Scarcity of Ideas and Optimal Prizes in Innovation Contests*

Nisvan Erkal† and Jun Xiao‡

December 14, 2019

Abstract

This paper studies the relationship between optimal prizes and scarcity of ideas in innovation contests. We consider a model of innovation where both ideas and effort are integral parts of the innovation process. Contest participants are privately informed about the quality of their ideas. We study how a contest designer's choice–the profit maximizing prize–should vary with the difficulty of the innovation challenge, which is represented by the distribution of idea quality. We introduce a new stochastic order to rank difficulty of challenges according to scarcity of high-quality ideas, and find that scarcity of high-quality ideas results in higher optimal prizes if and only if the benefit from a marginal improvement in the new technology's performance is sufficiently low.

JEL classification: D44, D82, O31

Keywords: innovation, idea distribution, contests, comparative statics, stochastic dominance

*We thank Suren Basov, Francis Bloch, Isa Hafalir, Vijay Krishna, Igor Letina, Priscilla Man, Claudio Mezzetti, Benny Moldovanu, Sergio Parreiras, Leo Simon, Satoru Takahashi, participants in 26th International Conference on Game Theory, 11th World Congress of Econometric Society, 18th Annual Society for the Advancement of Economic Theory Conference, and seminar participants at Deakin University, National University of Australia, National University of Singapore, Monash University, Penn State University, Queensland University of Technology, University of New South Wales, University of Melbourne, and University of Queensland for their comments and discussion. Jingong Huang and Eado Varon have provided excellent research assistance. Financial support from the Australian Research Council Discovery grant DP170103374 is greatly appreciated.

†Department of Economics, University of Melbourne. Email: n.erkal@unimelb.edu.au.

‡Department of Economics, University of Melbourne. Email: jun.xiao@unimelb.edu.au.

1 Introduction

The process of innovation starts with a problem. What turns a problem into a world-changing discovery is an insight or an idea. Hence, an idea represents an investment opportunity and can be turned into an innovation by exerting effort.

Contests have a long history in the procurement of innovation. In the case of an innovation contest, the problem to be solved is defined by the contest designer and is public information. Historically, innovation contests were predominantly used by governments seeking solutions to major innovation problems faced by their countries.[1] More recently, designing contests to solve innovation problems has been growing in popularity in the private sector also. Even major industry players, such as BMW (Füller et al., 2006), Cisco, IBM (Bjelland and Wood, 2008), Philips, and Qualcomm, utilize contests to enhance their innovative capacity.[2]

In this paper, we study the design of optimal prizes in innovation contests. There exists a large variation in the prizes awarded in innovation contests. For example, the average prize awarded by XPRIZE, an organization running public innovation contests, was \$4.5 million in 1996-2014 and \$8.4 million in 2014-2018. As another example, Netflix awarded \$1 million in 2009 to the computer algorithm that best improves its recommendation system. In 2013, the company awarded a much lower prize of \$10,000 to those who most improve its cloud computing services. One explanation for this variation may be that some problems are more difficult than others, and therefore higher prizes are awarded to solve them. However, is it always optimal to incentivize more difficult challenges with higher prizes?

Our starting point is a model of innovation where the key ingredients are ideas and effort. To solve a problem, a solver (i.e., potential innovator) must first have an idea and then the incentive to invest in that idea. We assume that ideas can be ranked in terms of their quality. Different solvers privately receive ideas with different qualities for the same problem.[3] They then decide whether to invest and how much to invest in solving the problem. The best solution wins the contest and receives the prize provided it is above a performance requirement.

[1]For example, the British Parliament offered a prize of £20,000 in 1714 for a method of determining longitude at sea. In 1795, the French military offered a cash prize of 12,000 francs for a new method to preserve food (which resulted in the development of canning).

[2]Innovation contests are also sometimes used by venture capitalists in the allocation of funds. See for example QPrize, which is Qualcomm Ventures' seed investment competition.

[3]Kornish and Ulrich (2014) show empirically that the quality of ideas matters in determining success. They consider 'raw ideas,' which are the opportunities conceived at the outset of an innovation process, and investigate the importance of having a good idea (as opposed to resources) in determining success.

Different problems in our model are represented by different idea distributions. Some problems are easy and then many innovators are likely to have high-quality ideas to start with. In this case, competition is more likely to take place along the effort dimension. As the problem gets more difficult, the probability of having a high-quality idea decreases. In this case, solvers are more likely to compete along the idea dimension, i.e., having a good idea will give a solver a strong advantage.

Scarcity is a fundamental concept in economics and is considered to be one of the main determinants of an object's value. Does this mean that the winner in a contest where high-quality ideas are scarcer should receive a higher prize? In other words, to what extent should we expect the prize to reflect the scarcity of high-quality ideas?[4] We show that the answer to this seemingly intuitive question is not straightforward.

To explore how the optimal prize changes as the scarcity of high-quality ideas changes, we introduce a new order of stochastic dominance to capture the notion of scarcity. An innovator has a scarcer idea if the probability that someone else has a better idea than him is smaller. Since this probability will depend on the size of the solver base, the scarcity order also depends on the solver population size.

Our results uncover that the relationship between scarcity of high-quality ideas and the optimal prize critically depends on the market value of the innovation, specifically on the marginal value of solution performance for the seeker (i.e., contest designer). Consider two examples. The Wolfskehl Prize awards DM 100,000 to the first person to rediscover the proof of Fermat's Last Theorem. As long as a proof is correct, it serves the purpose of validating the theorem. Hence, the benefits of a proof depends only on whether it meets a minimal requirement. Compare this with the Netflix Prize, which awarded $1 million in 2009 to the best algorithm to predict user ratings for films. In this case, the minimum requirement was for the new algorithm to outperform the existing one. Otherwise, there was no benefit to Netflix. Beyond meeting this minimum requirement, the more an algorithm improved the existing one, the more valuable it was. The winning algorithm

[4] The prize in an innovation contest is similar to a patent. Through the patentability requirements, society selects which innovations should receive a patent (i.e., prize). For example, in US patent law, an innovation should satisfy the nonobviousness requirement. According to this requirement, an invention is considered nonobvious if someone with ordinary skill or training in the relevant field could not easily make the invention based on prior art. In European patent law, the same idea is captured by the inventive step requirement, which states that an invention should be sufficiently inventive in order to be patented. In both cases, an invention is considered to be patentable if it meets a legally defined "scarcity" requirement.

bested Netflix's own algorithm for predicting ratings by 10.06%.[5]

We show that if the market value is not very sensitive to the quality of the solution (i.e., as long as the solution meets the minimal quality, additional quality does not have much impact on the market value of the innovation), then the goal of the seeker is to maximize the probability of having an idea that meets the minimal quality requirement. In this case, as the scarcity of high-quality ideas increases, the seeker compensates by increasing the prize level. Hence, the optimal prize is increasing in the scarcity of high-quality ideas. However, if the marginal value of solution performance for the seeker is sufficiently high, then the goal of the seeker is to obtain as good a solution as possible. In this case, the optimal prize is decreasing in the scarcity of high-quality ideas. Intuitively, when the seeker cares highly about the solution performance, the expected return from a pool of solvers decreases as the scarcity of high-quality ideas increases. The seeker is willing to invest more in a contest where the likelihood of a high-performance solution is higher.

We consider two extensions of our model. The results generalize in a straightforward way to the case where the seeker has a nonlinear benefit function. We also consider a variation of the model where the seeker sets the minimum performance requirement in addition to the prize.

In general, our results imply that contest designers, while adjusting prize levels with the difficulty of challenges, should pay attention to the how much they will benefit from a marginal increase in performance. Although our paper is couched in the language of innovation contests, our analysis applies to a wider range of scenarios. Our results have implications for any contest environment with private types where the participants have to exert effort and the contest designer cares about the best performance only. In other contexts, the reason for the private information may be past experience, access to different resources, genetic make-up, etc.

The remainder of the paper is organized as follows. Section 2 discusses the related literature. Section 3 introduces a model of contests and Section 4 characterizes the equilibrium strategies and the optimal prize. Section 5 presents our main result on how the optimal prize changes with the scarcity of ideas. The proof of the main result is presented in Section 6. Section 7 explores extensions of our benchmark model to the cases of nonlinear benefit functions and endogenous minimum requirements. Finally,

[5]The winning team BellKor's Pragmatic Chaos improved the predictions by 10.06% on the test data set, which Netflix used to determine the final winner. The contest structure for the Netflix Prize was different from the one we consider in this paper. It was dynamic, where the first submission triggered a deadline for further submissions. Nevertheless, the payoff function of the seeker is the same as to ours.

Section 8 concludes. All proofs, except for the proof of our main result (Proposition 2), are relegated to the appendix.

2 Related Literature

Our paper contributes to a large literature on contests with incomplete information. See, e.g., Moldovanu and Sela (2001), Chawla et al. (2015), Liu et al. (2018) and Olszewski and Siegel (2018). There also exists a growing literature specifically on innovation contests. See, e.g., Taylor (1995), Fullerton and McAfee (1999), Fullerton et al. (2002), Che and Gale (2003), Terwiesch and Xu (2008) and Korpeoglu and Cho (2018).[6]

Our paper has a different focus from all these papers because we are interested in the question of how the optimal prize changes as the distribution of ideas changes. An important feature of our paper is that each participant privately has access to a solution idea that they can invest in. This allows us to focus on the importance of ideas and idea quality (i.e., creativity) in the innovation process.

The ideas in our model represent different solution approaches which can be ranked in terms of quality. The possibility of different approaches is also considered in Ganuza and Hauk (2006), Erat and Krishnan (2012) and Letina and Schmutzler (2019). In these papers, participants choose between a variety of approaches before exerting effort. However, as different from our model, all potential approaches are available to all solvers at the same time. In contrast, each solver has access to one and only one (different) solution approach in our model.

Another strand of literature that is related to our paper are the studies on monotone comparative statics. These studies investigate how the solutions to a maximization problem change as the parameters of the problem change. Topkis (1978) and Milgrom and Shannon (1994) consider the question in non-stochastic environments while Athey (2002) consider it in stochastic environments. Quah (2007) studies how the solution to a maximization problem changes as the constraint set changes. In this paper, we ask how the contest designer's optimal choice changes as the idea distribution changes. As the idea distribution changes, both the objective function (the seeker's expected profit), and the participants' incentive and participation constraints change.

Our stochastic order is a modified version of the monotone likelihood ratio property used in Athey (2002). It is different from the stochastic orders used in the literature

[6]Some papers in the literature consider hybrid systems. For example, Fu et al. (2012) analyze how a fixed budget should be allocated between subsidies and prizes in order to motivate innovation. Galasso et al. (2018) study environments where patent rights and cash rewards are complements.

to investigate how the distribution of types affect properties of equilibrium outcomes. For example, Maskin and Riley (2000) show that dominance in terms of reverse hazard rate implies more aggressive bidding in first price auctions. Pesendorfer (2000) shows similar results in procurement auctions using hazard rate dominance. Hoppe et al. (2009) study the role of costly signaling in matching markets with privately informed agents. They use a variation of second order stochastic dominance to consider how increased heterogeneity affects the matching outcome.[7] Note that these papers in the literature study how equilibrium behavior changes with the distribution of private information, while we study how the mechanism designer's choice changes with the distribution of private information.

3 Model

We study a contest type which is commonly observed in innovation environments. Consider a seeker (e.g., a pharmaceutical company) who is searching for a solution (e.g., a vaccine for a new disease) and sponsors an innovation contest with a monetary prize of value $v \in [0, \bar{v}]$, where $\bar{v}$ is a finite upper bound for the prize. The upper bound ensures that the optimal prize is finite. To avoid the uninteresting case in which the optimal prize is always $\bar{v}$, we also assume the bound is not too small: $\bar{v} > 1$.[8] The assumption that the prize is a fixed amount and does not vary with solution performance is a commonly observed feature of innovation (as well as some other type of) contests in the real world.

A set of solvers (e.g., researchers), $\{1, 2, ..., n\}$, may participate in the contest, where $n \geq 2$. Each solver i has a private idea of quality $q_i \geq 0$, which is independently and identically distributed according to a cumulative distribution function (c.d.f.) F with support $[0, w_F]$.[9] The idea quality may not be bounded, in which case $w_F = +\infty$. Different solvers may have ideas with different qualities due to differences in creativity, experience, etc. At any $q \in (0, w_F)$, F is twice continuously differentiable and its density function, $F' \equiv f$, is positive. It is possible that $F(0) > 0$, which means that there is a mass point at idea quality 0. This captures the innovation challenges in which there is a positive probability that a solver does not have an idea to start with. There is no mass point in the distribution at any $q > 0$.

[7]The stochastic dominance they use is introduced by Barlow and Proschan (1966). For distributions with the same mean, it implies second order stochastic dominance. In contrast, the stochastic orders considered in Maskin and Riley (2000) and Pesendorfer (2000) imply first order stochastic dominance.

[8]This assumption is not important for our results, and the consequence of relaxing it is discussed in Example 1.

[9]Our analysis can be generalized to $[m_F, w_F]$ with $w_F > m_F \geq 0$.

We assume that the idea distribution F satisfies the following regularity condition:

Assumption 1 $q + F(q)/f(q)$ is non-decreasing.

Assumption 1 is satisfied if F is log-concave, which implies that this assumption is less restrictive than log-concavity.[10] Moreover, Assumption 1 is similar to the regularity condition of Myerson (1981) which requires $q - (1 - F(q))/f(q)$ to be non-decreasing. The difference is that Myerson's definition uses the hazard rate function $f(q)/(1-F(q))$, while our assumption uses the reverse hazard rate function $f(q)/F(q)$.

After learning the prize and his idea quality, a solver develops his idea into a solution. If solver i has an idea of quality $q_i = 0$, his performance level is 0 independent of how much effort he exerts. If the solver's idea quality is $q_i > 0$, he can submit a solution of performance level $x_i \geq 0$ at a cost of x_i/q_i. Hence, achieving a given performance level costs less if the solver has an idea of higher quality. We use the cost function x_i/q_i for cleaner exposition, but our approach applies to a more general multiplicatively separable cost function $C(x_i, q_i) = x_i L(q_i)$ also, where L is positive, continuously differentiable, and strictly decreasing.[11] Solver i's payoff is $v - x_i/q_i$ if he wins the prize and $-x_i/q_i$ otherwise. All solvers are risk-neutral.

It is reasonable to assume in innovation contests that for a solution to yield value to the seeker, its performance level must be sufficiently high. Consider, for example, the contest to develop a new drug. The new drug must have sufficiently low side effects before it can be utilized by the seeker. Hence, not all winning solutions will be of use to the seeker. To this end, we assume that there is a publicly-known and verifiable exogenous threshold $t > 0$ and solutions with performance levels below t has no benefit to the seeker. For example, the threshold may represent the minimal quality that a vaccine must satisfy for it to be approved by the regulatory authorities. We assume that $\bar{v} - t/w_F > 0$ to ensure that the threshold is not so high that all solvers choose zero performance.

Let $x_{(1)} = \max\{x_1, ..., x_n\}$ stand for the solution with the highest performance. If $x_{(1)} < t$, then none of the solvers wins the prize. Otherwise, the solver with the highest performance wins the prize. In case of a tie, the prize is allocated with equal probability among the tying solvers. We relax the assumption of an exogenously given threshold level in Section 7.2.

The above set-up is built on the contest model of Moldovanu and Sela (2001). The key difference is that in their model, the contest designer maximizes the total expected

[10] See Bagnoli and Bergstrom (2005) for a survey of applications of log-concave distributions. They also provide a comprehensive list of parametric distributions that are log-concave.

[11] Redefining the idea quality as $\tilde{q}_i = 1/L(q_i)$ yields the same set-up.

performance. In contrast, the seeker's profit in our set-up is a function of the maximum performance.[12] Specifically, it consists of two parts. The seeker potentially cares about obtaining a solution that is above the threshold level, and the difference between the threshold and the maximum performance:

$$\Pi(x_{(1)}, v) = \begin{cases} 1 + \lambda(x_{(1)} - t) - v & \text{if } x_{(1)} \geq t \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

If the maximum performance is below the threshold, the solution generates zero profit. Any solution with a performance above the threshold t yields a fixed benefit which is normalized to 1. In addition, $\lambda \geq 0$ stands for the marginal benefit that the seeker receives from an increase in solution performance. If $\lambda = 0$, the seeker does not profit from extra performance beyond the threshold. The higher λ is, the more she benefits from an increase in performance. We study the case in which the marginal benefit is constant and relax this assumption in Section 7.1. The discontinuity in the seeker's profit function allows us to consider the case where the seeker only cares about having a solution that is above the minimum threshold. However, as we discuss in Remark 3, our analysis also applies to the case of a continuous profit function.

The seeker is assumed to be risk-neutral and chooses the prize level v to maximize her expected profit.

4 Optimal Prize

We start by deriving the Bayesian Nash equilibria in a contest with a prize of v. Since the solvers are ex ante symmetric, we focus on the symmetric equilibrium. If $v \leq t/w_F$, all solvers optimally choose zero performance. In the next lemma, we characterize the equilibrium assuming $v > t/w_F$.

Lemma 1 *In a symmetric equilibrium, a solver with idea quality q submits a solution with performance*

$$\beta(q) = \begin{cases} t + vA_t(q) & \textit{if } q \geq q_t \\ 0 & \textit{otherwise} \end{cases}$$

[12] This feature also differentiates our set-up from the all-pay auction model where the designer maximizes the expected revenue, or equivalently, the bidders' total expected payments. Our set-up is also different from other auction models. For example, in a first price auction, the seller's revenue depends on the highest bid, but only the winner pays.

where $A_t(q) = \int_{q_t}^{q} s dF^{n-1}(s)$ *and* q_t *solves*

$$q_t F^{n-1}(q_t) \quad = \quad t/v. \tag{2}$$

According to the equilibrium strategy, if a solver's idea quality is too low, he chooses zero performance. If his idea quality is q_t, he is indifferent between submitting a solution of performance t and 0. If his idea quality is higher than q_t, the solver submits a solution with a performance level above t, and the performance increases with his idea quality. The discontinuity of $\beta(q)$ is a result of the threshold assumption, without which the equilibrium strategy is continuous (as shown in Moldovanu and Sela, 2001).

Given the solvers' equilibrium strategy, the seeker's expected profit can be written in the following way as a function of v:

$$\Pi_F(v) = \int_{q_t}^{w_F} [1 + \lambda(\beta(q) - t) - v] dF^n(q) = L_F(v) + \lambda K_F(v) \tag{3}$$

where

$$L_F(v) \quad = \quad (1-v)(1 - F^n(q_t)) \tag{4}$$

$$K_F(v) \quad = \quad v \int_{q_t}^{w_F} \int_{q_t}^{q} s dF^{n-1}(s) dF^n(q) \tag{5}$$

Equation (4) stands for the seeker's expected profit level when $\lambda = 0$. Since equilibrium performance level is increasing in idea quality, the probability that there is at least one solution with a performance at or above the threshold is given by $1 - F^n(q_t)$. Hence, when $\lambda = 0$, the seeker's profit is $1 - v$ with probability $1 - F^n(q_t)$ and zero otherwise. Her expected profit is $L_F(v) = (1 - v)(1 - F^n(q_t))$.

Equation (5) states the additional expected profit that the seeker makes when $\lambda > 0$. Note that the winner has the highest idea quality $q_{(1)} = \max\{q_1, ..., q_n\}$. From Lemma 1, his equilibrium performance exceeds t by $v \int_{q_t}^{q_{(1)}} s dF^{n-1}(s)$. Hence, $K_F(v)$ in equation (5) stands for the expected difference between t and the winner's performance, and $\lambda K_F(v)$ stands for the seeker's expected profit from this difference.

We consider the seeker's maximization problem. When $\lambda = 0$, the seeker maximizes $L_F(v)$. Notice that if $v \leq t/w_F \equiv \underline{v}_t$, equation (2) implies $q_t \geq w_F$. This means a solver participates only if his idea quality is w_F or higher. Since this occurs with probability zero, $L_F(v) = (1 - v)(1 - 1) = 0$. Moreover, $L_F(v) \leq 0$ if $v \geq 1$. Thus, an optimal prize must be in $(\underline{v}_t, 1)$. In addition, if the seeker's profit $L_F(v)$ is strictly concave in v, there is a unique optimal prize and it satisfies the first order condition $L'_F(v) = 0$. The following lemma states that strict concavity of $L_F(v)$ follows from Assumption 1.

Lemma 2 *Under Assumption* 1, $L_F(v)$ *is strictly concave for* $v \in (\underline{v}_t, 1)$.

While $L_F(v)$ is concave, $K_F(v)$ may be convex.[13] Thus, when $\lambda > 0$, the profit may not be a concave function of the prize. The following condition ensures that there is still a unique optimal prize in $(0, \bar{v})$.

Assumption 2 $L_F(v) - L'_F(v)K_F(v)/K'_F(v)$ is strictly increasing.

Intuitively, Assumption 2 ensures that if the objective function $\Pi_F(v)$ has multiple local maximizers, those maximizers achieve different local maxima. At the end of Section 5, we discuss how the results change when we relax this assumption.

Proposition 1 states the comparative statics result with respect to λ for a given distribution of idea quality. If λ is sufficiently large, the optimal prize may reach the upper boundary $\bar{v}$. Let $\bar{\lambda}_F$ stand for the smallest value of λ such that $\bar{v}$ is an optimal prize. Recall that the optimal prize is below 1 if $\lambda = 0$, so $\bar{\lambda}_F > 0$.

Proposition 1 *Under Assumptions* 1 *and* 2, *there is a unique optimal prize* $V_F(\lambda)$ *for all marginal benefit* $\lambda \neq \bar{\lambda}_F$. *Moreover,* $V_F(\lambda)$ *is weakly increasing in* λ.

Proposition 1 states that as the marginal benefit of solution performance to the seeker increases, the seeker finds it optimal to offer a higher prize. Offering a higher prize encourages the solvers to submit solutions with higher performance levels.

When $\lambda = \bar{\lambda}_F$, there may be two optimal prizes: an interior prize in $(0, \bar{v})$ and $\bar{v}$. Let $V_F(\bar{\lambda}_F)$ be one of the optimal prizes. Our results do not depend on the choice of optimal prize. When there are two optimal prizes, $V_F(\lambda)$ is not continuous in λ.

5 Scarcity of Ideas

In this section, we study comparative statics of the optimal prize with respect to the scarcity of high-quality ideas for a given value of λ.

5.1 Stochastic Dominance Notion of Idea Scarcity

Intuitively, the concept of idea scarcity depends on both the difficulty of the question and the number of solvers. In particular, for a given number of solvers, as the difficulty

[13] For example, $K_F(v)$ is convex if $F(q) = q$ and $t = 1/2$.

level of a challenge increases, high-quality ideas would become scarcer. Similarly, in a challenge with a given level of difficulty, as the number of participating solvers decreases, high-quality ideas would become scarcer. To capture both of these contributing factors, we first define an "effective quality" of a solver's idea, and then use it to introduce a stochastic dominance notion of idea scarcity.

Definition 1 If a solver has an idea of quality q drawn from distribution F, we define *effective quality* of his idea as $\phi_F(q) \equiv qF^{n-1}(q)$. Then, the c.d.f. of the effective quality is $\hat{F} : [0, w_F] \to [0, 1]$ that takes the form $\hat{F}(x) \equiv F(\phi_F^{-1}(x))$. We refer to $\hat{F}$ as the effective distribution of F.

For a solver with idea q_i, his effective quality is q_i discounted by $F^{n-1}(q_i)$, which is the probability that his idea quality is higher than that of all other solvers. Hence, the concept of effective quality, which is crucial for our analysis below, contains information on both a solver's marginal cost, through q_i, and his probability of winning, through $F^{n-1}(q_i)$. Both pieces of information are important for a solver's decision. In comparison, q_i and its distribution F, or $q_{(1)}$ and its distribution F^n contain only one piece of information.

Remark 1 *In auction theory, the expected surplus from trading with a bidder is defined in a similar way to the effective quality defined in Definition 1 (e.g., Bulow and Roberts 1989). To see this, consider a first-price or second-price auction with n symmetric bidders whose values are i.i.d. according to cumulative distribution function F. If the seller trades with a bidder with value v, the total surplus is v. In equilibrium, trade happens with probability $F^{n-1}(v)$, which represents the probability that the bidder's value is higher than the values of the other bidders. Therefore, the expected surplus from trading with a bidder is $x = vF^{n-1}(v)$ and its c.d.f. is $\hat{F}$.*

The effective distribution preserves important properties of the original distribution, such as first order stochastic dominance (FOSD) and log-concavity:

Lemma 3 *Distribution G first order stochastically dominates F, written as $F \prec_{FOSD} G$, if and only if $\hat{F} \prec_{FOSD} \hat{G}$.*

Lemma 4 *If F is log-concave, $\hat{F}$ is also log-concave.*

We now introduce a stochastic dominance order in order to rank idea scarcity.

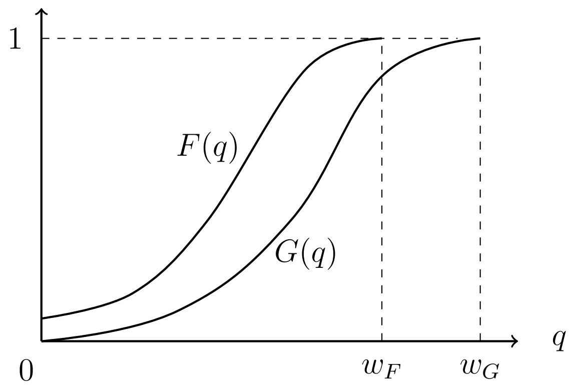

Figure 1: Distributions of Ideas

Definition 2 A distribution $\hat{G}$ dominates $\hat{F}$ in the likelihood ratio order, written as $\hat{F} \prec_{LR} \hat{G}$, if

$$\frac{\hat{g}(x)}{\hat{f}(x)} \leq \frac{\hat{g}(x')}{\hat{f}(x')}$$

for any $x < x'$ and x, x' in the common support of $\hat{G}$ and $\hat{F}$. We say F represents scarcer ideas than G, written as $F \prec G$, if $\hat{F} \prec_{LR} \hat{G}$.

Our stochastic order definition captures some important features of innovative environments. Figure 1 illustrates two distributions with $F \prec G$. The distributions in the figure have different supports and F has a mass point at 0. In other words, in an environment where ideas are scarce, there may be more solvers with zero-quality ideas and/or some high-quality solution ideas may not be available at all.

The above stochastic dominance order is defined indirectly using effective distributions. This approach is similar to the definition of (reverse) hazard rate dominance, which is specified using (reverse) hazard rates. Reverse hazard rate dominance and hazard rate dominance are widely used in the literature on auctions. We need a different stochastic order concept because of the all-pay feature of our design and the fact that the seeker's profit depends on the maximum performance instead of total performance.

The following result discusses how the stochastic order introduced in Definition 2 relates to FOSD. As in the case of hazard rate or reverse hazard rate dominance, the stochastic order stated in Definition 2 is stronger than FOSD. However, as we show in Appendix E, it is equivalent to FOSD for many widely used parametric distribution families, such as exponential, log-normal, Pareto, power function and uniform distributions.

Lemma 5 *$F \prec G$ implies $F \prec_{FOSD} G$.*

It is worth noting that the comparison of two distributions in terms of scarcity of ideas depends on n, the (exogenously given) number of potential solvers. The following result shows that the scarcity order of two distributions is never reversed as n varies.

Lemma 6 *If $F \prec G$ for some $n \in \mathbb{N}$, there is no $n' \in \mathbb{N}$ such that $G \prec F$.*

To see why, suppose $F \prec G$ for n and $G \prec F$ for n'. Then, Lemma 5 implies $F \prec_{FOSD} G$ and $G \prec_{FOSD} F$ which cannot happen. Although it does not happen for the parametric distributions considered in Appendix E, it is possible that $F \prec G$ holds for some but not all $n \in \mathbb{N}$.

Recall that (2) implies that $\phi_F^{-1}(t/v) = q_t$, so $1 - \hat{F}(t/v) = 1 - F(q_t)$ is the ex ante probability for a solver to actively participate (i.e., submit a solution with positive performance) in a contest. We refer to $P_F(v) \equiv 1 - \hat{F}(t/v)$ as the *participation rate* and $1 - P_F(v)$ as the *non-participation rate*. Then, the expected number of actively participating players is $nP_F(v)$. The elasticity of the participation rate w.r.t. the prize is $vP_F'(v)/P_F(v)$, and the elasticity of the non-participation rate w.r.t. the prize is $-vP_F'(v)/(1 - P_F(v))$.

The following lemma shows that scarcer ideas lead to more elastic participation and non-participation rates. Intuitively, when high-quality ideas are scarce, solvers are more responsive to marginal changes in the prize level.

Lemma 7 *$F \prec G$ implies that for any $v > 0$,*

$$P_G'(v)/P_G(v) \quad < \quad P_F'(v)/P_F(v) \tag{6}$$

$$-P_G'(v)/(1 - P_G(v)) \quad < \quad -P_F'(v)/(1 - P_F(v)) \tag{7}$$

The proof is straightforward from the definitions and is therefore omitted. Note that all our results remain the same if we replace $F \prec G$ with (6) and (7). This is discussed further in Remark 2 below. Lemma 7 implies that if a seeker could obtain information, for instance through a survey, on the participation rates of two challenges at different prize levels, she could compare the scarcity of ideas between the two challenges. In the challenge with scarcer ideas, the participation rate would be lower at all prize levels.

5.2 Comparative Static Analysis

Consider two distributions, F with support $[0, w_F]$ and G with support $[0, w_G]$, which are twice continuously differentiable. Assume that the density functions $F' \equiv f$ and $G' \equiv g$ are positive. Recall that $L_F(v)$ and $K_F(v)$ described in (4) and (5) only depend on F. Let $L_G(v)$ and $K_G(v)$ be the counterparts for distribution G. They only depend on G.

Assumption 3 $L'_F(v)/K'_F(v)$ and $L'_G(v)/K'_G(v)$ cross at most once.

Assumptions 3 ensures that $V_F(\lambda) - V_G(\lambda)$ crosses zero exactly once.[14]

The following proposition is the main result of the paper. We present its proof in Section 6.

Proposition 2 *Under Assumptions 1-3, if $F \prec G$, there exists a unique $\hat{\lambda} > 0$ such that*

(i) $V_G(\lambda) < V_F(\lambda)$ if $\lambda < \hat{\lambda}$;

(ii) $V_G(\lambda) \geq V_F(\lambda)$ if $\lambda > \hat{\lambda}$.

That is, scarcer ideas lead to a lower optimal prize if and only if the marginal benefit of solution performance is sufficiently high.

It is worth mentioning that $V_G(\hat{\lambda}) = V_F(\hat{\lambda})$, $V_G(\hat{\lambda}) < V_F(\hat{\lambda})$ or $V_G(\hat{\lambda}) > V_F(\hat{\lambda})$ is possible. In the last two cases, $V_G(\lambda)$ may not change continuously at $\hat{\lambda}$, and it may jump to $\bar{v}$ if λ is slightly above $\hat{\lambda}$.

Proposition 2 implies that the value of λ plays a critical role in determining how the optimal prize changes with the distribution of ideas. We first illustrate Proposition 2 with the following example and then explain the intuition behind it.

Example 1 *Suppose $t = 0.4$, $n = 4$, $\bar{v} = 1.5$ and consider two Pareto distributions $F(q) = 1 - (1+q)^{-4}$ and $G(q) = 1 - (1+q)^{-2}$. The distributions satisfy Assumptions 1-3 and $F \prec G$. As in Figure 2, there exists $\hat{\lambda} \approx 0.9$ such that $V_F(\lambda) > V_G(\lambda)$ if $\lambda \in [0, \hat{\lambda})$, $V_F(\lambda) = V_G(\lambda)$ if $\lambda = \hat{\lambda}$, and $V_F(\lambda) < V_G(\lambda)$ if $\lambda \in (\hat{\lambda}, 3.6)$. For $\lambda \geq 3.6$, $V_F(\lambda) = V_G(\lambda) = \bar{v}$.*[15]

For the intuition behind the result, consider the seeker's expected payoff function given in (3). When $\lambda = 0$, the seeker maximizes (4). What is important is to have a solution with performance above the minimum requirement. In this case, if the distribution of ideas shifts such that high-quality ideas become scarcer, it puts downward pressure on the probability of success (i.e., the probability that at least one solution meets the requirement). In addition, scarcer ideas reduce competition among the solvers and therefore increase participation, which puts upward pressure on the probability of success. It turns out the first effect always dominates the second, that is, scarcer ideas lead to lower

[14]In general, $V_F(\lambda) - V_G(\lambda)$ may not be monotone in λ. For example, Figure 2 shows that $V_F(\lambda) - V_G(\lambda)$ is decreasing for small values of λ and increasing for large values of λ.

[15]Recall that we assume $\bar{v} > 1$. If $\bar{v}$ is too small, for instance $\bar{v} < 0.4$, the optimal prize may be forced to be the upper bound: $V_G(\lambda) = V_F(\lambda) = \bar{v}$ for all λ. Moreover, if $\bar{v} < V_G(\hat{\lambda}) = V_G(\hat{\lambda})$, then $V_F(\lambda)$ and $V_G(\lambda)$ are forced to be $\bar{v}$ for $\lambda > \hat{\lambda}$. In this case, Proposition 2 still holds, but the strict inequality $V_G(\lambda) > V_F(\lambda)$ illustrated in Example 1 may not arise.

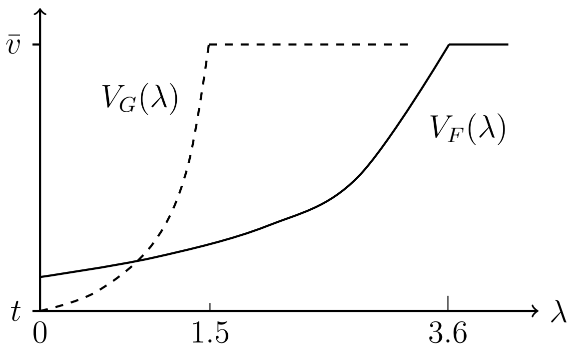

Figure 2: Optimal Prizes

probability of success. Hence, the seeker compensates by increasing the prize level (as implied by Lemma 8 in the next section). On the other hand, if high-quality ideas are abundant, the probability of success is high to start with, so the seeker does not need to have as high a prize. The marginal return of the prize is lower.

When $\lambda > 0$, the seeker maximizes the sum of (4) and (5). Now the seeker cares both about meeting the threshold and receiving as good a solution as possible. The magnitude of λ determines how much weight the seeker will put on (5). If the seeker increases the prize, more solvers participate and each participating solver increases their performance. Thus, an increase in the prize pushes the expected performance further beyond the requirement. When ideas are abundant, the average idea quality (solvers' productivity) is higher, so the marginal return of the prize is higher. Hence, if λ is sufficiently high, then this effect dominates and in challenges where ideas are more abundant (scarce), the seeker finds it optimal to set a higher (lower) prize (as implied by Lemma 9 in the next section).

In summary, while maximizing the expected payoff function $\Pi_F(v) = L_F(v) + \lambda K_F(v)$, idea abundance and prizes play substitute roles in the first term and complementary roles in the second term. This represents the main trade-off faced by a contest designer and λ determines the relative weight of the two elements.

Remark 2 For cleaner exposition, we use the likelihood ratio order defined in Definition 2, but our results can be generalized to less restrictive stochastic orders. It is well-known that the combination of hazard rate and reverse hazard rate orders is less restrictive than the likelihood ratio order.[16] All of our results and their proofs remain unchanged if we replace the likelihood ratio stochastic dominance used in Definition 2 with hazard rate and reverse hazard rate stochastic dominance. We can also relax $F \prec G$ to (6) and (7).

[16]See, for instance, Shaked and Shanthikumar (2007), p. 55.

This is because the two inequalities imply that $\hat{G}$ stochastically dominates $\hat{F}$ in terms of reverse hazard rate and hazard rate, which is sufficient for all our results.

Remark 3 Suppose that if the maximum performance is above the minimum performance requirement, the seeker's profit is given by

$$\Pi(x_{(1)}, v) = \begin{cases} \lambda(x_{(1)} - t) - v & \text{if } x_{(1)} \geq t \\ 0 & \text{otherwise} \end{cases}$$

instead of (1). That is, the profit is continuous in the maximum performance. Then, $L_F(v) = -v(1 - F^n(q_t))$ and $K_F(v)$ remains the same. Since there is no profit obtained from reaching the minimum performance requirement, if the marginal benefit is $\lambda = 0$, then $V_F(\lambda) = V_G(\lambda) = 0$. By virtually the same proof as Proposition 2, we obtain the critical value $\hat{\lambda} = 0$ and case (i) in Proposition 2 never arises. This means that in this case, scarcer ideas always lead to lower prizes.

Before ending this section, we discuss how the results are different if we relax Assumptions 2 and 3. Without Assumption 2, there may be multiple optimal prizes. Then, $V_F(\lambda)$ and $V_G(\lambda)$ represent the set of optimal prizes. To compare sets of possibly multiple optimal prizes, we use the strong set order (see, e.g., Topkis, 1978).

Definition 3 For two sets $A, B \subset \mathbb{R}^m$, A is greater than B in the strong set order, written as $A \geq B$, if, for any $a \in A$ and any $b \in B$, the pointwise maximum $a \vee b \in A$ and the pointwise minimum $a \wedge b \in B$.

If $m = 1$, then $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. For example, $\{2, 3, 4\} \geq \{1, 2, 3\}$ and $[2, 4] \geq [1, 3]$. If both A and B are singletons, the strong set order reduces to the order over real numbers. The following proposition shows that a result similar to Proposition 2 can be obtained if Assumptions 2 and 3 are relaxed.

Proposition 3 *Under Assumption 1, if $F \prec G$, there exist $\hat{\lambda}' \geq \hat{\lambda} > 0$ such that*

i) $V_F(\lambda) \geq V_G(\lambda)$ if $\lambda < \hat{\lambda}$.

ii) $V_F(\lambda) \leq V_G(\lambda)$ if $\lambda > \hat{\lambda}'$.

Notice that $\hat{\lambda}'$ and $\hat{\lambda}$ may be different, which is the key difference between Propositions 2 and 3. This is partly because the strong set order is only a partial order if $V_F(\lambda)$ or $V_G(\lambda)$ contains multiple prizes. In contrast, when the optimal prize is unique as in Proposition 2, the strong set order becomes the complete order over real numbers.

6 Proof of Proposition 2

This section contains the proof of our main result. Consider first the case with $\lambda = 0$. Then, the seeker's profit is given by $L_F(v)$ as in (4). Similarly, define $L_G(v) = (1-v)(1-G^n(q'_t))$ where q'_t solves

$$q'_t G^{n-1}(q'_t) = t/v \tag{8}$$

The corresponding marginal profits are $L'_F(v)$ and $L'_G(v)$. The marginal profit functions have two properties as illustrated in Figure 3. First, as established in Lemma 2, they are strictly decreasing in the prize value. Second, as the prize value increases, the marginal profit function under more abundant ideas, $L'_G(v)$, reaches the horizontal axis before the marginal profit function under scarcer ideas, $L'_F(v)$, does. This property is formalized below.

Lemma 8 *Suppose $F \prec G$. Then, $L'_F(v) = 0$ implies $L'_G(v) < 0$.*

When $\lambda = 0$, the unique optimal prize satisfies the first order conditions $L'_F(v) = 0$ and $L'_G(v) = 0$. Lemma 8 implies that, when $\lambda = 0$, $V_F(\lambda) > V_G(\lambda)$ (as shown in Figure 3). Hence, as high-quality ideas become scarcer, the seeker should increase the prize. Abundance of high-quality ideas and the prize amount play substitute roles in profit maximization.

Many comparative static results in the literature rely on submodularity or supermodularity (e.g., Topkis, 1978). However, the objective function $L_F(v)$ is neither submodular nor supermodular in $(v; F)$. Specifically, notice that the submodularity of $L_F(v)$ in $(v; F)$ requires $L'_G(v) < L'_F(v)$ for all v, but Lemmas A.2 and A.4 in Appendix B show otherwise (as illustrated in Figure 3).[17] Other comparative static results in the literature rely on single-crossing conditions (e.g., Milgrom and Shannon, 1994).[18] Our comparative static analysis is with respect to distributions, and Lemma 8 states a functional form of the single-crossing condition for $L_F(v)$ in $(v; F)$.

Consider next the case when $\lambda > 0$. Then, the expected profit is $\Pi_F(v) = L_F(v) + \lambda K_F(v)$, where $K_F(v)$, given in (5), is the expected additional performance beyond the threshold t. The following lemma states two relevant properties of K'_F and K'_G.

Lemma 9 *$K'_F(v) > 0$. Moreover, $F \prec G$ implies $K'_F(v) < K'_G(v)$.*

[17] Lemma A.2 implies $L'_F(1) - L'_G(1) > 0$ and Lemma A.4 implies $L'_F(t) - L'_G(t) < 0$ if F and G have a common support $[0, 1]$. Therefore, $L'_F(v)$ and $L'_G(v)$ must intersect as in Figure 3.

[18] They also require quasi-submodularity of the objective function. Lemma 2 proves quasi-submodularity of $L_F(v)$ in v.

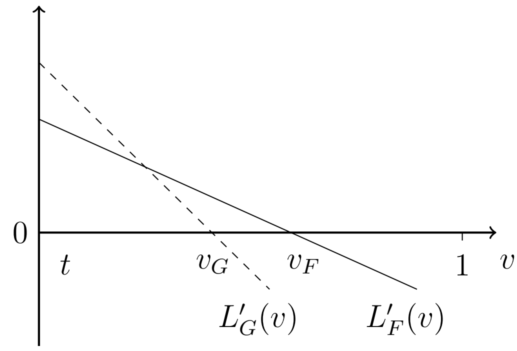

Figure 3: Optimal Prizes if $\lambda = 0$

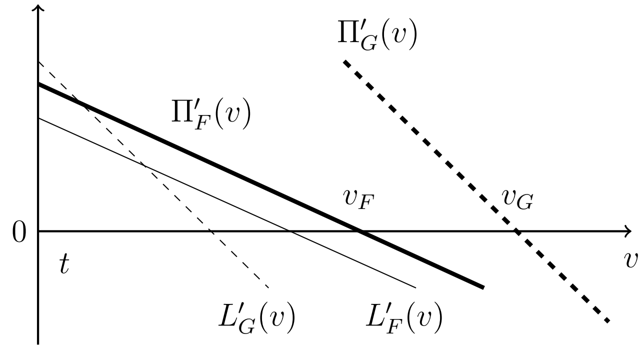

Figure 4: Optimal Prizes if λ is Large

Figure 4 illustrates Lemma 9. Note that in the figure, $\lambda K'_F(v)$ corresponds to the vertical distance between the curves $\Pi'_F(v)$ and $L'_F(v)$. The first property stated in the lemma is that the expected additional performance beyond the threshold is increasing in the prize. Given $K'_F(v) > 0$, a positive λ shifts $L'_F(v)$ upwards to be $\Pi'_F(v) = L'_F(v) + \lambda K'_F(v)$. Moreover, because $\lambda K'_F(v)$ is larger with a higher λ, $V_F(\lambda') \geq V_F(\lambda)$ if $\lambda' \geq \lambda$. In other words, because a larger λ leads to a higher marginal profit of prize, the optimal prize should not be lower.

The second property stated in Lemma 9 implies that $K_F(v)$ is supermodular in $(v; F)$, which means the marginal impact of the prize on $K_F(v)$ is smaller when ideas are scarcer. In Figure 4, the distance between $L'_G(v)$ and $\Pi'_G(v)$ is larger than the distance between $L'_F(v)$ and $\Pi'_F(v)$. As a result, if λ is sufficiently large, $\Pi'_G(v)$ is larger than $\Pi'_F(v)$.

In summary, Lemma 9 implies that a marginal increase in the prize amount pushes the expected performance further beyond the minimal requirement. Moreover, the marginal increase in the prize amount has higher returns if the solvers have more abundant ideas. As a result, as high-quality ideas become more abundant, the seeker should increase the prize. Abundance of high-quality ideas and the prize amount play complementary roles in profit maximization.

According to Proposition 1, the optimal prize reaches the upper boundary $\bar{v}$ when λ is sufficiently large. Therefore, define $\bar{\lambda}_F = \inf\{\lambda | V_F(\lambda) = \bar{v}\}$ and $\bar{\lambda}_G = \inf\{\lambda | V_G(\lambda) = \bar{v}\}$.[19] The following lemma shows that, as λ increases from 0, $V_G(\lambda)$ reaches the upper boundary before $V_F(\lambda)$ does.

Lemma 10 *If $F \prec G$, then $0 < \bar{\lambda}_G < \bar{\lambda}_F < +\infty$.*

We now combine the lemmas above to prove Proposition 2.

[19] See the proof of Proposition 1 in Appendix A for the expressions of $\bar{\lambda}_F$ and $\bar{\lambda}_G$.

Proof of Proposition 2. We proceed in three steps.

Step I. $V_F(\lambda) > V_G(\lambda)$ for λ close to 0. To see this, recall that if $\lambda = 0$, since Lemma 2 states that the profit is strictly concave in the prize value, its maximum is reached at an interior prize value. Therefore, both $V_F(0)$ and $V_G(0)$ satisfy the first order conditions $L'_F(V_F(0)) = 0$ and $L'_G(V_G(0)) = 0$, which combined with Lemma 8 imply that $V_G(0) < V_F(0)$. By continuity of $\Pi_F(v)$ in λ, if λ is sufficiently close to 0, we still have $V_G(\lambda) < V_F(\lambda)$.

Step II. $V_F(\lambda) \leq V_G(\lambda)$ if λ is large enough. Lemma 10 implies that $V_F(\lambda) \leq V_G(\lambda) = \bar{v}$ for $\lambda > \bar{\lambda}_G$.

Step III. In this step, we show existence and uniqueness of $\hat{\lambda}$. From Step I, $V_F(\lambda) > V_G(\lambda)$ for λ close to 0. Consider three possibilities: First, $V_F(\lambda) > V_G(\lambda)$ for all $\lambda \leq \bar{\lambda}_G$. Then, $\hat{\lambda} = \bar{\lambda}_G$ and the proposition holds.

Second, $V_F(\lambda) = V_G(\lambda)$ for some $\lambda \in (0, \bar{\lambda}_G]$. Notice that for any λ such that $V_F(\lambda) = V_G(\lambda) = v$, the first order conditions are

$$L'_F(v) + \lambda K'_F(v) = 0 \tag{9}$$

$$L'_G(v) + \lambda K'_G(v) = 0 \tag{10}$$

Solving (9) for λ and substituting into (10) yields

$$L'_F(v)/K'_F(v) = L'_G(v)/K'_G(v)$$

This equation has a unique solution because $L'_F(v)/K'_F(v)$ and $L'_G(v)/K'_G(v)$ cross at most once due to Assumption 3. Thus, there is a unique λ such that $V_F(\lambda) = V_G(\lambda)$. Letting $\hat{\lambda}$ be this unique λ gives us the result stated in the proposition.

Third, $V_F(\lambda) < V_G(\lambda)$ for some $\lambda_0 \in (0, \bar{\lambda}_G]$. Notice that if V_F is discontinuous at an interior point of $(0, \bar{\lambda}_F)$, there are at least two optimal prizes at this interior λ. This contradicts with Proposition 1, so V_F is continuous over $(0, \bar{\lambda}_F)$. Similarly, V_G is continuous over $(0, \bar{\lambda}_G)$. Recall that $V_F(\lambda_0) < V_G(\lambda_0)$, so $\lambda_0 < \bar{\lambda}_F < \bar{\lambda}_G$ and therefore V_F and V_G are continuous over $(0, \lambda_0)$. Moreover, $V_F(\lambda) > V_G(\lambda)$ if λ is sufficiently close to 0 and that $V_F(\lambda_0) < V_G(\lambda_0)$, so the intermediate value theorem implies that there is a $\lambda \in (0, \lambda_0)$ such that $V_F(\lambda) = V_G(\lambda)$. Then, we can prove the proposition as in the second case above. ■

7 Extensions

7.1 Nonlinear Benefits

So far we have assumed that the marginal benefit of additional performance above the threshold is constant. In this section, we show that our results continue to hold when we relax this assumption by generalizing the seeker's benefit to

$$\Pi(x_{(1)}) = \begin{cases} 1 + \lambda B(x_{(1)} - t) - v & \text{if } x_{(1)} \geq t \\ 0 & \text{otherwise} \end{cases}$$

where B is differentiable and satisfies $B(0) = 0$, $B'(x) > 0$ for all $x \geq 0$. It is not necessarily concave or convex. In previous sections, we considered the special case where $B(x) = x$.

Since the change in the seeker's benefit function does not have an impact on the solvers' behavior, the solvers' equilibrium strategies remain the same as in Lemma 1. As in Section 3, we can also write the expected profit as $\Pi_{F,B}(v) = L_F(v) + \lambda K_{F,B}(v)$, where $L_F(v)$ is the same as in (4) and

$$K_{F,B}(v) = \int_{q_t}^{w_F} B\left(v \int_{q_t}^{q} s dF^{n-1}(s)\right) dF^n(q)$$

is the expected benefit from additional performance above the threshold t.

The following two assumptions are generalizations of Assumptions 2 and 3 to nonlinear benefit functions. They imply Assumptions 2 and 3 as special cases.

Assumption 2′ $L_F(v) - L'_F(v)K_{F,B}(v)/K'_{F,B}(v)$ is strictly decreasing.

Assumption 3′ $L'_F(v)/K'_{F,B}(v)$ and $L'_G(v)/K'_{F,B}(v)$ cross at most once and $\lim_{x\to\infty} B'(x) = \underline{b} > 0$.

The first part of Assumption 3′ is similar to Assumption 3, but the second part is new. The second part requires that the marginal value $B'(x)$ converges to a positive value as x goes to infinity. If this is violated and the marginal value converges to zero, for sufficiently large λ, the comparative statics are similar to $\lambda = 0$. This is because the benefit function for sufficiently large λ has a similar shape to that with $\lambda = 0$.

By replacing Assumptions 2 and 3 with Assumptions 2′ and 3′, we can generalize Propositions 1-3 to accommodate nonlinear benefits. This is done in Propositions A.1-A.3. Because these results and their proofs are similar to those of Propositions 1-3, we relegate them to the appendix.

7.2 Endogenous Reservation Performance

In this section, we assume that the seeker can choose both a prize value v and a performance level r such that the prize is awarded only if the highest performance is at least r. Since r plays a similar role to a reservation price in the auction literature, we call it reservation performance.[20]

Note that given the seeker's choice (v, r), the solvers' equilibrium strategy is the same as in Lemma 1 except that t is replaced by r:

$$\beta(q_i) = \begin{cases} r + vA_r(q_i) & \text{if } q_i \geq q_r \\ 0 & \text{otherwise} \end{cases}$$

where $A_r(q_i) = \int_{q_r}^{q_i} s dF^{n-1}(s)$ and q_r is the unique solution of

$$q_r F^{n-1}(q_r) = r/v. \tag{11}$$

An increase in r causes an upward shift in $\beta(q_i)$. Hence, the seeker may want to set a higher r in order to elicit solutions with higher performance levels, which becomes more important as λ increases.

Although the seeker may choose any $r \geq 0$, we first note that choosing a reservation performance level $r < t$ is never optimal.[21] This is because performance levels below t are worthless to the seeker. Thus, we only need to consider $r \geq t$. Given such a $r \geq t$, solvers with $q_i < q_r$ do not participate and those with $q_i > q_r$ choose a performance level above $r \geq t$. As a result, the expected profit of the seeker is given by

$$\int_{q_r}^{w_F} [1 + \lambda(\beta(q) - t)] dF^n(q) - v(1 - F^n(q_r))$$

Substituting $\beta(q)$ into the profit, we can rewrite it as a function of v and q_r:[22]

$$\Pi_F(v, q_r) = L_F(v, q_r) + \lambda K_F(v, q_r) \tag{12}$$

[20] Another question of interest is the optimal mechanism (see, for example, Chawla et al., 2015), and how it varies with the distribution of idea qualities. However, this question is beyond the scope of this paper, which focuses on a common type of innovation contests.

[21] See Lemma A.6.

[22] For more details, see Appendix D.

where

$$\begin{aligned} L_F(v, q_r) &= (1-v)(1-F^n(q_r)) \\ K_F(v, q_r) &= \int_{q_r}^{w_F} \left[vF^{n-1}(q) \left(q - \frac{1-F^n(q)}{(F^n)'(q)} \right) - t \right] dF^n(q) \end{aligned}$$

Hence, the seeker chooses (v, q_r) instead of (v, r). Since $r \geq t$ is equivalent to $q_r \geq q_t$, the seeker chooses $v \geq 0$ and $q_r \geq q_t$ to maximize $\Pi_F(v, q_r)$. Given (v, q_r), r can be recovered from (11). Let $VR_F(\lambda)$ denote the set of optimal pairs of $(v, r) \in [0, \bar{v}] \times \mathbb{R}_+$ when the idea quality distribution is F.

Lemma 11 *If $\lambda = 0$, the optimal reservation performance level is t.*

According to Lemma 11, when $\lambda = 0$, even though the seeker has the option to choose a r value that is different from t, it is optimal to choose $r = t$. Announcing $r < t$ is not optimal because it increases expected prize payments without any benefits. It is not optimal to announce $r > t$ either since it causes the seeker to miss on profitable opportunities.

The only reason for setting $r > t$ may be to increase performance further beyond the threshold t which becomes important for sufficiently high values of λ. The following lemma formalizes this intuition.

Lemma 12 *If λ is sufficiently large, the optimal reservation performance level is higher than threshold t.*

Assumption 4 $q - \dfrac{1-F^n(q)}{(F^n)'(q)}$ increases in q, and $\lim_{q \to w_F} \left(q - \dfrac{1-F^n(q)}{(F^n)'(q)} \right) > 0$.[23]

Assumption 4 states that the virtual valuation of $q_{(1)}$, the best idea quality, is increasing, and it is positive at the highest value in the support. This assumption ensures that for all λ, each solver's probability of participation is bounded away from zero under the optimal reservation performance and prize.[24]

In order to compare $VR_F(\lambda)$ for different λ, we need to use the strong set order given in Definition 3 for higher dimensions with $m = 2$. We can also use the order to define monotonicity of $VR_F(\lambda)$ and $VR_G(\lambda)$. For example, we say $VR_F(\lambda)$ is monotone non-decreasing/non-increasing in λ if for any $\lambda_1 > \lambda_2$, $VR_F(\lambda_1)$ is no lower/higher than $VR_F(\lambda_2)$. The following result is analogous to Proposition 1.

[23] This assumption allows for the case where $\lim_{q \to w_F} \left(q - \dfrac{1-F^n(q)}{(F^n)'(q)} \right) = +\infty$.

[24] See the proof of Proposition 4 in Appendix D.

Proposition 4 *Under Assumptions* 1 *and* 4*, there exist* $\lambda'' \geq \lambda' > 0$ *such that*

i) $VR_F(\lambda)$ *is monotone non-decreasing for* $\lambda < \lambda'$.

ii) $VR_F(\lambda)$ *is monotone non-increasing for* $\lambda > \lambda''$.

As different from Proposition 1, Proposition 4 considers both the optimal prize and the optimal reservation performance, but it focuses on sufficiently small or large λ values only. In the proof of Proposition 4, we show that for λ sufficiently small, the optimal reservation performance is equal to the threshold, and the optimal prize is increasing in λ. In contrast, for sufficiently large λ, the optimal prize is equal to the upper boundary $\bar{v}$, and the optimal reservation performance is strictly decreasing in λ. Hence, the optimal prize is weakly increasing in λ for $\lambda < \lambda'$ and $\lambda > \lambda''$, which is in line with Proposition 1.

The intuition for the above result is as follows. Notice that $\frac{\partial^2 L_F(v,q_r)}{\partial v \partial q_r} > 0$, and q_r is strictly increasing in r due to (11), so v and r play complementary roles in $L_F(v, q_r)$. In contrast, $\frac{\partial^2 K_F(v,q_r)}{\partial v \partial q_r} < 0$, which means v and q_r play substitute roles in $K_F(v, q_r)$.[25] As a result, as λ increases from 0, the optimal v and q_r first change in the same direction due to their impact on $L_F(v, q_r)$, but eventually they may move in different directions due to their impact on $K_F(v, q_r)$.

For a given λ, let $VR_G(\lambda)$ be the set of optimal (v, r) associated with distribution G. Recall that $r = 0$ if $\lambda = 0$. Therefore, Proposition 2 implies that $VR_F(\lambda) \geq VR_G(\lambda)$ if $\lambda = 0$. The comparison of $VR_F(\lambda)$ and $VR_G(\lambda)$ for sufficiently large λ is closely related to the two equations below:

$$F^{n-1}(q)\left(q - \frac{1 - F^n(q)}{(F^n)'(q)}\right) = t/\bar{v} \tag{13}$$

$$G^{n-1}(q)\left(q - \frac{1 - G^n(q)}{(G^n)'(q)}\right) = t/\bar{v} \tag{14}$$

Equation (13) is the first order condition $\frac{\partial \Pi_F(v,q_r)}{\partial q_r} = 0$ evaluated at $\lambda \to \infty$, and (14) is the counterpart for G.[26] These equations have a unique solution under Assumption 4. Let q_F be the solution to (13) and q_G be the solution to (14). The following result compares $VR_F(\lambda)$ and $VR_G(\lambda)$:

Proposition 5 *Under Assumptions* 1 *and* 4*, if* $F \prec G$*, there exist* $\hat{\lambda}' \geq \hat{\lambda} > 0$ *such that*

i) for $\lambda < \hat{\lambda}$, $\quad VR_F(\lambda) \geq VR_G(\lambda)$.

ii) for $\lambda > \hat{\lambda}'$, $\quad VR_F(\lambda) \leq VR_G(\lambda)$ *if* $q_G G^{n-1}(q_G) \geq q_F F^{n-1}(q_F)$;
$VR_F(\lambda) \geq VR_G(\lambda)$ *if* $q_G G^{n-1}(q_G) < q_F F^{n-1}(q_F)$.

[25] Notice that $K_F(v, q_r) = \int_{q_r}^{w_F} [\beta(q) - t]\, dF^n(q)$, so $\frac{\partial^2 K_F(v,q_r)}{\partial v \partial q_r} = -A_r(q_r)(F^n)'(q_r) < 0$.

[26] See the proof of Proposition 5 in Appendix D.

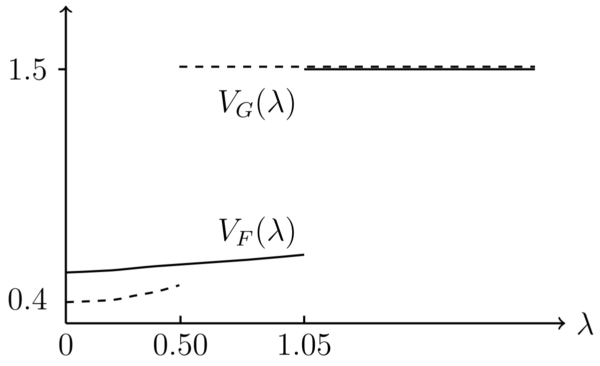

Figure 5: Optimal Prize

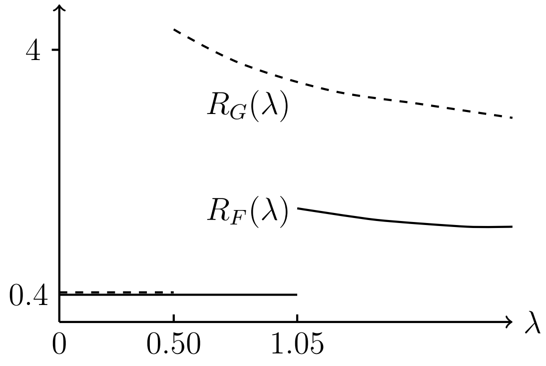

Figure 6: Optimal Reservation Performance

Proposition 5 is analogous to Proposition 3. We show in the proof of Proposition 5 that the optimal reservation performance is equal to the threshold for $\lambda < \hat{\lambda}$. Hence, part i) of Proposition 5 generalizes part i) of Proposition 3 by considering both the optimal prize and reservation performance. The difference in part ii) between Propositions 3 and 5 comes from the optimal reservation performance. As we show in the proof, the optimal reservation performance with G, representing the case of more abundant ideas, will be lower if the corresponding effective quality of q_G is lower than that of q_F.

The results in Proposition 5 follow from the following observations made in the proof. When λ is sufficiently small, the optimal prize associated with F is higher than that associated with G. The optimal reservation performance under both F and G is the same and equal to r. When λ is sufficiently large, the optimal prize under both F and G is equal to the upper boundary $\bar{v}$. If $\lambda \to +\infty$, the optimal reservation performance associated with F is above r and converges to $\bar{v}q_F F^{n-1}(q_F)$ and that associated with G is above r and converges to $\bar{v}q_G G^{n-1}(q_G)$.

The following example illustrates Proposition 5.

Example 2 *Consider Example 1, but now assume that the seeker chooses* (v, r)*. Recall that* $t = 0.4$*,* $n = 4$*,* $\bar{v} = 1.5$*. For each* λ*, there is a unique optimal prize, denoted as* $V_F(\lambda)$*, and a unique optimal reservation performance, denoted as* $R_F(\lambda)$*. There exists* $\hat{\lambda} = 0.50$ *such that* $V_F(\lambda) > V_G(\lambda)$ *and* $R_F(\lambda) = R_G(\lambda)$ *for* $\lambda < \hat{\lambda}$*, and* $V_G(\lambda) \geq V_F(\lambda)$ *and* $R_G(\lambda) > R_F(\lambda)$ *for* $\lambda > \hat{\lambda}$*. Figures 5 illustrates the optimal prizes and Figure 6 illustrates the optimal reservation performance.*

8 Conclusion

Although innovation contests have a long history, there has been an increase in their use in recent years. One of the reasons for this growth in popularity of innovation contests is that progress in information technology has made it easier to run innovation contests using the Internet. As a result, several innovation platforms have emerged on the Internet as the meeting place of seekers of innovative solutions and solvers of innovation problems.[27] Especially when it is not possible to identify ex ante who has the expertise to solve a specific challenge, it is useful to make the challenge public to many potential solvers.[28]

For designers of innovation contests, a significant challenge is what prize to set. We model innovation contests assuming both ideas and effort are integral parts of the innovation process. When the innovation challenge is a difficult one and high-quality ideas will not be commonly observed, is it always optimal to post a high prize? We analyze this question by introducing a novel way of capturing idea scarcity with a new order of stochastic dominance.

Our analysis uncovers that while determining the prize level, contest designers should consider how much they will benefit from a marginal increase in performance as well as how difficult the challenge is. It is not necessarily the case that they should incentivize harder challenges with higher prizes. If the marginal benefit of performance is low, the optimal prize increases with the scarcity of ideas; if the marginal benefit is higher, the optimal prize decreases with the scarcity of ideas.

[27]See, for example, InnoCentive, IdeaConnection, and OpenIDEO.

[28]Indeed, Jeppesen and Lakhani (2010) provide evidence that the winning solution may often come from “nonobvious individuals”.

References

Arnold, V. I. (1984), *Ordinary Differential Equations*, Springer.

Athey, S. (2002), "Monotone Comparative Statics under Uncertainty", *Quarterly Journal of Economics*, 117, pp. 187-223.

Bagnoli, M., and Bergstrom, T. (2005), "Log-concave Probability and Its Applications", *Economic theory*, 26, pp. 445-469.

Barlow, R. E. and Proschan, F. (1966), "Inequalities for Linear Combinations of Order Statistics from Restricted Families", *Annals of Mathematical Statistics*, 37, pp. 1593-1601.

Bjelland, O. M. and R. C. Wood. (2008), "An Inside View of IBM's 'Innovation Jam'", *MIT Sloan Management Review*, 50, pp. 32-40.

Bulow, J. and Roberts, J. (1989), "The Simple Economics of Optimal Auctions", *Journal of Political Economy*, 97, pp. 1060-1090.

Che, Y. and I. Gale. (2003), "Optimal Design of Research Contests", *American Economic Review*, 93, pp. 646-671.

Chawla, S., Hartline, J. D. and Sivan, B. (2015), "Optimal Crowdsourcing Contests", *Games and Economic Behavior.*

Erat, S. and V. Krishnan. (2012), "Managing Delegated Search Over Design Spaces", *Management Science*, 58, pp. 606-623.

Fu, Q., J. Lu, and Y. Lu (2012), "Incentivizing R&D: Prize or Subsidies?", *International Journal of Industrial Organization*, 30, pp. 67-79.

Füller, J., M. Bartl, H. Ernst, and H. Mühlbacher. (2006), "Community Based Innovation: How to Integrate Members of Virtual Communities into New Product Development", *Electronic Commerce Research*, 6, pp. 57-73.

Fullerton, R., B. Linster, M. McKee, and S. Slate. (2002) "Using Auctions to Reward Tournament Winners: Theory and Experimental Investigations", *RAND Journal of Economics*, 33, pp. 62-84.

Fullerton, R. and R. P. McAfee. (1999), "Auctioning Entry into Tournaments", *Journal of Political Economy*, 107, pp. 573-605.

Galassoa, A., M. Mitchell, and G. Virag (2018), “A Theory of Grand Innovation Prizes”, *Research Policy*, 47, pp. 343-362.

Ganuza, J.J. and E. Hauk. (2006), “Allocating Ideas: Horizontal Competition in Tournaments”, *Journal of Economics & Management Strategy*, 15, pp. 763-787.

Hoppe, H. C., Moldovanu, B. and Sela, A. (2009). “The Theory of Assortative Matching Based on Costly Signals”, *Review of Economic Studies*, 76, pp. 253-281.

Jeppesen, L. B. and K. R. Lakhani. (2010). “Marginality and Problem-Solving Effectiveness in Broadcast Search”, *Organization Science*, 21, pp. 1016-33.

Kornish L. J. and K. T. Ulrich. (2014), “The Importance of the Raw Idea in Innovation: Testing the Sow’s Ear Hypothesis”, *Journal of Marketing Research,* 51, pp. 14-26.

Korpeoglu, E. and S. Cho (2018), “Incentives in Contests with Heterogeneous Solvers”, *Management Science,* 64, pp. 2709-2715.

Letina, I. and A. Schmutzler. (2019), “Inducing Variety: A Theory of Innovation Contests”, *International Economic Review*, forthcoming.

Liu, B., Lu, J., Wang, R. and Zhang, J. (2018), “Optimal Prize Allocation in Contests: The Role of Negative Prizes”, *Journal of Economic Theory*, 175, pp. 291-317.

Maskin, E. and J. Riley. (2000), “Asymmetric Auctions”, *Review of Economic Studies*, 67, pp. 413-438.

Milgrom, P. and C. Shannon. (1994), “Monotone Comparative Statics”, *Econometrica*, 62, pp. 157-180.

Moldovanu, B. and A. Sela. (2001), “The Optimal Allocation of Prizes in Contests”, *American Economic Review*, 91, pp. 542-558.

Myerson, R. B. (1981), “Optimal Auction Design”, *Mathematics of Operations Research*, 6, pp. 58-73.

Olszewski, W. and Siegel, R. (2018), “Performance-Maximizing Contests”, Working paper.

Pesendorfer, M. (2000), “A Study of Collusion in First-Price Auctions”, *Review of Economic Studies*, 67, pp. 381-411.

Quah, J. K. H. (2007), "The Comparative Statics of Constrained Optimization Problems", *Econometrica*, 75, pp. 401-431.

Shaked, M. and Shanthikumar, J. G. (2007), *Stochastic Orders*, Springer.

Taylor, C. R. (1995), "Digging for Golden Carrots: An Analysis of Research Tournaments", *American Economic Review*, 85, pp. 872-890.

Terwiesch, C. and Y. Xu. (2008), "Innovation Contests, Open Innovation, and Multiagent Problem Solving", *Management Science*, 54, pp. 1529-43.

Topkis, D. (1978), "Minimizing a Submodular Function on a Lattice", *Operations Research*, 26, pp. 305-321.

Appendix

A Omitted Proofs in Sections 4-6

Proof of Lemma 1. First, consider the case without the performance threshold, i.e., $t = 0$. Assume that all other solvers choose performance according to the function β_0, and assume that this function is strictly increasing and differentiable. The subscript represents the value of t. A solver's problem is

$$\max_x vF^{n-1}(\beta_0^{-1}(x)) - x/q$$

and the first order condition is

$$v\frac{dF^{n-1}(\beta_0^{-1}(x))}{dx} - \frac{1}{q} = 0$$

In equilibrium, the solver chooses performance level $x = \beta_0(q)$, so $q = \beta_0^{-1}(x)$. Denoting $\beta_0^{-1}(x)$ as y and substituting the expressions of y and q into the first order condition, we obtain

$$1 = vy\frac{dF^{n-1}(y)}{dx} \tag{A.1}$$

As a solver's idea quality $q \to 0$, his cost of any given performance level goes to infinity. Hence, the optimal performance level must converge to 0, and this yields the boundary condition $y(0) = 0$.

Note that the right hand side of (A.1) is a function of y, which means it is a differential equation with separated variables. Thus, its solution with initial condition $y(0) = 0$ is given by[29]

$$\int_0^x ds \quad = \quad v\int_0^y s\frac{dF^{n-1}(s)}{ds}ds$$

Denoting $H_0(y) = v\int_0^y s\frac{dF^{n-1}(s)}{ds}ds$, we can rewrite the above equation as $x = H_0(y) = H_0(\beta_0^{-1}(x))$, and therefore $\beta_0(x) = H_0(x)$. Thus, the performance function for every solver is

$$\beta_0(q) = v\int_0^q s\frac{dF^{n-1}(s)}{ds}ds = v\int_0^q sdF^{n-1}(s)$$

which is clearly strictly increasing and differentiable. Notice that the only possible atom point of F is at $s = 0$, so the size of the atom, $F(0)$, does not affect the above integral.

[29]See Arnold (1984), p.42 for a detailed discussion of differential equations with separated variables.

Assuming that all solvers other than i play according to β_0, we need to show that, for any idea quality q of solver i, the performance $\beta_0(q)$ maximizes the expected payoff corresponding to that idea quality. Let $\pi(x,q) = vF^{n-1}(\beta_0^{-1}(x)) - x/q$ be the expected payoff of solver i with idea q that chooses performance level x. We will show that derivative $\pi_x(x,q)$ is nonnegative if x is smaller than $\beta_0(q)$ and nonpositive if x is larger than $\beta_0(q)$. As $\pi(x,q)$ is continuous in x, this implies that $x = \beta_0(q)$ maximizes $\pi(x,q)$. Notice that

$$\pi_x(x,q) = v(n-1)F^{n-2}(\beta_0^{-1}(x))f(\beta_0^{-1}(x))\frac{d\beta_0^{-1}(x)}{dx} - \frac{1}{q}$$

Let $x < \beta_0(q)$, and let $\hat{q}$ be the idea quality of a solver who is supposed to choose performance x, i.e., $\beta_0(\hat{q}) = x$. Note that $\hat{q} < q$ because β_0 is strictly increasing. Differentiating $\pi_x(x,q)$ with respect to q yields $\pi_{xq}(x,q) = 1/q^2 > 0$. Since $\hat{q} < q$, we obtain $\pi_x(x,q) \geq \pi_x(x,\hat{q})$. Since $x = \beta_0(\hat{q})$ we obtain by the first order condition that $\pi_x(x,\hat{q}) = 0$, and therefore that $\pi_x(x,q) \geq 0$ for every $x < \beta_0(q)$. A similar argument shows that $\pi_x(x,q) \leq 0$ for every $x > \beta_0(q)$.

So far we have derived the equilibrium strategy β_0 when $t = 0$. Next, we consider the case with $t > 0$ and derive the symmetric equilibrium strategy β_t. Since a solver has to choose at least t to possibly win, it is optimal for a solver with a sufficiently low idea quality to choose zero and not participate. Assume that a solver does not participate if his idea quality is below some critical level $q_t > 0$ (to be derived below). This means $\beta_t(q) = 0$ for $q < q_t$ and $\beta_t(q) \geq t$ for $q \geq q_t$. If $\beta_t(q_t) > t$, by deviating to $x = t$, a solver can achieve the same probability of winning at a lower cost. Thus, $\beta_t(q_t) = t$.

It remains to characterize $\beta_t(q)$ for $q > q_t$. Assume that β_t is strictly increasing and differentiable for $q > q_t$. Suppose all other solvers follow strategy β_t. Then, if a solver's idea quality is at the critical level q_t, he must be indifferent between choosing performance 0 and t, i.e.

$$vF^{n-1}(q_t) - t/q_t = 0$$

which can be rewritten as $q_t F^{n-1}(q_t) = t/v$ and uniquely determines q_t.

For $x \geq t$, $y = \beta_t^{-1}(x)$. Then we can verify that the first order condition is the same as (A.1). Thus, its solution y with initial condition $y(t) = q_t$ is given by

$$\int_t^x ds \quad = \quad v\int_t^y s\frac{dF^{n-1}(s)}{ds}ds$$

In the same way to derive β_0, the above equation implies $\beta_t(q) = t + v\int_{q_t}^q s dF^{n-1}(s)$ for

$q \geq q_t$. Notice that function β_t is equal to β_0 for $t = 0$. ■

Proof of Lemma 2. The proof consists of two steps. First, we show that $d \log F(q_t)/dv$ is negative and increasing in v. To see this, notice that (2) implies

$$\frac{dq_t}{dv} = -\frac{t}{v^2} \frac{1}{F^{n-1}(q_t) + (n-1)q_t F^{n-2}(q_t) f(q_t)}$$

so

$$\begin{aligned}
\frac{d \log F(q_t)}{dv} &= \frac{f(q_t)}{F(q_t)} \frac{dq_t}{dv} \\
&= -\frac{f(q_t)}{F(q_t)} \frac{t}{v^2} \frac{1}{F^{n-1}(q_t) + (n-1)q_t F^{n-2}(q_t) f(q_t)} \\
&= -\frac{t}{v^2} \frac{1}{n-1} \frac{1}{F^{n-1}(q_t)} \left(q_t + \frac{1}{n-1} \frac{F(q_t)}{f(q_t)} \right)^{-1} \\
&= -\frac{q_t}{v} \frac{1}{n-1} \left(q_t + \frac{1}{n-1} \frac{F(q_t)}{f(q_t)} \right)^{-1} \qquad \text{(A.2)}
\end{aligned}$$

where the last equality is from $F^{n-1}(q_t) = t/(q_t v)$. If v increases, q_t decreases due to (2), and $q_t + \frac{1}{n-1} \frac{F(q_t)}{f(q_t)}$ increases due to Assumption 1. Therefore, $d \log F(q_t)/dv$ in (A.2) increases in v.

Second, we prove the lemma. According to (4), we have

$$\begin{aligned}
L'_F(v) &= F^n(q_t) - (1-v) \frac{dF^n(q_t)}{dv} - 1 \qquad \text{(A.3)} \\
&= F^n(q_t) \left(1 + (1-v) \left(-\frac{1}{F^n(q_t)} \frac{dF^n(q_t)}{dv} \right) \right) - 1 \\
&= F^n(q_t) \left(1 + (1-v) n \left(-\frac{d \log F(q_t)}{dv} \right) \right) - 1
\end{aligned}$$

If v increases, $F^n(q_t)$ and $1 - v$ decrease. $-d \log F(q_t)/dv$ also decreases because of the first step. Hence, $L'_F(v)$ decreases in v. ■

Proof of Proposition 1. The proof has four steps.

Step 1. There exists $\bar{\lambda}_F$ such that for any $\lambda > \bar{\lambda}_F$, there is v_λ such that $\Pi'_F(v) > 0$ for $v > v_\lambda$. This means function Π_F has an increasing tail if λ is large enough.

To see this, recall that if $\lambda = 0$, there is a unique optimal prize in $(\underline{v}_t, 1)$. Thus, if $\bar{\lambda}_F$

exists, it is larger than 0. Taking derivatives of both sides of (5) w.r.t. v, we obtain

$$\begin{aligned} K'_F(v) &= \int_{q_t}^{w_F} \left(\int_{q_t}^{q} s dF^{n-1}(s) \right) dF^n(q) - v q_t (1 - F^n(q_t)) \frac{dF^{n-1}(q_t)}{dv} \\ &= \int_{q_t}^{w_F} \int_{s}^{w_F} s dF^n(q) dF^{n-1}(s) - v q_t (1 - F^n(q_t)) \frac{dF^{n-1}(q_t)}{dv} \\ &= \int_{q_t}^{w_F} s(1 - F^n(s)) dF^{n-1}(s) - v q_t (1 - F^n(q_t)) \frac{dF^{n-1}(q_t)}{dv} \end{aligned} \tag{A.4}$$

where the second equation comes from changing the order of integration. To show the existence of $\bar{\lambda}_F$, notice that (2) implies $q_t \to 0$ as $v \to +\infty$, so (A.4) implies

$$\begin{aligned} \lim_{v \to +\infty} K'_F(v) &= \int_0^{w_F} s(1 - F^n(s)) dF^{n-1}(s) \\ &\quad + \lim_{v \to +\infty} \left[\left(-\frac{dF(q_t)}{dv} \right) v(n-1) q_t \frac{1 - F^n(q_t)}{F^{n-1}(q_t)} \right] \\ &\geq \frac{n}{t} \int_0^{w_F} s(1 - F^n(s)) dF^{n-1}(s) > 0 \end{aligned} \tag{A.5}$$

where the first inequality comes from $dF(q_t)/dv < 0$ established in the proof of Lemma 2.

Next, we show $\lim_{v \to +\infty} L'_F(v) = -1$. Intuitively, if the prize is large enough, every solver chooses performance above the threshold, so the marginal effect is simply the marginal cost of the prize. Formally, recall that in the proof of Lemma 2 we obtain

$$L'_F(v) = F^n(q_t) \left(1 + (1 - v) n \left(-\frac{d \log F(q_t)}{dv} \right) \right) - 1$$

Substituting (A.2) into this equation, we can rewrite it as

$$L'_F(v) = F^n(q_t) \left(1 + \frac{(1 - v) q_t}{v} \frac{n}{n-1} \left(q_t + \frac{1}{n-1} \frac{F(q_t)}{f(q_t)} \right)^{-1} \right) - 1$$

Rewrite (2) as $v = \frac{t}{q_t F^{n-1}(q_t)}$ and substitute it into the above expression. Then, we can rewrite $L'_F(v)$ as a function of q_t:

$$L'_F(v) = F^n(q_t) + \frac{n}{t} \frac{q_t^2 F^{2n-1}(q_t) f(q_t) - t q_t F^n(q_t) f(q_t)}{(n-1) q_t f(q_t) + F(q_t)} - 1$$

Notice that (2) implies $q_t \to 0$ as $v \to +\infty$, which means solvers with any positive idea quality participate if the prize is high enough. Therefore, the above equation implies

$$\begin{aligned}
\lim_{v\to+\infty} L'_F(v) &= \frac{n}{t}\lim_{q_t\to 0}\frac{q_t^2F^{2n-1}(q_t)f(q_t)-tq_tF^n(q_t)f(q_t)}{(n-1)q_tf(q_t)+F(q_t)}-1\\
&= \frac{n}{t}\frac{\lim_{q_t\to 0}q_t^2F^{2n-1}(q_t)f'(q_t)-t\lim_{q_t\to 0}q_tF^n(q_t)f'(q_t)}{nf(0)+(n-1)\lim_{q_t\to 0}q_tf'(q_t)}-1\\
&= -1 \qquad\qquad (A.6)
\end{aligned}$$

where the second equality is due to L'Hôpital's rule and the last follows from $f(0) > 0$ and $f'(0) < +\infty$.

Let $\bar{\lambda}_F = 1/\lim_{q_t\to 0}\int_{q_t}^{w_F} s(1-F^n(s))dF^{n-1}(s)$, which is in $(0,+\infty)$. Then, for $\lambda > \bar{\lambda}_F$,

$$\lim_{v\to+\infty} L'_F(v)+\lambda\lim_{v\to+\infty} K'_F(v) \;>\; \lim_{v\to+\infty} L'_F(v)+\bar{\lambda}_F\lim_{v\to+\infty} K'_F(v)\geq -1+1=0$$

where the second inequality follows from (A.5) and $\lim_{v\to+\infty} L'_F(v) = -1$. Thus, for any given $\lambda > \bar{\lambda}_F$, there is v_λ such that $\Pi'_F(v) > 0$ for $v > v_\lambda$.

Step 2. We show that $V_F(\lambda) = \bar{v}$ if $\lambda > \bar{\lambda}_F$ and $V_F(\lambda) < \bar{v}$ if $\lambda < \bar{\lambda}_F$. If $\lambda > \bar{\lambda}_F$, Step 1 above shows that function Π_F has an increasing tail. Then, the optimal prize is $\bar{v}$. If $\lambda < \bar{\lambda}_F$, the definition of $\bar{\lambda}_F$ implies $V_F(\lambda) < \bar{v}$.

Step 3. We show that there is a unique optimal prize if $\lambda < \bar{\lambda}_F$. Suppose this is not true. Then, there are two prizes $v, v' \in V_F(\lambda)$ for some $\lambda < \bar{\lambda}_F$. Step 2 above implies that v and v' are smaller than $\bar{v}$. Then, the first order conditions are

$$L'_F(v)+\lambda K'_F(v)=0$$
$$L'_F(v')+\lambda K'_F(v')=0$$

Moreover, the objective function must have the same value at v and v', i.e.,

$$L_F(v)+\lambda K_F(v)=L_F(v')+\lambda K_F(v')$$

Solving for λ from the first order conditions and substituting into the above equation, we obtain

$$L_F(v)-\frac{L'_F(v)}{K'_F(v)}K_F(v)=L_F(v')-\frac{L'_F(v')}{K'_F(v')}K_F(v')$$

However, this is impossible because $L_F(v) - \frac{L'_F(v)}{K'_F(v)}K_F(v)$ is strictly monotone due to Assumption 2.

Step 4. We prove the monotonicity of $V_F(\lambda)$. From Step 2, $V_F(\lambda) = \bar{v}$ for $\lambda > \bar{\lambda}_F$, so it is weakly increasing. For $\lambda < \bar{\lambda}_F$, there is a unique optimal prize because of Step 3, and it is interior because of Step 2. Thus, the optimal prize solves $\Pi'_F(v) = 0$. In addition, from their definitions, both q_t and $K_F(v)$ increase in v. Therefore, a higher λ shifts $\Pi'_F(v) = L'_F(v) + \lambda K'_F(v)$ upwards. Hence, $V_F(\lambda)$ strictly increases in λ for $\lambda < \bar{\lambda}_F$. ■

Proof of Lemma 3. The proof has two steps.

Step 1. We show that $\hat{F} \leq_{FOSD} \hat{G}$ implies $F \leq_{FOSD} G$. We first show that $F(0) \geq G(0)$. Suppose otherwise that $F(0) < G(0)$. Notice that $\lim_{x\to 0} \hat{F}(x) = \lim_{x\to 0} F(\phi_F^{-1}(x)) = F(0)$. Similarly, $\lim_{x\to 0} \hat{G}(x) = G(0)$. Thus, $F(0) < G(0)$ implies $\lim_{x\to 0} \hat{F}(x) < \lim_{x\to 0} \hat{G}(x)$, which contradicts $\hat{F} \leq_{FOSD} \hat{G}$.

Next, we prove that $F \leq_{FOSD} G$. Suppose otherwise that there exists $q' \in [0, w_F]$ such that $F(q') < G(q')$. Then, the first step and the intermediate value theorem imply that there is $q \in (0, w_F)$ such that $F(q) = G(q)$. Define $\hat{q} = \max\{q \in (0, q') | F(q) = G(q)\}$. Then, $F(\hat{q}) = G(\hat{q})$ and $F(\hat{q} + \varepsilon) < G(\hat{q} + \varepsilon)$ for sufficiently small $\varepsilon > 0$. Moreover, let $\hat{x} = \hat{q}F^{n-1}(\hat{q}) = \hat{q}G^{n-1}(\hat{q})$. Then, by their definitions, $\phi_F(\hat{q} + \varepsilon) < \phi_G(\hat{q} + \varepsilon)$ and $\phi_F^{-1}(\hat{x} + \Delta) > \phi_G^{-1}(\hat{x} + \Delta)$ for sufficiently small $\Delta > 0$. Recall that $\phi_F^{-1}(x)F^{n-1}(\phi_F^{-1}(x)) = x$ and $\phi_G^{-1}(x)G^{n-1}(\phi_G^{-1}(x)) = x$, so

$$\phi_F^{-1}(x)F^{n-1}(\phi_F^{-1}(x)) = \phi_G^{-1}(x)G^{n-1}(\phi_G^{-1}(x)).$$

Since $\phi_F^{-1}(\hat{x} + \Delta) > \phi_G^{-1}(\hat{x} + \Delta)$, the above equation implies $F^{n-1}(\phi_F^{-1}(\hat{x} + \Delta)) < G^{n-1}(\phi_F^{-1}(\hat{x} + \Delta))$, which is equivalent to $\hat{F}(\hat{x} + \Delta) < \hat{G}(\hat{x} + \Delta)$. This contradicts $\hat{F} \leq_{FOSD} \hat{G}$.

Step 2. We show that $F \leq_{FOSD} G$ implies $\hat{F} \leq_{FOSD} \hat{G}$. Because $F \leq_{FOSD} G$, we have $qF^{n-1}(q) \geq qG^{n-1}(q)$. Then the solution of

$$qF^{n-1}(q) = x \tag{A.7}$$

must be smaller than that of

$$qG^{n-1}(q) = x \tag{A.8}$$

That is, $\phi_F^{-1}(x) \leq \phi_G^{-1}(x)$. Notice that (A.7) and (A.8) imply

$$\phi_F^{-1}(x)F^{n-1}(\phi_F^{-1}(x)) = \phi_G^{-1}(x)G^{n-1}(\phi_G^{-1}(x))$$

so the above inequality implies $F^{n-1}(\phi_F^{-1}(x)) \geq G^{n-1}(\phi_G^{-1}(x))$, which combined with the definitions of $\hat{F}$ and $\hat{G}$ implies $\hat{F}(x) \geq \hat{G}(x)$. ■

Proof of Lemma 4. Using the definition of $\hat{F}$, we have

$$\frac{d \log \hat{F}(x)}{dx} = \frac{d \log F(\phi_F^{-1}(x))}{dx} = \frac{f(\phi_F^{-1}(x))}{F(\phi_F^{-1}(x))} \frac{d\phi_F^{-1}(x)}{dx}$$

Denote $q_x \equiv \phi_F^{-1}(x)$ and substitute it into the equation above. We obtain

$$\begin{aligned} \frac{d \log \hat{F}(x)}{dx} &= \frac{f(q_x)}{F(q_x)} \frac{1}{\phi'_F(q_x)} \\ &= \frac{f(q_x)}{F(q_x)} \frac{1}{F^{n-1}(q_x) + (n-1) q_x F^{n-2}(q_x) f(q_x)} \\ &= \frac{1}{n-1} \frac{1}{F^{n-1}(q_x)} \left(q_x + \frac{1}{n-1} \frac{F(q_x)}{f(q_x)} \right)^{-1} \end{aligned} \tag{A.9}$$

Assumption 1 implies that $q_x + \frac{1}{n-1} \frac{F(q_x)}{f(q_x)}$ increases in q_x for $n \geq 2$. Moreover, if x increases, q_x increases. Therefore, (A.9) decreases in x, which means $\hat{F}$ is log-concave. ■

Proof of Lemma 5. It is well-known that $\hat{F} \prec_{LR} \hat{G}$ implies $\hat{F} \prec_{FOSD} \hat{G}$, which, combined with Lemma 3, implies $F \prec_{FOSD} G$. ■

Proof of Lemma 8. Using equation (A.3), we can rewrite $L'_F(v) = 0$ as

$$F^n(q_t) - (1-v) \frac{dF^n(q_t)}{dv} - 1 = 0 \tag{A.10}$$

Similarly, we can rewrite $L'_G(v) < 0$ as

$$G^n(q'_t) - (1-v) \frac{dG^n(q'_t)}{dv} - 1 < 0 \tag{A.11}$$

Suppose $L'_F(v) = 0$. Then equation (A.10) implies

$$1 - v = (F^n(q_t) - 1) \left(\frac{dF^n(q_t)}{dv} \right)^{-1}$$

Substituting it into (A.11), we can rewrite $L'_G(v) < 0$ as

$$G^n(q'_t) - (F^n(q_t) - 1) \left(\frac{dF^n(q_t)}{dv} \right)^{-1} \frac{dG^n(q'_t)}{dv} - 1 < 0 \tag{A.12}$$

It remains to show (A.12). Recall that in the first step to prove Lemma 2, we obtain

$dF^n(q_t)/dv < 0$ and $dG^n(q_t')/dv < 0$, so (A.12) can be rewritten as

$$-\frac{dF^n(q_t)}{dv}\frac{1}{1-F^n(q_t)} > -\frac{dG^n(q_t')}{dv}\frac{1}{1-G^n(q_t')} \tag{A.13}$$

Recall that the definition of $F \prec G$ requires $\hat{F} \prec_{LR} \hat{G}$, which implies $\hat{F}^n \prec_{LR} \hat{G}^n$ due to Theorem 1.C.33 of Shaked and Shanthikumar (2007). Then, (A.13) holds because

$$LHS \text{ of (A.13)} = \frac{(\hat{F}^n)'(t/v)}{1-\hat{F}^n(t/v)}\frac{t}{v^2} > \frac{(\hat{G}^n)'(t/v)}{1-\hat{G}^n(t/v)}\frac{t}{v^2} = RHS \text{ of (A.13)}$$

where the equalities follow from the definitions of $\hat{F}$ and $\hat{G}$ and the inequality follows from the hazard rate dominance implied by $\hat{F}^n \prec_{LR} \hat{G}^n$. Hence, (A.12) also holds. ■

Proof of Lemma 9. If v increases, q_t decreases. Therefore, (5) implies that $K_F(v)$ is increasing in v.

We prove $K_F'(v) < K_G'(v)$ in two steps. First, note that

$$q_t(1-F^n(q_t))\frac{dF^{n-1}(q_t)}{dv} \geq q_t'(1-G^n(q_t'))\frac{dG^{n-1}(q_t')}{dv} \tag{A.14}$$

To see why, notice that equation (2) implies

$$q_t = \frac{t}{v}\frac{1}{F^{n-1}(q_t)}$$

Substituting this expression into (A.14), we have

$$\begin{aligned} LHS \text{ of (A.14)} &= \frac{t}{v}(1-F^n(q_t))\frac{dF^{n-1}(q_t)}{dv}\frac{1}{F^{n-1}(q_t)} \\ &= \frac{t}{v}(1-F^n(q_t))\frac{dF^{n-1}(q_t)}{d(t/v)}\frac{1}{F^{n-1}(q_t)}\left(-\frac{t}{v^2}\right) \\ &= \frac{t}{v}(1-F^n(q_t))\frac{d\hat{F}^{\frac{n-1}{n}}(t/v)}{d(t/v)}\frac{1}{\hat{F}^{\frac{n-1}{n}}(t/v)}\left(-\frac{t}{v^2}\right) \end{aligned} \tag{A.15}$$

Notice that $\hat{F} \prec_{LR} \hat{G}$ implies $\hat{G}$ dominates $\hat{F}$ in terms of reverse hazard rate. Therefore, $\hat{G}^{\frac{n-1}{n}}$ also dominates $\hat{F}^{\frac{n-1}{n}}$ in terms of reverse hazard rate. That is,

$$\frac{d\hat{F}^{\frac{n-1}{n}}(t/v)}{d(t/v)}\frac{1}{\hat{F}^{\frac{n-1}{n}}(t/v)} \leq \frac{d\hat{G}^{\frac{n-1}{n}}(t/v)}{d(t/v)}\frac{1}{\hat{G}^{\frac{n-1}{n}}(t/v)}$$

As a result, equation (A.15) implies

$$LHS \text{ of } (A.14) \geq \frac{t}{v}(1 - F^n(q_t))\frac{d\hat{G}^{\frac{n-1}{n}}(t/v)}{d(t/v)}\frac{1}{\hat{G}^{\frac{n-1}{n}}(t/v)}\left(-\frac{t}{v^2}\right) \geq RHS \text{ of } (A.14)$$

Second, we use equation (A.14) to prove Lemma 9. Following the same argument as for (A.4), we have

$$K_G'(v) = \int_{q_t'}^{w_G} s(1 - G^n(s))dG^{n-1}(s) - vq_t'(1 - G^n(q_t'))\frac{dG^{n-1}(q_t')}{dv}$$

Because of (A.14), it is sufficient to show

$$\int_{q_t}^{w_F} s(1 - F^n(s))dF^{n-1}(s) \leq \int_{q_t'}^{w_G} s(1 - G^n(s))dG^{n-1}(s) \tag{A.16}$$

Lemma 3 implies that

$$G(q) \leq F(q) \tag{A.17}$$

Moreover, we have $q_t < q_t'$. Therefore,

$$\int_{q_t'}^{w_G} G^{n-1}(s)ds \leq \int_{q_t}^{w_F} F^{n-1}(s)ds \tag{A.18}$$

We also have

$$\begin{aligned}\int_{q_t}^{w_F} sdF^{n-1}(s) &= w_F - q_tF^{n-1}(q_t) - \int_{q_t}^{w_F} F^{n-1}(s)ds \\ &= w_F - \frac{t}{v} - \int_{q_t}^{w_F} F^{n-1}(s)ds \end{aligned} \tag{A.19}$$

where the first equality comes from integration by parts and the second one comes from (2). Similarly,

$$\int_{q_t'}^{w_G} sdG^{n-1}(s) = w_G - \frac{t}{v} - \int_{q_t'}^{w_G} G^{n-1}(s)ds \tag{A.20}$$

Therefore,

$$\begin{aligned}\int_{q_t}^{w_F} s dF^{n-1}(s) &= w_F - \frac{t}{v} - \int_{q_t}^{w_F} F^{n-1}(s)ds \\ &\leq w_G - \frac{t}{v} - \int_{q'_t}^{w_G} G^{n-1}(s)ds \\ &= \int_{q'_t}^{w_G} s dG^{n-1}(s) \end{aligned} \tag{A.21}$$

where the first equality comes from (A.19), the inequality comes from $w_F \leq w_G$ and (A.18), and the last equality comes from (A.20). Note that if w_G or $w_F = +\infty$, the above analysis applies as well.

Therefore, (A.17) implies $1 - F^n(s) \leq 1 - G^n(s)$, which combined with (A.21) implies (A.16). ■

Proof of Lemma 10. We prove in three steps. First, $\lim_{v\to+\infty} K'_F(v) < \lim_{v\to+\infty} K'_G(v)$. Because $\lim_{v\to+\infty} q_t = \lim_{v\to+\infty} q'_t = 0$, as in the proof of Lemma 9, it is sufficient to show that

$$\int_0^{w_F} s(1 - F^n(s))dF^{n-1}(s) < \int_0^{w_G} s(1 - G^n(s))dG^{n-1}(s) \tag{A.22}$$

Because $G(q) \leq F(q)$ and the inequality holds strictly for a set of positive measure, we have

$$\int_0^{w_G} G^{n-1}(s)ds < \int_0^{w_F} F^{n-1}(s)ds$$

Notice that this inequality is analogous to (A.18), which is used to prove (A.16). Hence, we can use this inequality to prove (A.22) following the same argument used to prove (A.16).

Second, $\lim_{v\to+\infty}(L'_F(v) + \bar{\lambda}_G K'_F(v)) < 0$. By the definition of $\bar{\lambda}_G$, we have

$$\lim_{v\to+\infty} L'_G(v) + \bar{\lambda}_G \lim_{v\to+\infty} K'_G(v) = 0 \tag{A.23}$$

Recall that $\lim_{v\to+\infty} K'_G(v) > 0$ due to (A.5), so (A.23) implies

$$\bar{\lambda}_G = -\frac{\lim_{v\to+\infty} L'_G(v)}{\lim_{v\to+\infty} K'_G(v)} \tag{A.24}$$

Hence, we have

$$\begin{aligned}
\lim_{v\to+\infty} L'_F(v) + \bar{\lambda}_G \lim_{v\to+\infty} K'_F(v) &= \lim_{v\to+\infty} L'_F(v) - \lim_{v\to+\infty} L'_G(v) \frac{\lim_{v\to+\infty} K'_F(v)}{\lim_{v\to+\infty} K'_G(v)} \\
&< \lim_{v\to+\infty} L'_F(v) - \lim_{v\to+\infty} L'_G(v) \\
&= 0
\end{aligned}$$

where the inequality follows from $\lim_{v\to+\infty} K'_F(v) < \lim_{v\to+\infty} K'_G(v)$ in the first step and $0 < \lim_{v\to+\infty} K'_F(v)$ in (A.5), and the last equality follows from (A.6).

Third, we prove the lemma. In the proof of Proposition 1, we obtain $\bar{\lambda}_F \in (0, +\infty)$, so similarly, $\bar{\lambda}_G \in (0, +\infty)$. In addition, by the definition of $\bar{\lambda}_F$, we have $\lim_{v\to+\infty}(L'_F(v) + \bar{\lambda}_F K'_F(v)) = 0$. Therefore, the second step above implies $\bar{\lambda}_G < \bar{\lambda}_F$. Hence, $0 < \bar{\lambda}_G < \bar{\lambda}_F < +\infty$. ■

Proof of Proposition 3. The proposition follows from Steps I and II in the proof of Proposition 2. Notice that in these two steps, we do not use Assumptions 2 and 3. ■

B Proof of the Properties stated in Footnote 15

We prove the properties through Lemmas A.1-A.4.[30]

Lemma A.1 *If $F \prec G$, then $q_t < q'_t$ and $F(q_t) > G(q'_t)$.*

Proof. The definition of $\hat{F} \prec_{LR} \hat{G}$ implies $\hat{F}(q) \geq \hat{G}(q)$, which combined with Lemma 3 implies $F(q) \geq G(q)$. Therefore, we have $\phi_F \geq \phi_G$, so $q_t \leq q'_t$. Since (2) and (8) imply $q_t F^n(q_t) = q'_t G^n(q'_t)$, inequality $q_t \leq q'_t$ implies $F(q_t) \geq G(q'_t)$. ■

Lemma A.2 $L'_F(1) > L'_G(1)$.

Proof. Suppose $v = 1$, equation (A.3) implies $L'_G(v) - L'_F(v) = G^n(q'_t) - F^n(q_t) < 0$, where the inequality comes from Lemma A.1. ■

Lemma A.3 *$F \prec G$ implies that $G(q'_t)/F(q_t)$ decreases in v.*

[30] We use different numbering for lemmas which only appear in the appendix.

Proof.

$$
\begin{aligned}
\frac{d}{dv}\left(\log\frac{G(q'_t)}{F(q_t)}\right) &= \frac{d\log G(q'_t)}{dv} - \frac{d\log F(q_t)}{dv} \\
&= \frac{d\log \hat{G}(t/v)}{d(t/v)}\left(-\frac{t}{v^2}\right) - \frac{d\log \hat{F}(t/v)}{d(t/v)}\left(-\frac{t}{v^2}\right) \\
&= \frac{\hat{g}(t/v)}{\hat{G}(t/v)}\left(-\frac{t}{v^2}\right) - \frac{\hat{f}(t/v)}{\hat{F}(t/v)}\left(-\frac{t}{v^2}\right) \\
&< 0
\end{aligned}
$$

where the inequality comes from the reverse hazard rate dominance in the definition of $F \prec G$. ■

Lemma A.4 *Suppose $F \prec G$ and F and G have a common support $[0,1]$. Then, $L'_F(t) < L'_G(t)$.*

Proof. If $v = t$, equation (2) becomes $q_t F^{n-1}(q_t) = 1$. Since the LHS is strictly increasing in q_t, this equation has a unique solution $q_t = 1$. Then, the common support $[0,1]$ implies

$$
q_t = q'_t = F(q_t) = G(q'_t) = 1 \tag{A.25}
$$

If $v = t$, we also have

$$
\frac{dF^n(q_t)}{dv} = \frac{d\hat{F}^n(t/v)}{d(t/v)}\left(-\frac{t}{v^2}\right) = -\frac{n}{t}\hat{f}(1) > -\frac{n}{t}\hat{g}(1) = \frac{dG^n(q'_t)}{dv} \tag{A.26}
$$

where the second equality comes from $v = t$ and the inequality comes from $\hat{f}(1) < \hat{g}(1)$ implied by the reverse hazard rate dominance in the definition of $F \prec G$.

Therefore,

$$
\begin{aligned}
L'_G(v) - L'_F(v) &= [G^n(q'_t) - F^n(q_t)] - (1-v)\left(\frac{dG^n(q'_t)}{dv} - \frac{dF^n(q_t)}{dv}\right) \\
&= (1-1) - (1-v)\left(\frac{dG^n(q'_t)}{dv} - \frac{dF^n(q_t)}{dv}\right) \\
&> 0
\end{aligned}
$$

where the second equality comes from (A.25) and the inequality comes from (A.26). ■

C Omitted Proofs in Section 7.1

We denote the optimal prizes for nonlinear benefits as $V_{F,B}$ and $V_{G,B}$ for distributions F and G, respectively. The following result generalizes Lemma 10 to the case of nonlinear benefits.

Lemma A.5 *Under Assumptions 1 and 2′, if $F \prec G$, then $0 < \bar{\lambda}_{G,B} < \bar{\lambda}_{F,B} < +\infty$, where $\bar{\lambda}_{F,B} = \inf\{\lambda | V_{F,B}(\lambda) = \bar{v}\}$ and $\bar{\lambda}_{G,B} = \inf\{\lambda | V_{G,B}(\lambda) = \bar{v}\}$.*

Proof. First we show $\lim_{v\to+\infty} K'_{F,B}(v) = \underline{b} \lim_{v\to+\infty} K'_F(v)$. Notice that the definition of $K_{F,B}$ implies

$$K'_{F,B}(v) = \int_{q_t}^{w_F} B'\left(v\int_{q_t}^{q} sdF^{n-1}(s)\right)\left(\int_{q_t}^{q} sdF^{n-1}(s) - vq_t\frac{dF^{n-1}(q_t)}{dv}\right)dF^n(q)$$

so

$$\begin{aligned}
\lim_{v\to+\infty} K'_{F,B}(v) &= \underline{b}\int_0^{w_F}\left(\int_0^{q} sdF^{n-1}(s) + \lim_{v\to+\infty} vq_t\frac{dF^n(q_t)}{dv}\right)dF^n(q) \\
&= \underline{b}\int_0^{w_F}\left(\int_0^{q} sdF^{n-1}(s)\right)dF^n(q) \\
&= \underline{b}\int_0^{w_F} s(1-F^n(s))dF^{n-1}(s) \qquad \text{(A.27)}
\end{aligned}$$

where the second equation comes from $\lim_{v\to\infty} vq_t dF^n(q_t)/dv = 0$, and the last from changing the order of integration. In addition, (A.4) implies that

$$\begin{aligned}
\lim_{v\to+\infty} K'_F(v) &= \int_0^{w_F} s(1-F^n(s))dF^{n-1}(s) - \lim_{v\to+\infty} vq_t(1-F^n(q_t))\frac{dF^{n-1}(q_t)}{dv} \\
&= \int_0^{w_F} s(1-F^n(s))dF^{n-1}(s) \qquad \text{(A.28)}
\end{aligned}$$

where the second equation is also from $\lim_{v\to\infty} vq_t dF^n(q_t)/dv = 0$. Hence, (A.27) and (A.28) imply that $\lim_{v\to+\infty} K'_{F,B}(v) = \underline{b}\lim_{v\to+\infty} K'_F(v)$.

Next, we prove the lemma. Notice that the proof of Lemma 10 only uses $\lim_{v\to+\infty} K'_F(v)$. Because of the first step above, the proof of Lemma 10 applies to nonlinear benefit function B as well. ■

Proposition A.1 *Under Assumptions 1 and 2′, there is a unique optimal prize $V_{F,B}(\lambda)$ for any $\lambda \neq \bar{\lambda}_{F,B}$. Moreover, $V_{F,B}(\lambda)$ is weakly increasing in λ.*

The proof is the same as that of Proposition 1 and is therefore omitted. The following result generalizes Proposition 2 to nonlinear benefit functions using a similar kind of argument.

Proposition A.2 *Under Assumptions* 1, 2′ *and* 3′, *if* $F \prec G$, *there exists* $\hat{\lambda} > 0$ *such that*

i) $V_{G,B}(\lambda) < V_{F,B}(\lambda)$ *if* $\lambda < \hat{\lambda}$

ii) $V_{G,B}(\lambda) \geq V_{F,B}(\lambda)$ *if* $\lambda > \hat{\lambda}$

Proof. If $\lambda = 0$, the expected profit is $\Pi_{F,B}(v) = L_F(v)$, which is the same as the expected profit in Section 4. Therefore, following the same argument as in the proof of Proposition 2, we have $V_F(\lambda) > V_G(\lambda)$ if λ is small.

Replacing Lemma 10 with Lemma A.5, we can follow the same arguments as in Steps II and III of the proof of Proposition 2 to prove this proposition. ■

Assumptions 2′ and 3′ ensure that the optimal prize is unique and that there is a unique $\hat{\lambda}$ at which the order of the optimal prizes switches. Without Assumption 2′, there may be multiple optimal prizes. Then, let $V_{F,B}(\lambda)$ and $V_{G,B}(\lambda)$ represent the set of optimal prizes. The following proposition states that a result similar to Proposition 3 holds. The proof is the same as that of Proposition 3, so we omit it here.

Proposition A.3 *Under Assumption* 1, *if* $F \prec G$, *there exist* $\hat{\lambda}' \geq \hat{\lambda} > 0$ *such that*

i) $V_{G,B}(\lambda) \leq V_{F,B}(\lambda)$ *if* $\lambda < \hat{\lambda}$

ii) $V_{G,B}(\lambda) \geq V_{F,B}(\lambda)$ *if* $\lambda > \hat{\lambda}'$

D Omitted Proofs in Section 7.2

Lemma A.6 *A reservation performance* $r < t$ *is never optimal.*

Proof. Suppose the seeker chooses $r < t$. The seeker's expected payoff is

$$\int_{q_t}^{w_F} [1 + \lambda(\beta(q) - t)] dF^n(q) - v(1 - F^n(q_r))$$

The first term is her gross profit from the best performance and the second term is the cost of giving up the prize v. Note that the solvers with ideas $q_i \in [r, t)$ choose performances above r, but their performances are worthless to the seeker because they are still below t. Thus, the integration in the first term is for $q \geq q_t$.

Note that q_r decreases as r increases. Thus, the gross profit (the first term) remains the same, but the cost (the second term) is lower. This means an increase in r leads to an increase in the seeker's profit. Hence, $r < t$ is never optimal. ■

Derivation of equation (12) According to the equilibrium strategies, the expected highest performance is

$$\begin{aligned}
\int_{q_r}^{w_F} \beta(q) dF^n(q) &= \int_{q_r}^{w_F} \left(r + v \int_{q_r}^{q} s dF^{n-1}(s) \right) dF^n(q) \\
&= \int_{q_r}^{w_F} \left(v q_r F^{n-1}(q_r) + v \int_{q_r}^{q} s dF^{n-1}(s) \right) dF^n(q) \\
&= \int_{q_r}^{w_F} \left(v q F^{n-1}(q) - v \int_{q_r}^{q} F^{n-1}(s) ds \right) dF^n(q) \\
&= v \int_{q_r}^{w_F} \left[q F^{n-1}(q) (F^n)'(q) - F^{n-1}(q)(1 - F^n(q) \right] dq \\
&= v \int_{q_r}^{w_F} F^{n-1}(q) \left(q - \frac{1 - F^n(q)}{(F^n)'(q)} \right) dF^n(q)
\end{aligned}$$

where the second equation is from (11), the third from integration by parts, and the fourth from changing the order of integration. Substituting the above expression into the seeker's expected profit, we can rewrite it as in (12).

Proof of Lemma 11. If $\lambda = 0$, the expected profit is $L_F(v, q_r) = (1 - v)(1 - F^n(q_r))$ with $q_r \geq q_t$, which is decreasing in q_r. Therefore, if (v, r) is optimal, it must ensure $q_r = q_t$. This means that $r = t$ whatever the optimal v is. ■

Proof of Proposition 4. We prove in three steps.

Step I. We prove part i). Specifically, Lemma 11 implies that $\partial \Pi_F(v, q_r)/\partial q_r < 0$ if $\lambda = 0$. Because $\partial \Pi_F(v, q_r)/\partial q_r$ is continuous in λ, there exists $\lambda' > 0$ such that $\partial \Pi_F(v, q_r)/\partial q_r < 0$ for $\lambda \in [0, \lambda']$. Then, for these λ values, the optimal q_r is at the lower bound q_t, which implies the optimal r equals t. Hence, Proposition 1 implies that $V_F(\lambda)$ is monotone non-decreasing in λ if $\lambda \leq \lambda'$.

Step II. The optimal q_r is bounded away from w_F as $\lambda \to +\infty$. Suppose not. That is, suppose that $q_r \to w_F$ as $\lambda \to +\infty$. Then,

$$\begin{aligned}
& \lim_{\lambda \to +\infty} \frac{\int_{q_r}^{w_F} \left[\lambda F^{n-1}(q) \left(q - \frac{1 - F^n(q)}{(F^n)'(q)} \right) \right] dF^n(q)}{1 - F^n(q_r)} \\
\geq & \lim_{\lambda \to +\infty} \frac{\lambda F^{n-1}(q_r) \left(q_r - \frac{1 - F^n(q_r)}{(F^n)'(q_r)} \right) \int_{q_r}^{w_F} dF^n(q)}{1 - F^n(q_r)} \\
= & \lim_{\lambda \to +\infty} \lambda F^{n-1}(q_r) \left(q_r - \frac{1 - F^n(q_r)}{(F^n)'(q_r)} \right) \\
= & +\infty
\end{aligned}$$

where the inequality follows from Assumption 4 and the last equality follows from the assumption that $q_r \to w_F$ as $\lambda \to +\infty$ and $\lim_{q \to w_F}\left(q - \frac{1-F^n(q)}{(F^n)'(q)}\right) > 0$ from Assumption 4.

Recall that the seeker chooses (v, q_r) and their marginal profits are

$$\frac{\partial \Pi_F(v, q_r)}{\partial v} = \int_{q_r}^{w_F} \left[\lambda F^{n-1}(q)\left(q - \frac{1-F^n(q)}{(F^n)'(q)}\right) - 1\right] dF^n(q) \tag{A.29}$$

$$\frac{\partial \Pi_F(v, q_r)}{\partial q_r} = -\left\{1 + \lambda\left[vF^{n-1}(q_r)\left(q_r - \frac{1-F^n(q_r)}{(F^n)'(q_r)}\right) - t\right] - v\right\}(F^n)'(q_r) \tag{A.30}$$

Then, if λ is large enough, (A.29) implies that $\partial \Pi_F(v, q_r)/\partial v > 0$ for any v, so $v = \bar{v}$. Recall that Assumption 4 implies that $\lim_{q \to w_F}\left(q - \frac{1-F^n(q)}{(F^n)'(q)}\right) > 0$ and denote the limit as $z > 0$. Then, if λ is large enough, (A.30) implies that $\partial \Pi_F(v, q_r)/\partial q_r$ for large enough q_r has the same sign as

$$-\lambda\left[\bar{v}\lim_{q_r \to w_F} F^{n-1}(q_r)\left(q_r - \frac{1-F^n(q_r)}{(F^n)'(q_r)}\right) - t\right] = -\lambda\left[\bar{v}z - t\right] < 0$$

where the equality follows from the large $\bar{v}$. Thus, q_r is bounded away from w_F as $\lambda \to +\infty$.

Step III. If λ is large enough, the optimal is $\bar{v}$ and the optimal r is weakly decreasing in λ. To see this, notice that q_r is bounded away from w_F, so for large enough λ, $\partial \Pi_F(v, q_r)/\partial v > 0$ for all v. Hence, the optimal prize is $\bar{v}$. With $v = \bar{v}$ in (A.30), a higher λ shifts $\partial \Pi_F(v, q_r)/\partial q_r$ downwards, which results in a lower q_r. Intuitively, because v cannot be increased further, we need to lower r in order to increase the profits. ■

Proof of Lemma 12. We know from above that if λ is large enough, the optimal prize is $\bar{v}$. With $v = \bar{v}$, the marginal profit in (A.30) at $q_r = q_t$ is

$$\frac{\partial \Pi_F(\bar{v}, q_t)}{\partial q_r} = \left\{\bar{v} - 1 + \lambda \bar{v} F^{n-1}(q_t)\frac{1-F^n(q_t)}{(F^n)'(q_t)}\right\}(F^n)'(q_t) > 0$$

which means the optimal $r > t$ whenever optimal $v = \bar{v}$. ■

Proof of Proposition 5. If λ is small enough, the optimal $r = t$, so Proposition 2 implies part i).

Notice that Lemma 12 implies that if λ is large enough, the optimal $r > t$. Therefore, the optimal q_r satisfies the first order condition $\partial \Pi_F(\bar{v}, q_r)/\partial q_r = 0$, which is equivalent to

$$F^{n-1}(q_r)\left(q_r - \frac{1-F^n(q_r)}{(F^n)'(q_r)}\right) = \frac{1}{\lambda} - \frac{1}{\lambda\bar{v}} + \frac{t}{\bar{v}} \tag{A.31}$$

The left hand side of the above equation is increasing in q_r due to Assumption 4, so the above equation has a unique solution. Let $q_F(\lambda)$ be the solution. Then, the corresponding reservation performance is $R_F(\lambda) = \bar{v}q_F(\lambda)F^{n-1}(q_F(\lambda))$. As $\lambda \to +\infty$, (A.31) becomes (13), and $q_F(\lambda)$ decreases and converges to q_F which is the solution of (13). As a result, $\lim_{\lambda\to+\infty} R_F(\lambda) = \bar{v}q_F F^{n-1}(q_F)$. Thus, part ii) of the proposition is also true. ■

E Parametric Distributions

Recall that Lemma 5 shows that $F \prec G$ implies $F \prec_{FOSD} G$. We show below that $F \prec_{FOSD} G$ implies $F \prec G$ for many distribution families. First, note that for F and G within the following distribution families, $\hat{F}(q)$ and $\hat{G}(q)$ have closed-form expressions.

- **Uniform Distributions** Consider $F(q) = q/w_F$ and $G(q) = q/w_G$ with $0 < w_F < w_G$. The supports are $[0, w_F]$ for F and $[0, w_G]$ for G. Then, $\phi_F(q) = q(q/w_F)^{n-1}$, so $\phi_F^{-1}(x) = (w_F)^{\frac{n-1}{n}} x^{\frac{1}{n}}$ and $\hat{F}(x) = (x/w_F)^{1/n}$. Similarly, $\hat{G}(x) = (x/w_G)^{1/n}$. Then, it is straightforward to verify $F \prec G$.

- **Power Function Distributions** Consider $F(q) = q^\alpha$ and $G(q) = q^{\alpha'}$ with $q \in [0,1]$ and $0 < \alpha < \alpha'$. Then, $\phi_F(q) = q^{(n-1)\alpha+1}$, so $\phi_F^{-1}(x) = x^{\frac{1}{(n-1)\alpha+1}}$ and $\hat{F}(x) = x^{\frac{\alpha}{(n-1)\alpha+1}}$. Similarly, $\hat{G}(x) = x^{\frac{\alpha'}{(n-1)\alpha'+1}}$. Thus, $F \prec G$.

For F and G within the following distribution families, $\hat{F}(q)$ and $\hat{G}(q)$ do not have closed-form expressions, so we verify $F \prec G$ numerically.

- **Pareto Distributions** Consider the c.d.f. $F(q;\alpha) = 1 - (1+q)^\alpha$ for $q \in [0, +\infty)$ with parameter $\alpha < -1$. We focus on $\alpha < -1$ to ensure that the mean $-\frac{1}{1+\alpha}$ is finite. Then, $F(\cdot;\alpha') \prec_{FOSD} F(\cdot;\alpha'')$ if and only if $\alpha' \leq \alpha''$. Notice that the density of effective ideas also depends on parameters α, so we write it as $\hat{f}(x;\alpha)$. A sufficient condition for $F(\cdot;\alpha') \prec F(\cdot;\alpha'')$ is $\frac{\partial^2 log \hat{f}(x;\alpha)}{\partial x \partial \alpha} > 0$ for all $\alpha \in (\alpha', \alpha'')$ and for all $x > 0$. We verify this condition for $-100 \leq \alpha \leq -1.05$ and $2 \leq n \leq 100$. Thus, for those α and n values, the first order stochastic dominance and the stochastic order in Definition 2 are equivalent.

- **Exponential Distributions** Consider the c.d.f. $F(q;\alpha) = 1 - e^{\alpha q}$ for $q \in [0, +\infty)$ and parameter $\alpha < 0$. As above, $F(\cdot;\alpha') \prec_{FOSD} F(\cdot;\alpha'')$ if and only if $\alpha' \leq \alpha''$.

We can verify the sufficient condition that $\frac{\partial^2 log\hat{f}(x;\alpha)}{\partial x \partial \alpha} > 0$ for $0.1 \le \alpha \le 20$ and $2 \le n \le 100$. Thus, for those α and n values, the first order stochastic dominance and the stochastic order in Definition 2 are equivalent.

- **Log-normal Distributions** Let $F(q; \mu, \sigma^2)$ for $q \ge 0$ be the c.d.f. of log-normal distribution with mean $\mu > 0$ and variance $\sigma^2 > 0$. For any fixed σ^2, we have $F(\cdot; \mu', \sigma^2) \prec_{FOSD} F(\cdot; \mu'', \sigma^2)$ if and only if $\mu' \le \mu''$. As above, we can verify the sufficient condition that $\frac{\partial^2 log\hat{f}(x;\mu,\sigma^2)}{\partial x \partial \mu} > 0$ for $0.1 \le \mu \le 20$ and $2 \le n \le 100$. Thus, for those μ and n values, the first order stochastic dominance and the stochastic order in Definition 2 are equivalent.