)$, Assumption 2 implies that Φ is single-peaked on $[0, \infty)$. Let $v^* \in (0, \infty)$ be this peak and $\Phi^* \equiv \Phi(v^*)$. There exist between zero and two positive solutions to the equation $\Phi(v) = (1 - \beta)V$. Let $v_L > 0$ and $v_H > v_L$ denote these solutions when two exist. Otherwise, let $v_L = v_H = v^*$ if $(1-\beta)V \geq \Phi^*$ and $v_H = \infty$ if $(1-\beta)V \leq \Phi(\infty)$. In all cases, $\overline{T}(v) > 0$ if and only if $v \in (v_L, v_H)$.

Case 1: $v_L = v_H$ or $\overline{v} \leq v_L$. In this case $\overline{T}(v) = 0$ for all $v \in [0, \overline{v})$. Since additionally either $\underline{T}(v) = 0$ or $T^*(v) > \underline{T}(v)$ for all $v < \overline{v}$, in this case the optimal bonus $T^*(v)$ is the unique solution to the first-order condition

$$\frac{d\Pi^*}{dT}(T; v) = F(v) \cdot \left((v - T)(1 - F(v)) \cdot \frac{R}{V^2} - 1 \right) \leq 0,$$

with equality if $T^*(v) > 0$. This condition yields the optimal bonus

$$T^*(v) = \max\left\{ v - \frac{V^2/R}{1 - F(v)}, 0 \right\}$$

for all $v \in [0, \overline{v})$. Continuity of T^* allows us to extend this identity to $v = \overline{v}$. Lemma A.1 ensures that the function

$$\tau(v) \equiv v - \frac{V^2/R}{1 - F(v)}$$

is strictly quasiconcave on $[0, v]$. Hence it is positive, if at all, on an interval, ensuring that $T^*(v) = \max\{\tau(v), 0\}$ is quasiconcave on $[0, \overline{v}]$.

Case 2: $v_L < \hat{v} \equiv \min\{v_H, \overline{v}\}$ and $T^*(v_L) > 0$. Then $T^*(v_L) > \underline{T}(v_L)$ and therefore $\frac{d\Pi^*}{dT}(\underline{T}(v_L); v_L) > 0$. Since $\overline{T}(v_L) = 0$, this condition is equivalently

$$(v_L - \underline{T}(v_L))(1 - F(v_L)) \cdot \frac{R}{V^2} > 1,$$

implying in particular

$$v_L(1 - F(v_L)) \cdot \frac{R}{V^2} > 1.$$

Since $v_L(1 - F(v_L)) = (1 - \beta)V$, this latter inequality is equivalently $(1 - \beta) \cdot \frac{R}{V} > 1$. But then whenever $\overline{T}(v) > \underline{T}(v)$, we have

$$\frac{d\Pi^*}{dT+}(\overline{T}(v); v) = F(v)\left[(v - \overline{T}(v))(1 - F(v)) \cdot \frac{R}{V^2} - 1\right] = F(v)\left[(1 - \beta) \cdot \frac{R}{V} - 1\right] > 0,$$

meaning that $T^*(v) > \overline{T}(v)$ whenever $\overline{T}(v) > \underline{T}(v)$. Reasoning very similar to the previous case therefore allows us to conclude that

$$T^*(v) = \max\left\{v - \frac{V^2/R}{1 - F(v)}, 0\right\}$$

on $v \in [0, \overline{v}]$, and therefore that T^* is quasiconcave on this interval.

Case 3: $v_L < \hat{v}$ **and** $T^*(v_L) = 0$. In this case the first-order condition

$$\frac{d\Pi^*}{dT}(0; v_L) = F(v_L) \cdot \left(\Phi(v_L) \cdot \frac{R}{V^2} - 1\right) = F(v_L) \cdot \left((1 - \beta) \cdot \frac{R}{V} - 1\right) \leq 0$$

must hold. Additionally, for $v < v_L$ we have $\overline{T}(v) = \underline{T}(v) = 0$ and

$$\frac{d\Pi^*}{dT}(0; v) = F(v) \cdot \left(\Phi(v) \cdot \frac{R}{V^2} - 1\right),$$

which is negative given that $\Phi(v) < (1 - \beta) \cdot V = \Phi(v_L)$ for all $v < v_L$. So $T^*(v) = 0$ for $v \leq v_L$. It is therefore sufficient to establish quasiconcavity on $[v_L, \overline{v}]$.

For $v \in (v_L, \hat{v})$ we have

$$\begin{aligned}
\frac{d\Pi^*}{dT+}(\max\{\underline{T}(v), \overline{T}(v)\}; v) &= F(v) \cdot \left((1 - F(v))(v - \max\{\underline{T}(v), \overline{T}(v)\}) \cdot \frac{R}{V^2} - 1\right) \\
&\leq F(v) \cdot \left((1 - F(v))(v - \overline{T}(v)) \cdot \frac{R}{V^2} - 1\right) \\
&= F(v) \cdot \left((1 - \beta) \cdot \frac{R}{V} - 1\right) \leq 0.
\end{aligned}$$

Hence $T^*(v) \leq \max\{\underline{T}(v), \overline{T}(v)\}$ on $(v_L, \hat{v})$. But also $T^*(v) \geq \underline{T}(v)$, and for every $v < \overline{v}$ and this inequality is strict if $\underline{T}(v) > 0$. Thus if $\underline{T}(v) \geq \overline{T}(v)$ on $(v_L, \hat{v})$, where $\overline{T}(v) > 0$, we would conclude that $\underline{T}(v) < T^*(v) \leq \underline{T}(v)$, a contradiction. So it must be that $\overline{T}(v) > \underline{T}(v)$ and $T^*(v) \leq \overline{T}(v)$ on $(v_L, \hat{v})$.

On $(v_L, \hat{v})$, the left-hand derivative of Π^* at $\overline{T}(v)$ is

$$\frac{d\Pi^*}{dT-}(\overline{T}(v); v) = F(v) \cdot \left[\left(\frac{1-\beta}{F(v)} + b \cdot \frac{1-F(v)}{F(v)} \right) \cdot \frac{R}{V} - 1 \right],$$

where the bracketed term is decreasing in v. Hence there exists a $v_0 \in [v_L, \hat{v}]$ such that $T^*(v) = \overline{T}(v)$ for $v \in [v_L, v_0]$ while $T^*(v) < \overline{T}(v)$ for $v \in (v_0, \hat{v})$.

We next establish quasiconcavity on the intervals $[v_L, v_0]$ and $[v_0, \hat{v}]$. This claim is trivial on a degenerate interval, so for this argument we assume that each interval is non-degenerate. On $[v_L, v_0]$, we have

$$T^*(v) = \overline{T}(v) = v - \frac{(1-\beta)V}{1-F(v)},$$

which is strictly quasiconcave by Lemma A.1. Meanwhile, on $(v_0, \hat{v})$ the optimal bonus is characterized by the first-order condition

$$\frac{d\Pi^*}{dT}(T; v) = F(v) \cdot \left[\frac{1-F(v)}{F(v)} \cdot \left(\frac{v-T}{V} + b \right) \cdot \frac{R}{V} - 1 \right] \leq 0,$$

with equality if $T^*(v) > 0$. This first-order condition has the unique solution

$$T^*(v) = \max \left\{ v - \frac{V^2/R}{1-F(v)} + V \cdot b + \frac{V^2}{R}, 0 \right\},$$

which by continuity of T^* must also hold at $v = v_0, \hat{v}$. Lemma A.1 ensures that the function

$$\tau(v) \equiv v - \frac{V^2/R}{1-F(v)}$$

is strictly quasiconcave. It therefore exceeds $-V \cdot B - V^2/R$ (if at all) on an interval, implying that $T^*(v) = \max\{\tau(v) + V \cdot b + V^2/R, 0\}$ is quasiconcave.

We next prove that T^* is quasiconcave on $[v_L, \hat{v}]$. If $v_0 \in \{v_L, \hat{v}\}$ then this claim is immediate, so suppose that $v_0 \in (v_L, \hat{v})$. Quasiconcavity on $[v_L, v_0]$ and $[v_0, \hat{v}]$ implies quasiconcavity on $[v_L, \hat{v}]$ so long as T^* has a (weakly) concave kink at v_0. Using the expressions for T^* to the left and right of v_0 derived above, we have

$$\frac{dT^*}{dv-}(v_0) = 1 - (1-\beta) \cdot V \cdot \frac{f(v_0)}{(1-F(v_0))^2}, \quad \frac{dT^*}{dv+}(v_0) = 1 - \frac{V^2}{R} \cdot \frac{f(v_0)}{(1-F(v_0))^2}.$$

Since $(1-\beta) \cdot \frac{R}{V} \leq 1$ in the current case, it follows that the left-hand derivative is no smaller than the right-hand one, as claimed.

If $\hat{v} = \overline{v}$, then we have established quasiconcavity on $[v_L, \overline{v}]$ and therefore $[0, \overline{v}]$. So suppose instead that $\hat{v} = v_H < \overline{v}$. Then for all $v \in [v_H, \overline{v}]$ we have $\Phi(v) \leq \Phi(v_H) = (1-\beta) \cdot V$ and therefore

$$\begin{aligned} \frac{d\Pi^*}{dT}(\underline{T}(v); v) &= F(v) \cdot \left((1 - F(v))(v - \underline{T}(v)) \cdot \frac{R}{V^2} - 1 \right) \\ &\leq F(v) \cdot \left(\Phi(v) \cdot \frac{R}{V^2} - 1 \right) \\ &\leq F(v) \cdot \left((1-\beta) \cdot \frac{R}{V} - 1 \right) \leq 0. \end{aligned}$$

Hence $T^*(v) = \underline{T}(v)$ on $[v_H, \overline{v}]$. But for $v < \overline{v}$ this is possible only if $\underline{T}(v) = 0$, and so $T^*(v) = 0$ on $[v_H, \overline{v})$. Thus by continuity also $T^*(\overline{v}) = 0$. This fact, combined with quasiconcavity of T^* on $[0, v_H]$, implies that T^* is quasiconcave on $[0, \overline{v}]$.

To complete the proof of the proposition, we construct an interval of parameters R for which $T^*(v)$ is non-monotone on $[0, \overline{v}]$ when V is sufficiently large. Suppose that $V > \Phi^*/(1-\beta)$. Then $\overline{T}(v) = 0$ for all v and therefore

$$T^*(v) = \max\left\{ v - \frac{V^2/R}{1 - F(v)}, 0 \right\}$$

on $[0, \overline{v}]$. For v sufficiently small, we have $T^*(v) = 0$. Non-monotonicity is therefore ensured whenever $T^*(\overline{v}) > 0$ and $\frac{dT^*}{dv}(\overline{v}) < 0$. Equivalently, we require

$$\frac{V^2}{\overline{v}(1 - F(\overline{v}))} < R < V^2 \frac{f(\overline{v})}{(1 - F(\overline{v}))^2}.$$

These two inequalities define a non-degenerate interval if and only if $\varphi(\overline{v}) > 0$. Since $\overline{v} \geq \overline{v}_0$, it is sufficient that $\varphi(\overline{v}_0) > 0$. Note that $\overline{v}_0$ is continuous and increasing in V, and for V large enough that $F(V) > \beta$ we have $\overline{v}_0 F(\overline{v}_0) = V \cdot \beta$, implying that $\overline{v}_0 \to \infty$ as $V \to \infty$. Hence if V is chosen large enough, there exists a non-degenerate interval of R on which $T^*(v)$ is non-monotone.

A.8 Proof of Theorem 1

Throughout this proof, we will write v explicitly as an argument of quantities which depend on it. The objective function $\Pi^*(T; v)$ stated in the proof of Proposition 2 may be equivalently written in terms of Δq using the targeting constraint $v = T + V \cdot \Delta q$, yielding

$$\Pi^*(\Delta q; v) \equiv \Pi^{FB} - F(v) \cdot (v - V \cdot \Delta q) - (1 - F(v)) \int_{1-\beta}^{\max\{\Delta q, 1-\beta+F(v)\Delta q\}} \rho(q)\, dq$$
$$+ F(v) \int_{\max\{0, 1-\beta-(1-F(v))\Delta q\}}^{1-\beta} \rho(q)\, dq$$

over the constraint set $\Delta q \in [0, \overline{Q}(v)]$, where

$$\overline{Q}(v) \equiv \min\{v/V, \beta/F(v), 1\}.$$

Note that $\underline{T}(v) > 0$ is equivalent to $\overline{Q}(v) < v/V$. Hence for every $v < \overline{v}$, either $\overline{Q}(v) = v/V$ or else $\Delta q^*(v) < \overline{Q}(v)$. Additionally, since Π^* is continuous in Δq and v and $\overline{Q}(v)$ is continuous in v, the maximum theorem implies that $\Delta q^*(v)$ is continuous in v. And whenever $v > 0$, Lemma 2 implies that $\Delta q^*(v) > 0$.

Define

$$\underline{Q}(v) \equiv \frac{1-\beta}{1-F(v)}$$

and

$$\Delta\rho(\Delta q; v) \equiv \rho\left(\max\{\Delta q, 1 - \beta + F(v)\Delta q\}\right) - \rho\left(\max\{0, 1 - \beta - (1 - F(v))\Delta q\}\right)$$

for $\Delta q \in [0, \overline{Q}(v)]$. Then Π^* is differentiable for all $\Delta q \neq \underline{Q}(v)$, with derivative

$$\frac{d\Pi^*}{d\Delta q}(\Delta q; v) = F(v) \cdot [V - (1 - F(v))\Delta\rho(\Delta q; v)]$$

for $\Delta q < \underline{Q}(v)$ and

$$\frac{d\Pi^*}{d\Delta q}(\Delta q; v) = F(v) \cdot \left[V - \frac{1-F(v)}{F(v)}(\Delta\rho(\Delta q; v) + \rho(0))\right]$$

for $\Delta q > \Delta\underline{Q}(v)$. The profit function additionally has one-sided derivatives at $\Delta q = \underline{Q}(v)$ (whenever $\underline{Q}(v) \in [0, \overline{Q}(v)]$) with

$$\frac{d\Pi^*}{d\Delta q-}(\underline{Q}(v); v) = \frac{d\Pi^*}{d\Delta q}(\underline{Q}(v)-; v), \quad \frac{d\Pi^*}{d\Delta q+}(\underline{Q}(v); v) = \frac{d\Pi^*}{d\Delta q}(\underline{Q}(v)+; v).$$

For every $v < \overline{v}$, the optimal relative prize priority $\Delta q^*(v)$ is the unique Δq satisfying the appropriate first-order condition with respect to $\Pi^*(\Delta q; v)$ over $[0, v/V]$. Note in particular that $d\Pi^*/d\Delta q > 0$ at $\Delta q = 0$, so that $\Delta q^*(v) > 0$. Meanwhile, the optimal absolute priority satisfies $\Delta\rho^*(v) = \Delta\rho(\Delta q^*(v); v)$.

Lemma A.2. *If $\Delta q^*(v) < \Delta q^*(v')$ for $v, v' \in [0, \overline{v}]$ satisfying $v < v'$, then $\Delta\rho^*(v) < \Delta\rho^*(v')$.*

Proof. The inequality $\Delta q^*(v) < \Delta q^*(v')$ implies that $\Delta q^*(v) \in [0, \overline{Q}(v'))$, and so $\Delta\rho(\Delta q; v')$ is defined for $\Delta q \in [\Delta q^*(v), \Delta q^*(v')]$. Note that $\Delta\rho(\Delta q; v')$ is increasing in Δq on $[0, \overline{Q}(v')]$, since ρ is increasing and $\max\{\Delta q, 1-\beta+F(v)\Delta q\}$ is increasing in Δq while $\max\{0, 1-\beta-(1-F(v))\Delta q\}$ is nonincreasing in Δq. So we have

$$\Delta\rho(\Delta q^*(v'); v') > \Delta\rho(\Delta q^*(v); v').$$

We next establish that $\Delta q^*(v) \leq \overline{Q}(w)$ for every $w \in [v, v']$. Note that $\overline{Q}(w) = w/V$ for $w \leq \overline{v}_0$ and $\overline{Q}(w) = \min\{\beta/F(v), 1\}$ for $w \geq \overline{v}_0$. Hence $\overline{Q}$ is increasing to the left of $\overline{v}_0$ and nonincreasing to the right. Thus the desired result holds so long as $\Delta q^*(v) \leq \min\{\overline{Q}(v), \overline{Q}(v')\}$. The inequality $\Delta q^*(v) \leq \overline{Q}(v)$ holds by feasibility, while $\Delta q^*(v) \leq \overline{Q}(v')$ holds given the hypothesis $\Delta q^*(v) < \Delta q^*(v')$ combined with feasibility of $\Delta q^*(v')$.

To complete the proof, we establish that $\Delta\rho(\Delta q^*(v); w)$ is nondecreasing in w on $[v, v']$. If $\Delta q^*(v) = 0$ then $\Delta\rho(\Delta q^*(v), w) = 0$ for all w, so this result is immediate. Otherwise, whenever $F(w) > (1-\beta)/\Delta q^*(v) - 1$ we have

$$\Delta\rho(\Delta q^*(v); w) = \rho(1-\beta+F(w)\Delta q^*(v)) - \rho(1-\beta-(1-F(w))\Delta q^*(v)),$$

which has derivative

$$\frac{\partial}{\partial w}\Delta\rho(\Delta q^*(v); w) = \Delta q^*(v) \cdot f(w) \cdot \Delta\rho(\Delta q^*(v); w) > 0.$$

Hence $\Delta\rho(\Delta q^*(v); w)$ is increasing in w on this range. Meanwhile, whenever $F(w) \leq (1-\beta)/\Delta q^*(v) - 1$ we have

$$\Delta\rho(\Delta q^*(v); w) = \rho(\Delta q^*(v)) - \rho(0),$$

So $\Delta\rho(\Delta q^*(v); w)$ is independent of w on this range. It follows that $\Delta\rho(\Delta q^*(v); w)$ is nondecreasing in w on $[v, v']$, allowing us to conclude that

$$\Delta\rho(\Delta q^*(v'); v') > \Delta\rho(\Delta q^*(v); v') \geq \Delta\rho(\Delta q^*(v); v)$$

and therefore that $\Delta\rho^*(v') > \Delta\rho^*(v)$. □

Lemma A.3. *For every $v \geq 0$, the inequality*

$$\Delta\rho^*(v) \leq \frac{V}{1-F(v)}$$

holds.

Proof. The inequality is trivial if $v = 0$, so suppose $v > 0$. If $\Delta q^*(v) \leq \underline{Q}(v)$, then the fact that $\Delta q^*(v) > 0$ for every $v > 0$ means that $\Delta q^*(v)$ must satisfy the one-sided first-order condition

$$\frac{d\Pi^*}{d\Delta q-}(\Delta q^*(v); v) = F(v) \cdot [V - (1 - F(v))\Delta\rho^*(v)] \geq 0$$

(with equality whenever $\Delta q^*(v) < \underline{Q}(v)$), which is equivalent to the stated inequality. On the other hand, if $\Delta q^*(v) > \underline{Q}(v)$, then the first-order condition

$$\frac{d\Pi^*}{d\Delta q}(\Delta q^*(v); v) = F(v) \cdot \left[V - \frac{1 - F(v)}{F(v)}(\Delta\rho^*(v) + \rho(0))\right] \geq 0$$

must hold (with equality whenever $\Delta q^*(v) < \overline{Q}(v)$), or equivalently

$$\Delta\rho^*(v) \leq F(v) \cdot \frac{V}{1 - F(v)} - \rho(0).$$

Since $F(v) < 1$ and $\rho(0) \geq 0$, this inequality implies the desired one. $\square$

We now prove that $\Delta\rho^*(v)$ is increasing in v on $[0, \overline{v}]$. It is sufficient to prove this property on $[0, \overline{v})$, since continuity of $\Delta q^*(v)$ implies continuity of $\Delta\rho^*(v)$, so that strict monotonicity on $[0, \overline{v})$ implies the same property on $[0, \overline{v}]$.

Fix any $v_0, v_1 \in [0, \overline{v})$ satisfying $v_1 > v_0$. Suppose first that $\Delta q^*(v_1) > \overline{Q}(v_0)$. Then mechanically $\Delta q^*(v_0) \leq \overline{Q}(v_0) < \Delta q^*(v_1)$, implying $\Delta\rho^*(v_0) < \Delta\rho^*(v_1)$ by Lemma A.2.

Suppose next that $\underline{Q}(v_1) \leq \overline{Q}(v_0)$ and $\Delta q^*(v_1) \in [\underline{Q}(v_1), \overline{Q}(v_0)]$. If $\overline{Q}(v_1) = v_1/V$, then $\overline{Q}(v_0) \leq v_0/V < \overline{Q}(v_1)$, and so $\Delta q^*(v_1) < \overline{Q}(v_1)$. On the other hand, if $\overline{Q}(v_1) < v_1/V$, then $v_1 < \overline{v}$ implies that $\Delta q^*(v_1) < \overline{Q}(v_1)$. Hence in either case the one-sided first-order condition

$$\frac{d\Pi^*}{d\Delta q+}(\Delta q^*(v_1); v_1) = F(v_1) \cdot \left[V - \frac{1 - F(v_1)}{F(v_1)}(\Delta\rho(\Delta q^*(v_1); v_1) + \rho(0))\right] \leq 0$$

must hold (with equality if $\Delta q^*(v_1) > \underline{Q}(v_1)$), implying

$$V \leq \frac{1 - F(v_1)}{F(v_1)}(\Delta\rho(\Delta q^*(v_1); v_1) + \rho(0)) < \frac{1 - F(v_0)}{F(v_0)}(\Delta\rho(\Delta q^*(v_1); v_1) + \rho(0)).$$

Now, $\overline{Q}(v) = v/V$ for $v \leq \overline{v}_0$ and $\overline{Q}(v) = \min\{\beta/F(v), 1\}$ for $v \geq \overline{v}_0$. Hence $\overline{Q}$ is increasing to the left of $\overline{v}$ and nonincreasing to the right. Since $\Delta q^*(v_1) < \overline{Q}(v_1)$ and $\Delta q^*(v_1) \leq \overline{Q}(v_0)$, it follows that $\Delta q^*(v_1) \leq \overline{Q}(v)$ for all $v \in [v_0, v_1]$. Additionally, since $\underline{Q}(v)$ is increasing in v and $\Delta q^*(v_1) \geq \underline{Q}(v_1)$, we must have $\Delta q^*(v_1) > \underline{Q}(v)$ for all $v \in [v_0, v_1]$, in which case $\rho(\Delta q^*(v_1); v) = \rho(\Delta q^*(v_1)) - \rho(0)$ for $v \in [v_0, v_1]$ and therefore $\rho(\Delta q^*(v_1); v_1) = \rho(\Delta q^*(v_1); v_0)$. Hence

$$V < \frac{1 - F(v_0)}{F(v_0)} (\Delta\rho(\Delta q^*(v_1); v_0) + \rho(0)),$$

an inequality equivalent to

$$\frac{d\Pi^*}{d\Delta q}(\Delta q^*(v_1); v_0) < 0$$

and therefore implying $\Delta q^*(v_0) < \Delta q^*(v_1)$. Lemma A.2 then allows us to conclude that $\Delta\rho^*(v_0) < \Delta\rho^*(v_1)$.

Finally, suppose that $\Delta q^*(v_1) < \underline{Q}(v_1)$. Then $\Delta q^*(v_1) > 0$ means that the first-order condition

$$\frac{d\Pi^*}{d\Delta q}(\Delta q^*(v_1); v_1) = F(v_1) \cdot [V - (1 - F(v_1))\Delta\rho^*(v_1)] = 0$$

must hold, implying

$$\Delta\rho^*(v_1) = \frac{V}{1 - F(v_1)} > \frac{V}{1 - F(v_0)}.$$

Lemma A.3 therefore yields the inequality $\Delta\rho^*(v_1) > \Delta\rho^*(v_0)$.

We next prove that $q_B^*(v)$ is increasing in v on $[0, \overline{v}]$. The optimal absolute prize priority may be written $\rho^*(v) = \rho(q_B^*(v)) - \rho(q_A^*(v))$, where

$$\begin{aligned} q_B^*(v) &= \max\{\Delta q^*(v), 1 - \beta + F(v)\Delta q^*(v)\}, \\ q_A^*(v) &= \max\{0, 1 - \beta - (1 - F(v))\Delta q^*(v)\}. \end{aligned}$$

Fix v_0 and $v_1 > v_0$ in $[0, \overline{v}]$. If $q_A^*(v_1) \geq q_A^*(v_0)$, then the fact that $\rho^*(v_1) > \rho^*(v_0)$ implies $q_B^*(v_1) > q_B^*(v_0)$. So suppose that $q_A^*(v_0) > q_A^*(v_1)$. Then in particular $q_A^*(v_0) > 0$, and therefore

$$q_A^*(v_0) = 1 - \beta - (1 - F(v_0))\Delta q^*(v_0) > q_A^*(v_1) \geq 1 - \beta - (1 - F(v_1))\Delta q^*(v_1),$$

or after rearrangement

$$\Delta q^*(v_1) \geq \frac{1 - F(v_0)}{1 - F(v_1)} \Delta q^*(v_0) > \Delta q^*(v_0).$$

Then since $\max\{\Delta q, 1 - \beta + F(v)\Delta q\}$ is increasing in Δq and nondecreasing in v, it follows that $q_B^*(v_1) > q_B^*(v_0)$.

Finally, we show that $q_A^*(v)$ is non-monotone whenever V is sufficiently large. Let $\Phi(v) \equiv v(1-F(v))$. This function has derivative

$$\Phi'(v) = (1-F(v))\left(1 - v\frac{f(v)}{1-F(v)}\right).$$

Under Assumption 4, there exists a $v^* > 0$ such that $\Phi'(v) < 0$ for all $v > v^*$. Then since $\Phi(v) \le v$ for all $v \ge 0$, it follows that Φ has a finite maximum Φ^* on $\mathbb{R}_+$. It is sufficient for non-monotonicity that $\Delta q^*(v) = v/V$ and $\underline{Q}(v) \ge v/V$ for all $v \in [0, \overline{v}_0]$ and $\overline{v}_0 > v^*$, since the first two assumptions imply that

$$q_A^*(v) = 1 - \beta - \frac{\Phi(v)}{V}$$

on $[0, \overline{v}_0]$, and the final assumption implies non-monotonicity of $\Phi(v)$ on $[0, \overline{v}_0]$. We will show that all three assumptions hold for sufficiently large V.

Note that $\overline{v}_0$ is continuous and increasing in V, and for V large enough that $F(V) > \beta$ we have $\overline{v}_0 F(\overline{v}_0) = V \cdot \beta$, implying that $\overline{v}_0 \to \infty$ as $V \to \infty$. Hence $\overline{v}_0 > v^*$ for V sufficiently large. Meanwhile, the condition $\underline{Q}(v) \ge v/V$ is equivalent to $\Phi(v) \le (1-\beta) \cdot V$. Then $V > \Phi^*/(1-\beta)$ ensures that $\underline{Q}(v) \ge v/V$ for all $v \in [0, \overline{v}_0]$. Finally, whenever $v \le \overline{v}_0$ we have $\overline{Q}(v) = v/V$, so that $\Delta q = v/V$ is feasible, and whenever $\underline{Q}(v) \ge v/V$ we have $\Delta q^*(v) = v/V$ whenever the first-order condition

$$V \ge (1-F(v)) \cdot \left[\rho\left(1 - \beta + \frac{vF(v)}{V}\right) - \rho\left(1 - \beta - \frac{\Phi(v)}{V}\right)\right]$$

holds. This condition is satisfied for all v whenever $V > \rho(1)$.

A.9 Proof of Theorem 2

Throughout this proof, we will write v explicitly as an argument of quantities which depend on it. We will additionally make free use of concepts and notation developed in the proof of Lemma 5.

We begin by proving that $T^*(v) = 0$ for v sufficiently small. Note that $\overline{T}(v) = \underline{T}(v) = 0$ for v sufficiently small. In this regime, $T^*(v) = 0$ if and only if the first-order condition

$$\frac{d\Pi^*}{dT}(0; v) = F(v) \cdot \left\{\frac{1-F(v)}{V} \cdot \left[\rho\left(1 - \beta + \frac{vF(v)}{V}\right) - \rho\left(1 - \beta - \frac{v(1-F(v))}{V}\right)\right] - 1\right\} \le 0$$

holds, which is true for sufficiently small v given continuity of ρ.

We next prove that $\max_{v \in [0,\overline{v}]} T^*(v) > 0$ for R sufficiently large. For this it is sufficient to show that $T^*(\overline{v}_0) > 0$ for R large. Since $\underline{T}(\overline{v}_0) = 0$, we have $T^*(\overline{v}_0) > 0$ if and only if the first-order condition

$$\frac{d\Pi^*}{dT}(0; v) = F(\overline{v}_0) \cdot \left((1 - F(\overline{v}_0)) \cdot \Delta\bar{\rho}_0(0; \overline{v}_0) \cdot \frac{R}{V} - 1 \right) > 0,$$

where $\overline{v}_0$ and $\Delta\bar{\rho}_0(0; \overline{v}_0) \equiv \Delta\bar{\rho}(0; \overline{v}_0)/R$ are independent of R. Since $\Delta\bar{\rho}_0(0; \overline{v}_0) > 0$, this condition holds whenever R is sufficiently large.

We next derive conditions on R under which T^* is non-monotonic. The threshold costs $\overline{T}(v) = \underline{T}(v) = 0$ for sufficiently small v, in which case the slope of the profit function at $T = 0$ equals

$$\frac{d\Pi^*}{dT}(0; v) = F(v) \cdot \left\{ \frac{1 - F(v)}{V} \cdot \left[\rho \left(1 - \beta + \frac{vF(v)}{V} \right) - \rho \left(1 - \beta - \frac{v(1 - F(v))}{V} \right) \right] - 1 \right\}.$$

This expression is continuous in v and equals $-F(0) < 0$ at $v = 0$, meaning that $T^*(v) = 0$ for v sufficiently small. It is therefore sufficient for non-monotonicity of $T^*(v)$ that $T^*(\overline{v}_0) = 0$ and $\max_{v \in [0,\overline{v}_0]} T^*(v) > 0$.

Recall that $\overline{T}(v) = 0$ if and only if $\Phi(v) \leq (1 - \beta)V$, where $\Phi(v) = v(1 - F(v))$. As noted in the proof of Theorem 1, Assumption 4 ensures that Φ has a finite maximum Φ^* on $\mathbb{R}_+$. Hence whenever $V > \Phi^*/(1 - \beta)$, we have $\overline{T}(v) = 0$ for all $v \geq 0$. We maintain this lower bound on V going forward.

Note that $\overline{v}_0$ is continuous and increasing in V, and for V sufficiently large that $F(V) > \beta$, we have $\overline{v}_0 F(\overline{v}_0) = V \cdot \beta$. We maintain this lower bound on V going forward. An implication of this identity is that $\overline{v}_0 \to \infty$ as $V \to \infty$ and therefore $\overline{v}_0/V = \beta/F(\overline{v}_0) \to \beta$ as $V \to \infty$.

For $v \in [0, \overline{v}_0]$, define

$$\Gamma(v, V) \equiv \frac{1 - F(v)}{V} \cdot \left[\rho_0 \left(1 - \beta + \frac{vF(v)}{V} \right) - \rho_0 \left(1 - \beta - \frac{v(1 - F(v))}{V} \right) \right].$$

When $v \leq \overline{v}_0$, the inequality $T^*(v) > 0$ is equivalent to $R > 1/\Gamma(v, V)$. Note that at $v = \overline{v}_0$ we have

$$\Gamma(\overline{v}_0, V) = \frac{1 - F(\overline{v}_0)}{V} \cdot \left[\rho_0 (1) - \rho_0 \left(1 - \frac{\overline{v}_0}{V} \right) \right].$$

Under Assumption 3, ρ_0 is continuously differentiable in a neighborhood of 1 and $1 - \beta$. Then since $1 - \overline{v}/V$ approaches $1 - \beta$ as $V \to \infty$, it follows that $\Gamma(v, V)$ is differentiable in a neighborhood of $\overline{v}$ for sufficiently large V. We maintain this lower bound on V going forward.

The following lemma completes the proof by establishing that $\overline{\Gamma}(V) \equiv \max_{v\in[0,\overline{v}_0]} \Gamma(v,V) > \Gamma(\overline{v}_0, V)$ for V sufficiently large, in which case $T^*(\overline{v}_0) = 0$ and $\max_{v\in[0,\overline{v}_0]} T^*(v) > 0$ for R in the non-empty interval $(1/\overline{\Gamma}(V), 1/\Gamma(\overline{v}_0, V))$.

Lemma A.4. $\frac{\partial \Gamma}{\partial v}(\overline{v}_0, V) < 0$ *for* V *sufficiently large.*

Proof. Differentiating Γ with respect to v and evaluating at $v = \overline{v}_0$ yields

$$\frac{\partial \Gamma}{\partial v}(\overline{v}_0, V) = \frac{f(\overline{v})}{V} \cdot (\gamma(V) - \Delta\rho_0(V)),$$

where

$$\gamma(V) \equiv \frac{1 - F(\overline{v}_0)}{V} \cdot \left[\rho_0'(1) \cdot \left(\overline{v}_0 + \frac{F(\overline{v}_0)}{f(\overline{v}_0)}\right) - \rho_0'\left(1 - \frac{\overline{v}_0}{V}\right) \cdot \left(\overline{v}_0 - \frac{1 - F(\overline{v}_0)}{f(\overline{V}_0)}\right)\right]$$

and

$$\Delta\rho_0(V) \equiv \rho_0(1) - \rho_0\left(1 - \frac{\overline{v}_0}{V}\right).$$

It is sufficient to establish that $\gamma(V) < \Delta\rho_0(V)$ when V is large. Note that $\Delta\rho_0(V) \to \rho_0(1) - \rho_0(1-\beta) > 0$ as $V \to \infty$. Meanwhile, $\gamma(V)$ may be equivalently written

$$\gamma(V) = \frac{\overline{v}_0}{V} \cdot \left[\rho_0'(1) \cdot \left(1 - F(\overline{v}_0) + \frac{F(\overline{v}_0)}{\overline{v}_0 h(\overline{v}_0)}\right) - \rho_0'\left(1 - \frac{\overline{v}_0}{V}\right) \cdot \left(1 - \frac{1}{\overline{v}_0 h(\overline{v}_0)}\right)(1 - F(\overline{v}_0))\right],$$

where $h(v) = f(v)/(1-F(v))$ is the hazard rate of F. Under Assumption 4, $vh(v) \to \infty$ as $v \to \infty$. Hence $\gamma(V) \to 0$ as $V \to \infty$, establishing the result. ☐

A.10 Proof of Lemma 6

Throughout this proof, we will write β explicitly as an argument of quantities which depend on it. Define

$$\underline{\beta}_0 \equiv \frac{vF(v)}{V}, \quad \overline{\beta}_0 \equiv 1 - \frac{v(1-F(v))}{V}.$$

These prize endowments are interior whenever $v \le V$. Additionally, $\overline{\beta}_0 - \underline{\beta}_0 = 1 - v/V$, so that $\overline{\beta}_0 > \underline{\beta}_0$ when $v < V$. Recall from the proof of Lemma 5 that

$$\underline{T}(\beta) = \max\left\{0, v - V, v - \frac{V \cdot \beta}{F(v)}\right\}.$$

Then when $v < V$, the minimal feasible bonus $\underline{T}(\beta)$ vanishes if and only if $\beta \ge \underline{\beta}_0$. Meanwhile, the threshold bonus

$$\overline{T}(\beta) = \max\left\{0, v - \frac{(1-\beta)V}{F(v)}\right\}$$

defined in the proof of Lemma 5 vanishes if and only if $\beta \leq \overline{\beta}_0$.

When $\beta \in (\underline{\beta}_0, \overline{\beta}_0)$, the cost-minimizing match standards $q_A^{**}(T;\beta)$ and $q_B^{**}(T;\beta)$ corresponding to any feasible bonus $T \in [0, v]$ (constructed in the proof of Lemma 3) are interior. Hence in particular the optimal match standards $q_A^*(\beta)$ and $q_B^*(\beta)$ must be interior.

Meanwhile, for $\beta \leq \underline{\beta}_0$ we have $\overline{T}(\beta) = 0$ and $\underline{T}(\beta) > 0$. In this regime, $q_A^{**}(T;\beta) > 0$ for every feasible $T \in [\underline{T}(\beta), v]$. However, q_B^{**} is decreasing in T and $q_B^{**}(\underline{T}(\beta);\beta) = 1$. Hence $q_B^*(\beta) < 1$ if and only if $T^*(\beta) > \underline{T}(\beta)$, i.e., if

$$\frac{d\Pi^*}{dT}(\underline{T}(\beta);\beta) > 0.$$

Using the expression for this derivative derived in the proof of Proposition 2, this condition is equivalently

$$\rho(1) - \rho\left(1 - \frac{\beta}{F(v)}\right) > \frac{V}{1 - F(v)}.$$

The left-hand side of this condition is increasing in β and vanishes at $\beta = 0$. So define

$$\underline{\beta} \equiv \max\left\{\beta \leq \underline{\beta}_0 \ : \ \rho(1) - \rho\left(1 - \frac{\beta}{F(v)}\right) \leq \frac{V}{1 - F(v)}\right\}.$$

Then $q_B^*(\beta) < 1$ if and only if $\beta > \underline{\beta}$.

Finally, for $\beta \geq \overline{\beta}_0$ we have $\underline{T}(\beta) = 0$ and $\overline{T}(\beta) > 0$. In this regime, $q_B^{**}(T;\beta) < 1$ for every feasible $T \in [0, v]$. However, $q_A^{**}(T;\beta) > 0$ if and only if $T > \overline{T}(\beta)$. Hence $q_A^*(\beta) > 0$ if and only if $T^*(\beta) > \underline{T}(\beta)$, i.e., if

$$\frac{d\Pi^*}{dT+}(\overline{T}(\beta);\beta) > 0.$$

Using the expression for this derivative derived in the proof of Lemma 5, this condition is equivalently

$$\rho\left(\frac{1-\beta}{1 - F(v)}\right) - \rho(0) > \frac{V}{1 - F(v)}.$$

The left-hand side of this condition is decreasing in β and vanishes at $\beta = 1$. So define

$$\overline{\beta} \equiv \min\left\{\beta \geq \overline{\beta}_0 \ : \ \rho\left(\frac{1-\beta}{1 - F(v)}\right) - \rho(0) \leq \frac{V}{1 - F(v)}\right\}.$$

Then $q_A^*(\beta) > 0$ if and only if $\beta < \overline{\beta}$.

A.11 Proof of Proposition 3

Throughout this proof, we will write β explicitly as an argument of quantities which depend on it. Lemma 6 ensures that $q_B^*(\beta) < 1$ and $q_A^*(\beta) > 0$ whenever $\beta \in (\underline{\beta}, \overline{\beta})$. The former inequality implies $T^*(\beta) > \underline{T}(\beta)$ whenever $\underline{T}(\beta) > 0$, while the latter implies $T^*(\beta) > \overline{T}(\beta)$ whenever $\overline{T}(\beta) > 0$, where $\overline{T}$ is as defined in the proof of Lemma 5. As established in the proof of Proposition 2, under Assumption 1 the optimal bonus $T^*(\beta)$ is therefore

$$T^*(\beta) = \max\left\{v - \frac{V^2/R}{1 - F(v)}, 0\right\}$$

for all $\beta \in (\underline{\beta}, \overline{\beta})$, which does not vary with β. Since additionally the objective $\Pi^*(T; \beta)$ and the constraint set $[\underline{T}(\beta), v]$ are continuous in β, the optimal bonus $T^*(\beta)$ is continuous in β, so that $T^*(\beta)$ must also be independent of β on the larger set $[\underline{\beta}, \overline{\beta}]$. The targeting constraint then implies that $\Delta q^*(\beta) = (v - T^*(\beta))/V$ is likewise independent of β, and since $\Delta\rho^*(\beta) = R \cdot \Delta q^*(\beta)$ under Assumption 1, so is $\Delta\rho^*$.

A.12 Proof of Theorem 3

Throughout this proof, we will write β explicitly as an argument of quantities which depend on it. First note that $\Pi^*(T; \beta)$ is continuous in (T, β) and $\underline{T}(\beta)$ is continuous in β. As a result $T^*(\beta)$ is continuous in β.

Next, as discussed in the proof of Proposition 3, for $\beta \in (\underline{\beta}, \overline{\beta})$ the inequalities $T^*(\beta) > \overline{T}(\beta)$ and $T^*(\beta) > \underline{T}(\beta)$ must hold whenever, respectively, $\overline{T}(\beta) > 0$ and $\underline{T}(\beta) > 0$. Then given the expressions for $d\Pi^*/dT$ derived in the proof of Lemma 5, for $\beta \in (\underline{\beta}, \overline{\beta})$ the optimal bonus $T^*(\beta)$ must satisfy the first-order condition

$$\frac{d\Pi^*}{dT}(T; \beta) = F(v) \cdot \left((1 - F(v)) \cdot \Delta\rho_0(T; \beta) \cdot \frac{R}{V} - 1\right) \leq 0,$$

with equality if $T^*(\beta) > 0$, where

$$\Delta\rho_0(T; \beta) \equiv \rho_0\left(1 - \beta + F(v)\frac{v - T}{V}\right) - \rho_0\left(1 - \beta - (1 - F(v))\frac{v - T}{V}\right).$$

Since this expression for $d\Pi^*/dT$ is continuous in both T and β and $T^*(\beta)$ is continuous in β, it follows that this first-order condition must additionally hold at $\beta = \underline{\beta}, \overline{\beta}$.

In particular, $T^*(\beta) = 0$ for $\beta \in [\underline{\beta}, \overline{\beta}]$ if and only if $\beta \in [\underline{\beta}_0, \overline{\beta}_0]$ (with these quantities as defined in the proof of Lemma 6) and

$$R \leq \frac{V}{(1 - F(v))\Gamma(\beta)},$$

where $\Gamma(\beta) \equiv \Delta\rho_0(0; \beta)$.

Lemma A.5. Γ *is single-troughed on* $[\underline{\beta}_0, \overline{\beta}_0]$.

Proof. Under Assumption 5, Γ is differentiable and

$$\Gamma'(\beta) = \rho_0'\left(1 - \beta - \frac{v(1 - F(v))}{V}\right) - \rho_0'\left(1 - \beta + \frac{vF(v)}{V}\right).$$

At the endpoints $\underline{\beta}_0$ and $\overline{\beta}_0$ we have

$$\Gamma'(\underline{\beta}_0) = \rho_0'\left(1 - \frac{v}{V}\right) - \rho_0'(1)$$

and

$$\Gamma'(\overline{\beta}_0) = \rho_0'(0) - \rho_0'\left(1 - \frac{v}{V}\right).$$

The assumption that ρ_0' is single-troughed with $\rho_0'(0) = \rho_0'(1)$ implies that $\rho_0'(0), \rho_0'(1) > \rho_0'(q)$ for every $q \in (0, 1)$. The fact that $v \in (0, V)$ therefore implies $\Gamma'(\underline{\beta}_0) < 0 < \Gamma'(\overline{\beta}_0)$.

Continuity of ρ_0' implies that Γ' is also continuous, so there exists some $\beta_0^* \in (\underline{\beta}_0, \overline{\beta}_0)$ such that $\Gamma'(\beta_0^*) = 0$. To complete the proof, we must show that this crossing point is unique. Letting q^* be the unique trough of ρ_0', we must have

$$1 - \beta_0^* - \frac{v(1 - F(v)}{V} < q^* < 1 - \beta_0^* + \frac{v(F(v)}{V},$$

or else ρ_0' would be strictly monotone between these two endpoints, violating $\Gamma(\beta_0^*) = 0$. But then $\rho_0'(1 - \beta - v(1 - F(v))/V)$ is locally increasing in β around β_0^*, while $\rho_0'(1 - \beta + vF(v)/V)$ is locally decreasing, implying that Γ' is locally increasing in β near β_0^*. Since this argument applies to any zero of Γ', it has at most one zero, establishing uniqueness. □

Let $\beta_0^* \in (\underline{\beta}_0, \overline{\beta}_0)$ be the unique minimizer of Γ on $[\underline{\beta}_0, \overline{\beta}_0]$, and define

$$\overline{R} \equiv \frac{V}{(1 - F(v))\Gamma(\beta_0^*)}.$$

Then if $R > \overline{R}$, the optimal bonus is positive for all $\beta \in [\underline{\beta}, \overline{\beta}]$, and if $R = \overline{R}$ then the optimal bonus is positive except at $\beta = \beta_0^*$. Let $\beta_L = \beta_H \in (\underline{\beta}, \overline{\beta}$ in either case, with their values to be defined later. Otherwise, $T^*(\beta) = 0$ on a non-degenerate interval $[\beta_L, \beta_H]$ with $\underline{\beta}_0 \leq \beta_L < \beta_0^* < \beta_H \leq \overline{\beta}_0$. Additionally define

$$\underline{R} \equiv \frac{V}{(1 - F(v)) \min\{\Gamma(\underline{\beta}_0), \Gamma(\overline{\beta}_0)\}}.$$

If $R > \underline{R}$, then $\beta_L > \underline{\beta}_0$ and $\beta_H < \overline{\beta}_0$. The fact that Γ is single-troughed implies $\Gamma(\beta_0) < \min\{\Gamma(\underline{\beta}_0), \Gamma(\overline{\beta}_0)\}$, so that $\overline{R} > \underline{R}$ and there exists a non-degenerate interval of parameters R for which $\underline{\beta}_0 < \beta_L < \beta_H < \overline{\beta}_0$, and hence also $\underline{\beta} < \beta_L < \beta_H < \overline{\beta}$.

Suppose that the interval (β_L, β_H) is non-empty. By definition, $T^*(\beta) = 0$ on this interval. Clearly then $\Delta q^*(\beta) = (v - T^*(\beta))/V$ is also independent of β on the interval. Additionally,

$$\Delta\rho^*(\beta) = R \cdot \Delta\rho_0(T^*(\beta); 0) = R \cdot \Delta\rho_0(0; \beta) = R \cdot \Gamma(\beta)$$

for $\beta \in (\beta_L, \beta_H)$. Then since (β_L, β_H) contains β_0^*, the optimal absolute prize priority $\Delta\rho^*$ is single-troughed on (β_L, β_H).

Next, suppose that $\beta_L > \underline{\beta}$. Then on $[\underline{\beta}, \beta_L)$ we have $T^*(\beta) > 0$, in which case it must satisfy the first-order condition

$$(1 - F(v)) \cdot \Delta\rho_0(T^*(\beta); \beta) \cdot \frac{R}{V} = 1.$$

Since β does not appear in this condition except indirectly through ρ_0, it follows that $\rho_0(T^*(\beta); \beta)$ is independent of β and therefore so is $\Delta\rho^*(\beta) = R \cdot \rho_0(T^*(\beta); \beta)$. An identical argument holds on $(\beta_H, \overline{\beta}]$ whenever $\beta_H < \overline{\beta}$.

We complete the proof by signing the slope of $T^*(\beta)$ outside (β_L, β_H), with the corresponding comparative statics for $\Delta q^*(\beta)$ then following from the targeting equation $\Delta q^*(\beta) = (v - T^*(\beta))/V$. If $\beta_L = \underline{\beta}$ or $\beta_H = \overline{\beta}$, then there is nothing to prove in that region, so for what follows we assume that $\beta_L > \underline{\beta}$ and $\beta_H < \overline{\beta}$. Note that $\Delta\rho_0(T; \beta)$ is continuously differentiable with respect to both T and β, and its derivative with respect to T is negative everywhere. The implicit function theorem therefore implies that $T^*(\beta)$ is differentiable on $(\underline{\beta}, \beta_L)$ and $(\beta_H, \overline{\beta})$ and

$$\frac{dT^*}{d\beta}(\beta) = V \cdot \frac{\rho_0'\left(q_A^*(\beta)\right) - \rho_0'\left(q_B^*(\beta)\right)}{F(v) \cdot \rho_0'\left(q_B^*(\beta)\right) + (1 - F(v)) \cdot \rho_0'\left(q_A^*(\beta)\right)},$$

where

$$q_A^*(\beta) = 1 - \beta - (1 - F(v))\frac{v - T^*(\beta)}{V}, \quad q_B^*(\beta) = 1 - \beta + F(v)\frac{v - T^*(\beta)}{V}.$$

It additionally possesses one-sided derivatives at the endpoints of these intervals, with the same formula applying.

Note that any stationary point of T^* on $[\underline{\beta}, \beta_L]$ or $[\beta_H, \overline{\beta}]$ must satisfy

$$\rho_0'\left(q_A^*(\beta)\right) = \rho_0'\left(q_B^*(\beta)\right).$$

Additionally, at a stationary point $q_A^*(\beta)$ and $q_B^*(\beta)$ are both locally decreasing in β, so that $dT^*/d\beta > 0$ just above the stationary point while $dT^*/d\beta < 0$ just below it. In other words, any stationary point of T^* on $[\underline{\beta}, \beta_L]$ or $[\beta_H, \overline{\beta}]$ must be a local minimum.

Additionally, at $\beta = \underline{\beta}$ we have $q_B^*(\underline{\beta}) = 1$ and $q_A^*(\underline{\beta}) \in (0, q_B^*(\underline{\beta}))$. Then under Assumption 5 it must be that $\rho_0'\left(q_A^*(\underline{\beta})\right) < \rho_0'\left(q_B^*(\underline{\beta})\right)$ and therefore $\frac{dT^*}{d\beta+}(\underline{\beta}) < 0$. Similarly, at $\beta = \overline{\beta}$ we have $q_B^*(\overline{\beta}) = 0$ and $q_A^*(\overline{\beta}) \in (q_B^*(\overline{\beta}), 1)$, in which case $\rho_0'\left(q_A^*(\overline{\beta})\right) > \rho_0'\left(q_B^*(\overline{\beta})\right)$ and therefore $\frac{dT^*}{d\beta-}(\underline{\beta}) > 0$.

Consider first the case that $\beta_L = \beta_H$. Then T^* is differentiable everywhere on $[\underline{\beta}, \overline{\beta}]$, and in light of the results of the previous paragraphs it must have at least one stationary point $\beta^* \in (\underline{\beta}, \overline{\beta})$. Since any stationary point is a local minimum, this stationary point is therefore unique. Letting $\beta_L = \beta_H = \beta^*$ yields the claimed behavior of T^*.

Now consider the case that $\beta_L < \beta_H$. Then $T^*(\beta_L) = T^*(\beta_H) = 0$, and so it must be that $\frac{dT^*}{d\beta-}(\beta_L) \leq 0$ and $\frac{dT^*}{d\beta+}(\beta_H) \geq 0$. But since any stationary point of T^* on $[\underline{\beta}, \beta_L]$ must be a local minimum and therefore unique, and since $\frac{dT^*}{d\beta-}(\underline{\beta}) < 0$, it follows that $\frac{dT^*}{d\beta}(\beta) < 0$ for all $\beta \in [\underline{\beta}, \beta_L)$. In other words, T^* is decreasing on $[\underline{\beta}, \beta_L]$. A very similar argument implies that T^* is increasing on $[\beta_H, \overline{\beta}]$.

A.13 Proof of Lemma 7

Recall from the proof of Lemma 3 that

$$q_B^* = \max\{\Delta q^*, 1 - \beta + F(v)\Delta q^*\}.$$

If $q_B^* < 1$, then $\Delta q^* < 1$ and $1 - \beta + F(v)\Delta q^* < 1$, with the latter bound equivalent to $\Delta q^* < \beta/F(v)$. Thus $\Delta q^* < \min\{\beta/F(v), 1\}$, and so prize incentives are not exhausted.

Conversely, if $q_B^* = 1$, then one of two cases is possible. If $\Delta q^* \geq (1 - \beta)/(1 - F(v))$, then $\Delta q^* = q_B^* = 1$, in which case the hypothesized bound on Δq^* implies that $\beta \geq F(v)$. In this case $\Delta q^* = \min\{1, \beta/F(v)\}$ and incentives are exhausted. The remaining possibility is that $\Delta q^* < (1 - \beta)/(1 - F(v))$, in which case $\Delta q^* = \beta/F(v)$ and the hypothesized bound on Δq^* implies that $F(v) > \beta$. Thus in this case too $\Delta q^* = \min\{1, \beta/F(v)$ and incentives are exhausted.

A.14 Proof of Proposition 4

Using the notation and concepts developed in the proof of Lemma 5, when $v > V$ and β is sufficiently close to 1 we have $\beta > F(v)$ and $\underline{T} = v - V$ and $\overline{T} > \underline{T}$. In this regime, $T^* = \underline{T}$ if and only if the first-order condition

$$\frac{d\Pi^*}{dT}(\underline{T}) = F(v) \cdot \left(\frac{1 - F(v)}{F(v)} \cdot \rho_0(1) \cdot \frac{R}{V} - 1 \right) \leq 0$$

holds, which is true for R sufficiently small. Whenever $\beta > F(v)$ and $T^* = \underline{T} = v - V$, we have $\Delta q^* = 1 = \min\{1, \beta/F(v)\}$. Thus $\mathcal{C}^*$ exhausts prize incentives in this regime.

A.15 Proof of Lemma 8

The set of feasible schemes is non-empty if and only if $\underline{T} \leq M/F(v)$. This inequality trivially holds if $v \leq \overline{v}_0$. Otherwise, using the expression for $\underline{T}$ derived in the proof of Lemma 5, it holds whenever

$$\max \left\{ v - V, v - \frac{V \cdot \beta}{F(v)} \right\} \leq \frac{M}{F(v)}.$$

The left-hand side of this inequality is continuous and increasing in v, vanishes at $v = \overline{v}_0$, and grows without bound as $v \to \infty$. Meanwhile the right-hand size is continuous and decreasing in v and is positive at $v = \overline{v}_0$. Hence there exists a unique $v^* \in (\overline{v}_0, \infty)$ at which this inequality holds with equality, and it is satisfied if and only if $v \in [\overline{v}_0, v^*]$.

Given M and $v^*(M)$, for every $M' > M$ we have

$$\max \left\{ v^*(M) - V, v^*(M) - \frac{V \cdot \beta}{F(v^*(M))} \right\} = \frac{M}{F(v^*(M))} < \frac{M'}{F(v^*(M))},$$

from which it follows that $v^*(M) < v^*(M')$. Additionally, when $M = 0$ we have

$$\max \left\{ v - V, v - \frac{V \cdot \beta}{F(v)} \right\} > 0 = \frac{M}{F(v)}$$

for every $v > \overline{v}_0$, so that $v^* = M$ in this limit. In the other direction, given any v we have

$$\max \left\{ v - V, v - \frac{V \cdot \beta}{F(v)} \right\} < \frac{M}{F(v)}$$

for M sufficiently large, so that $\lim_{M \to \infty} v^*(M) = \infty$.

A.16 Proof of Lemma 9

Let $\overline{M} \equiv \max_{v \in [0,\overline{v}]} F(v)T^*(v)$, where $T^*(v)$ is the optimal unregulated bonus under target v. If $M \geq \overline{M}$, the unconstrained optimal bonus is no larger than $M/F(v)$ for all $v \leq \overline{v}$. Meanwhile, for $v > v$ Lemma 5 ensures that $T^*(v) = \underline{T}(v)$. Since $\underline{T}(v) < M/F(v)$ for all $v < v^*$, it follows that $T^*(v) \leq M/F(v)$ for all $v \in (\overline{v}, v^*]$.

A.17 Proof of Proposition 5

Throughout this proof, we will write v explicitly as an argument of quantities which depend on it. Additionally, we will freely employ concepts and notation developed in the proof of Lemma 5.

We first show that the financial constraint does not bind on $[0, v^*)$ for sufficiently small R. For this it is sufficient that $T^*(v) = \underline{T}(v)$ for all $v \leq v^*$, since $\underline{T}(v) \leq M/F(v)$ for all $v \leq v^*$. The optimal bonus equals $\underline{T}(v)$ if and only if the first-order condition

$$\frac{d\Pi^*}{dT}(\underline{T}(v); v) = F(v) \cdot \left\{(1 - F(v)) \cdot \Delta\bar{\rho}_0(\underline{T}(v); v) \cdot \frac{R}{V} - 1\right\} \leq 0$$

holds, where $\Delta\bar{\rho}_0(T; v) \equiv \Delta\bar{\rho}(T; v)/R$ is independent of R. Note that $\Delta\bar{\rho}_0(T; v) \leq \rho_0(1)/F(v)$ for every v and $T \in [\underline{T}(v), v]$, and so it is sufficient that

$$\frac{1 - F(v)}{F(v)} \cdot \rho_0(1) \cdot \frac{R}{V} \leq 1$$

for all $v < v^*$. Since the left-hand side is decreasing in v, this condition holds whenever

$$R \leq \frac{F(0)}{1 - F(0)} \cdot \frac{V}{\rho_0(1)}.$$

We now show that the financial constraint binds close to v^* for sufficiently large R. Suppose that $T^*(v) > M/F(v)$ for some $v \leq \overline{v}^*$, where $T^*(v)$ is the optimal bonus with no financial constraint. Then $\frac{d\Pi^*}{dT-}(M/F(v); v) > 0$, so that the optimal bonus under a financial constraint equals $M/F(v)$. Since $T^*(v)$ and $M/F(v)$ are continuous in v, it is therefore sufficient for the result to show that $T^*(v^*) > M/F(v^*)$ when R is sufficiently large.

Recall from the proof of Lemma 8 that v^* satisfies the identity $\underline{T}(v^*) = M/F(v^*)$. Hence we need to show that $T^*(v^*) > \underline{T}(v^*)$ when R is large. This inequality is satisfied if and only if the first-order condition

$$\frac{d\Pi^*}{dT}(\underline{T}(v^*); v^*) = F(v^*) \cdot \left\{(1 - F(v^*)) \cdot \Delta\bar{\rho}_0(\underline{T}(v^*); v^*) \cdot \frac{R}{V} - 1\right\} > 0$$

holds. This inequality holds whenever R is sufficiently large.