# Skewing the odds: Strategic risk taking in contests*

Dawei Fang
Department of Economics, University of Gothenburg

Thomas Noe
Saïd Business School & Balliol College, University of Oxford

5 February, 2015

**Abstract:** We study contests where, subject only to a capacity constraint on mean performance, contestants compete for identical prizes by choosing random performance levels. The capacity constraint combined with the rank-contingent rewards makes win-small/lose-big strategies optimal. Equilibrium strategies are generally skewed but never risk maximizing. When capacity is known and symmetric, we derive a closed-form solution for the game and analyze the effects of contest selectivity and size on equilibrium outcomes. We next consider contests where capacity is private information and show that, contrary to the risk-taking-and-ruin intuition, weak contestants do not always gamble on high-risk strategies. In fact, when the capacities of weak and strong contestants differ sufficiently, weak contestants choose low-variance performance distributions that never top strong contestants' performance, ensuring that the equilibrium is perfectly selection efficient. Finally, we consider the implications of our analysis for mutual fund tournaments, R&D competition, and stochastic contests.

*Keywords:* contests, risk taking, rank-based selection, skewness preference, mutual fund tournaments, R&D competition, stochastic contests, relaxed Colonel Blotto game

*We thank Stefan Ankirchner, Alexander Herbertsson, Martin Holmén, Chris Jones, Jurjen Kamphorst, Margaret Meyer, John Quah, Christian Seel, and Casper de Vries for helpful comments, and seminar participants at EEA-ESEM 2014, Erasmus School of Economics, University of Gothenburg, MEA 2014, University of Oxford, and RES 2014 Conference. Any remaining errors are our responsibility.

# 1 Introduction

Contests where contestants compete for a fixed number of prizes using a fixed bundle of resources are ubiquitous. Many of these contests have a "use-it-or-lose-it" character, i.e., the resources which could be applied to win the contest, if not applied, are either lost or have very little residual value. To win a prize, contestants are willing to devote all their resources into competition and they have to fight with the resources they have. Thus, resources represent a contestant's capacity. Higher capacity gives a contestant an advantage in competition but does not guarantee victory, since contestants with lower capacity can adopt risky strategies that make victory possible. History offers many examples. For instance, in the Battle of the Granicus, Alexander the Great, with a force of at most ten thousand Macedonians, defeated a Persian army of at least one hundred thousand through an uphill, frontal, cross-river cavalry charge. The Persians, rightly believing that such a charge was extremely risky, even foolhardy, made no preparations to defend against it, and thus were surprised by the Macedonian attack and utterly routed (Siculus (1718)).

Strategic risk taking can take various forms.[1] A portfolio manager, aiming to be a top 10 manager, changes the return distribution by changing the portfolio composition. A research unit, competing against other research units in the speed of innovation, varies the probability distribution of discovery dates by varying the riskiness of the research process. A politician, aiming to win an election, considers how to deploy a fixed campaign war chest over alternative advertising strategies. A student, reviewing for an exam, decides how to allocate his limited time and attention into different topics.

In these settings, it is natural to represent risk-taking strategies as the choices of probability distributions over realized performance and capacities as upper bounds on the expectations of these distributions. Winning a prize depends only on the rank of a contestant's performance relative to the performance of his competitors. This approach has been widely adopted in the literature to model strategic risk taking in rank-order contests under various contexts: mutual fund tournaments (Taylor (2003), Chen, Hughson, and Stoughton (2012)), R&D competition (Bhattacharya and Mookherjee (1986) and Klette and de Meza (1986)), electoral campaigns (Myerson (1993), Lizzeri (1999), Lizzeri and Persico (2001), and Sahuguet and Persico (2006)), sales contests (Gaba and Kalra (1999)), and promotion contests (Hvide (2002), Hvide and Kristiansen (2003), Goel and Thakor (2008), Gilpatric (2009), and Han, Hirshleifer, and Persons (2009)).

[1] The focus of this paper is on strategic risk taking, that is, manipulable risks. This approach contrasts with the contest models featuring exogenous uncertainties produced by non-manipulable additive noises (Lazear and Rosen (1981)) or pre-specified contest success functions (Tullock (1980)). See Konrad (2007) and Van Long (2013) for a brief review of the exogenous-uncertainty contest models.

While this literature has yielded interesting results, the models considered often reduce risk choice to variance choice by assuming that contestant performance distributions are symmetric.[2] Under this assumption, papers in this literature reach one common conclusion: if risk taking is costless, the optimal risk level is always a corner solution, either the safest or the riskiest feasible strategy.

Academics and practitioners measure risk by variance, and the insights from Dubins and Savage (1965) on optimal gambling strategies in games with unfair odds also suggest that variance is the key parameter for risk taking. However, we show, in this paper, that restricting attention to variance is not without loss of generality. Contests are fundamentally different from casino gambles. In contrast to casino gambles, for a contestant in a rank-order contest, there is no extra gain from winning big as opposed to merely winning; there is no extra loss from losing big as opposed to merely losing. Given the existence of a capacity constraint, an asymmetric win-small/lose-big strategy is optimal against any predictable competitor performance level.[3] Because symmetric distribution models force desired tail risk on one side of the distribution to be accompanied by undesired tail risk on the other side, they force contestants with skewness preference to exhibit variance preference.[4]

In this paper, we relax the symmetry assumption. The only restrictions we impose on the performance distributions are that realized performance is nonnegative and expected performance never exceeds capacity. Our model is also flexible along a number of other dimensions: the number of contestants, the number of prizes, and, when information is incomplete, the variance of contestants' capacities.

We start the analysis in Section 2 by investigating the simplest contest in which two equally matched contestants compete for one prize. We show that, if one contestant plays a symmetric strictly unimodal distribution, the other can always obtain a probability of winning strictly greater than one half by adopting a win-small/lose-big strategy. Since a contestant's best reply will never produce a winning probability less than one half, the win-small/lose-big strategy prevents strictly unimodal distributions from ever being best replies in the two-contestant case and leads to an equilibrium in which both contestants submit uniformly distributed performance levels.

However, when more than two contestants compete for the prize, uniformly distributed performance levels would generate a probability of winning function that was

[2]The Colonel-Blotto electoral campaign models introduced by Myerson (1993) are an exception.

[3]Kamenica and Gentzkow (2011) show that a strategy analogous to win-small/lose-big strategies in our paper is optimal in a model of rational Bayesian persuasion where the sender maximizes the probability that the receiver's posterior probability belief exceeds a fixed threshold. The optimal signal produces posterior beliefs in the receiver that either meet the threshold or are very low.

[4]Dijk, Holmén, and Kirchler (2014) show experimentally that contestants have skewness preference.

unimodal which again would be bested by a win-small/lose-big strategy.[5] To prevent this from occurring, the probability weight placed on high performance levels must decrease, i.e., the performance distribution must become right skewed. This result suggests, and our model verifies, that, typically, equilibrium distributions are highly skewed.[6] Contestants do aim for dispersion locally in the sense that their equilibrium distributions are always absolutely continuous, but they do not aim to maximize variance. Thus, in contrast to the corner-solution result found in the literature, the equilibrium level of risk taking is always interior.

The analysis for the multi-contestant/multi-prize case in a symmetric, complete information setting is carried out in Section 3. In this setting, the equilibrium is unique. Every contestant plays a Complementary Beta distribution (Jones (2002)) with one shape parameter being the number of losers and the other being the number of winners (prizes). Contestants' skewness preferences have an interesting relation to the contest structure. When the contest is *selective*, i.e., less than one half of contestants win a prize, equilibrium performance distributions are positively skewed. When the contest is *inclusive*, i.e., more than one half of contestants win a prize, equilibrium distributions are negatively skewed. Equilibrium distributions have zero skew only when exactly one half of contestants win, in which case equilibrium distributions are symmetric and weakly U-shaped. Equilibrium distributions are never strictly unimodal, and, except in the one-winner/one-loser case, never even weakly unimodal. Increasing contest selectivity induces contestants to increase both performance variance and skewness.

Section 4 extends the analysis to contests with incomplete information. Contestants are either strong, with higher capacity, or weak, with lower capacity, and capacity is private information. When strong contestants have significantly larger capacity, the equilibrium is unique. Weak contestants adopt concession strategies that result in prizes only when the realized number of strong contestants is less than the number of prizes. In this case, contestant behavior is not consistent with the risk-taking-and-ruin intuition — that agents facing a high probability of loss prefer high-variance strategies.[7] In fact, weak contestants' performance distributions may exhibit significantly less variance than those of strong contestants. When strong contestants' capacity is only marginally higher, weak contestants adopt strategies that produce a positive probability of winning

[5]In equilibrium, the probability of winning function is a cumulative distribution function. Thus, when we say "unimodal probability of winning function," we mean that the probability of winning function satisfies the conditions for unimodality imposed on a probability distribution.

[6]Skewness preference has been derived in an number of contexts other than contests. See, for example, Tsiang (1972), Brunnermeier, Gollier, and Parker (2007), Barberis and Huang (2008), and Bordalo, Gennaioli, and Shleifer (2012).

[7]For a discussion of risk taking and the probability of ruin, see, for example, Pyle and Turnovsky (1970) and Rose-Ackerman (1991).

even when the realized number of strong contestants exceeds the number of prizes. In this case, there are many equilibria. However, all equilibria produce the same type-conditioned probability of winning.

Section 5 investigates the effects of three commonly used contest modifications—scoring caps, penalty triggers, and localizing contests—on selection efficiency, the efficiency of a contest mechanism in allocating prizes to the strongest contestants.[8] We show that penalizing contestants whose performance fails to reach a threshold weakly improves selection efficiency while dividing a grand contest into smaller local contests weakly harms selection efficiency. Interestingly, capping contestants' performance levels has no impact on selection efficiency (under a weak condition), but it induces contestants to play safer strategies.

Section 6 applies the model to mutual fund tournaments, R&D contests, and stochastic contests. Because, in practice, only a small number of mutual funds are identified as "stars" relative to the population of funds and only stars receive significantly positive rank-based capital inflows, mutual fund tournaments are selective. Thus, our theory predicts that the unsystematic returns of mutual funds have positive skewness, which is consistent with the empirical evidence in Wagner and Winter (2013).

Our R&D contest model follows Klette and de Meza (1986) but relaxes their symmetry assumption. In their model and other R&D competition models (Dasgupta and Stiglitz (1980) and Bhattacharya and Mookherjee (1986)), the winner's prize is not fixed but, rather, strictly increasing in winner performance. Thus, winning small and winning big are no longer payoff equivalent. However, rewards remain rank dependent. Because of rank dependency, winning creates a jump in contestant payoff. For this reason, contestants still have win-small/lose-big preferences. We show that these preferences, through strategic interaction, *concavify* the performance-payoff relation. Thus, we find, in contrast to much of the literature (Bhattacharya and Mookherjee (1986) and Klette and de Meza (1986)), not only that R&D contests do not lead to excessive risk taking but also that, as argued by Dasgupta and Stiglitz (1980), the R&D contest structure actually biases contestants against risk taking.

In the last part of Section 6, we apply our model to "stochastic contests" in which each contestant decides when to stop a privately observed stochastic process and contestants are ranked according to their stopped values (Seel and Strack (2013)). Inter-

[8] The effects of these contest modifications on strategies and outcomes have been analyzed in the context of all-pay contests, in which contestants make effort choices rather than risk choices and the contest designer's goal is effort relevant but not selection relevant. See, for example, Che and Gale (1998) and Gavious, Moldovanu, and Sela (2002) for scoring caps, Gilpatric (2009), Akerlof and Holden (2012), and Moldovanu, Sela, and Shi (2012) for penalty triggers, and Moldovanu and Sela (2006) and Fu and Lu (2009) for localizing contests.

estingly, when the stochastic process is a geometric Brownian motion or a Brownian motion absorbed at zero, stochastic contests and our capacity-constrained contests are strategically equivalent, with a contestant's capacity in stochastic contests represented by the initial value of his stochastic process. This equivalence is founded on Skorokhod embedding theorems (Skorokhod (1965)), which determine the conditions under which a probability distribution can be induced by stopping a stochastic process.[9] We show that an increase in contest selectivity or contest size induces contestants to postpone the stopping of the stochastic process.

Section 7 concludes by summarizing the main findings of the paper and providing a brief discussion of the possible extensions of the model. All the proofs are relegated to the Appendix. The online Supplementary material contains the proof of the equilibrium uniqueness in the certain-capacity contest game studied in Section 3.

## 1.1 Related literature

Our paper makes several contributions to the contest literature. First, our paper contributes to the contest literature on risk-taking strategies. As mentioned earlier, this literature restricts contestants' distributional choices either to symmetric distributions (Klette and de Meza (1986), Hvide (2002), Gaba, Tsetlin, and Winkler (2004), Goel and Thakor (2008), Kräkel (2008), and Gilpatric (2009)) or to mixtures of two exogenously specified distributions (Hvide and Kristiansen (2003), Taylor (2003), Kräkel and Sliwka (2004) and Nieken and Sliwka (2010)). Our results show that these restrictions have profound effects. For example, Gaba, Tsetlin, and Winkler (2004) show that, when contestants can choose any symmetric distribution about the same mean, contestants play safe when the contest is inclusive while take extreme risk when the contest is selective. In our framework, which imposes no restrictions on the shape of the distributions, neither the safest nor the riskiest strategy is played in equilibrium, and symmetric unimodal distributions are the antithesis of the equilibrium distributional choices.

Second, our paper contributes to the literature on contests where contestants have win-small/lose-big incentives. One class of contests that induce such incentives is the all-pay contests in which contestants compete for fixed prizes by submitting costly and non-refundable bids (Varian (1980), Ellingsen (1991), Baye, Kovenock, and de Vries (1993), Che and Gale (1998), Gavious, Moldovanu, and Sela (2002), and Siegel (2009)). Because of bidding costs, conditional on winning or losing, a contestant's payoff decreases with his bid, and the win-small/lose-big incentive is hence induced by cost-saving concerns. In our model, the win-small/lose-big incentive is generated by a differ-

[9] See Obłój (2004) for a survey of Skorokhod embedding theorems.

ent but analogous mechanism. Performance levels determine the probability of winning in our model and thus are analogous to bids in an all-pay contest. In contrast to bids in an all-pay contest, these "performance-level bids" have no direct cost to contestants and only directly affect contestants through their effect on the probability of winning. However, because such bids use up capacity and capacity is constrained, the capacity constraint indirectly imposes a shadow price. This shadow price ensures that, as in an all-pay contest, conditional on winning the contest, it is optimal to win by the smallest possible margin and, conditional on losing, to lose by the largest possible margin.

Although the win-small/lose-big incentive leads to randomized bids in both all-pay contests and our model, the different driving forces behind this incentive affect contestants' randomization strategies differently in at least two respects. First, in all-pay contests, the supports of the bid distributions are always bounded above by contestants' prize valuations. Thus, changes in contest structure have little impact on the supports of the bid distributions. In contrast, in our model, although the "win-small" incentive induces contestants to bid on a bounded support, since maximal rational bids are not capped by prize valuations, the upper bounds often vary when contest structure changes. This difference significantly affects comparative statics. For example, Hillman and Samet (1987) find that, in an all-pay contest, increasing the number of contestants induces each contestant to behave less aggressively in terms of first-order stochastic dominance.[10] In contrast, in our capacity-constrained model, we show that increasing the number of contestants induces each contestant to behave more aggressively, i.e., choose riskier distributions in terms of second-order stochastic dominance. Second, even in symmetric, complete information all-pay contest models, some contestants may not be active bidders. This can lead to a continuum of asymmetric equilibria in which some contestants are inactive (Baye, Kovenock, and de Vries (1996)). In contrast, in our model, contestants are always active. Moreover, in our model, when information is complete and contestants are symmetric, the equilibrium is unique and symmetric.

Third, our paper contributes to the literature on selection efficiency of contests. Some works take a statistical approach with contestants' performance distributions exogenously specified (Meyer (1991), Ryvkin and Ortmann (2008) and Ryvkin (2010)) while others take into account contestants' effort-bidding strategies (Clark and Riis (2001) and Kawamura and Moreno de Barreda (2014)). The effect of risk taking on selection properties of contests is analyzed in Hvide and Kristiansen (2003) under the assumption that contestants can only choose between two fixed strategies. They find

[10] This result is not explicitly stated in Hillman and Samet (1987). However, simple calculation shows that it is implied by their Proposition 1. Similarly, the results in Clark and Riis (1998) imply the same conclusion in a multi-prize all-pay setting.

that increasing the number of contestants or contestant quality can sometimes lower winner quality. In contrast, in our more flexible setting, winner quality is always weakly increasing in both the number of contestants and contestant quality.

Finally, our paper contributes to the literature on stochastic contests and Colonel Blotto games. Seel and Strack (2013) develop the stochastic contest framework, in which each contestant privately observes a stochastic process and chooses a stopping time and is ranked by the resulting stopped value of the process. Capacity is represented by the initial value of his stochastic process. Seel and Strack consider the case where several contestants with the same capacity compete for one prize and the stochastic process is a Brownian motion absorbed at zero.[11] We extend their analysis to the multi-prize case with both complete and incomplete information on capacity and we also allow the process to be a geometric Brownian motion. Moreover, we also examine the effect of a change in contest structure on stopping strategies. As we will discuss in detail in Section 7, because our analysis of the capacity-certainty version of the model provides a closed-form solution for an arbitrary number of prizes and contestants, for a relaxed version of the Colonel Blotto game, our paper also contributes to solving the technically challenging problems related to characterizing the solutions to this game.

# 2 Besting a fixed distribution

## 2.1 Framework

Consider the problem of a contestant picking a distribution function $F$ for a nonnegative random variable, $X$, so as to maximize $X$'s probability of exceeding the realized value of another nonnegative random variable, $Y$, whose distribution $P$ is exogenously determined and statistically independent of $X$. We call the distribution that the contestant is attempting to surpass the "fixed distribution" and the distribution selected by the contestant the "challenge distribution." To abstract from the problem of ties, in this section, we assume that the fixed distribution has no point mass and is thus continuous. No continuity restriction is imposed on the challenge variable. Let $F(\cdot)$ and $P(\cdot)$ denote, respectively, the cumulative distribution functions (CDFs) of the challenge and the fixed random variables, and let $dF$ and $dP$ denote the measures associated with the random variables. We can express the probability that the challenge variable wins as

$$\mathbb{P}\{X \geq Y\} = \int_0^\infty \mathbb{P}\{Y \leq x\}\, dF(x) = \int_0^\infty P(x) dF(x). \tag{1}$$

[11] Seel and Strack (2013) also consider the case in which contestants have different capacities, but this is done only for the two-contestant case.

The contestant's problem is to maximize this probability. It is more convenient to express this problem as one of choosing probability measures over the nonnegative real line rather than random variables on a measure space. Thus, we can formulate the contestant's problem as one of choosing a challenge measure *dF* to use against the fixed measure *dP*. The challenge measure has to satisfy two constraints: (a) it has to be a probability measure and (b) its expectation is constrained to be weakly below some value, say $\mu > 0$. We call the latter *the capacity constraint*. We assume that $P(\mu) < 1$. Otherwise, the problem is trivial, since $P(\mu) = 1$ implies that simply choosing performance $\mu$ will ensure winning with certainty. Since reducing the total mass of the challenge measure below 1 will never strictly increase the contestant's payoff, the solution to this problem coincides with the solution to the following relaxed problem:

$$\max_{dF\geq 0} \int_0^\infty P(x)dF(x) \quad s.t. \int_0^\infty dF(x) \leq 1 \quad \& \int_0^\infty x dF(x) \leq \mu. \tag{2}$$

Since $P$ is a CDF, $P$ is nondecreasing, bounded, and upper semicontinuous. Thus, given that the feasible set of measures is compact in the weak topology, an optimal solution to problem (2) exists. Since problem (2) is a linear program, feasible and bounded, by duality theory, strong duality holds.[12] Therefore, the optimal solution to (2) must also be a maximizer of the following Lagrangian:

$$\mathscr{L}(dF,\alpha,\beta) = \int_0^\infty P(x)dF(x) - \alpha\left(\int_0^\infty dF(x) - 1\right) - \beta\left(\int_0^\infty x dF(x) - \mu\right), \tag{3}$$

where $\alpha$ and $\beta$ are nonnegative optimal dual variables that solve the following dual problem:

$$\min_{\alpha,\beta\geq 0} \sup_{dF\geq 0} \mathscr{L}(dF,\alpha,\beta). \tag{4}$$

Rewrite equation (3) as

$$\mathscr{L}(dF,\alpha,\beta) = \int_0^\infty [P(x) - (\alpha + \beta x)]\, dF(x) + \alpha + \beta\mu. \tag{5}$$

Equation (5) implies that the optimal dual variables that solve (4) must satisfy

$$P(x) - (\alpha + \beta x) \leq 0 \quad \forall x \geq 0, \tag{6}$$

since otherwise $\sup_{dF\geq 0}\mathscr{L}(dF,\alpha,\beta)$ goes to positive infinity. Thus, $\alpha + \beta x$ is an upper bound for $P(x)$.

[12]For a discussion of duality theory, see Boyd and Vandenberghe (2004).

When condition (6) is satisfied, the value of the dual problem (4) equals

$$\alpha + \beta\mu, \tag{7}$$

which is strictly increasing in both $\alpha$ and $\beta$. Thus, the nonnegative optimal dual variables must minimize $\alpha + \beta\mu$ subject to condition (6). Hence, given that $P$ is upper semicontinuous, condition (6) must be binding for some $x \geq 0$, i.e, there exists some point(s) $x' \geq 0$ such that $P(x') - (\alpha + \beta x') = 0$.

Thus, $\alpha + \beta x$ is an *upper support line* for $P$. Placing any probability weight on points at which $P(x) - (\alpha + \beta x) < 0$ lowers the Lagrangian. Thus, the optimal challenge distribution will place no weight on such points. Therefore, the optimal challenge measure is always concentrated on points at which the upper support line, $\alpha + \beta x$, meets the distribution function, $P$. Thus, the optimal challenge measure, $dF$, and the associated optimal dual variables, $\alpha$ and $\beta$, must satisfy the following conditions:

$$\begin{aligned} P(x) &\leq \alpha + \beta x \quad \forall x \geq 0; \\ dF\{x \geq 0 : P(x) &< \alpha + \beta x\} = 0. \end{aligned} \tag{8}$$

By strong duality, the primal problem (2) must have its optimal value equal to that of the dual problem (4), so the contestant's optimal winning probability is given by equation (7) with $\alpha$ and $\beta$ being optimal dual variables.

**Remark 1.** *Since we have assumed that $P(\mu) < 1$, given that P is a CDF and, hence, there exists $x^o > \mu$ such that $P(x^o) > P(\mu)$, it is clear that the capacity constraint must be binding at the solution of problem* (2)*. In other words, the optimal challenge distribution must have its mean equal to $\mu$. By condition* (8)*, this requires the upper support line, $\alpha + \beta x$, to meet P on some point weakly below $\mu$ and also on some point weakly above $\mu$.*

## 2.2 Besting specific distributions

The optimal challenge measure depends on the shape of the fixed distribution, $P$. If $P$ is strictly concave over its support, then at each point on the graph of $P$ over its support, there is an upper support line that is strictly above $P$ at all other points. Thus, by Remark 1, the optimal challenge distribution places all weight on the point $\mu$. If $P$ is strictly convex on its support, its support must be bounded for $P$ to be a CDF. Assume that the support of $P$ is $[0, \zeta]$. Recall that, to make the discussion nontrivial, we have assumed $P(\mu) < 1$, which is equivalent to assuming $\zeta > \mu$ here. By Remark 1, it is clear that the only upper support line associated with the optimal challenge measure is the one that connects the value of $P$ at $x = 0$ to its value at its upper endpoint, $x = \zeta$. Thus, the optimal challenge measure only places weight on 0 and $\zeta$, the lower and upper endpoints of the support of $P$.

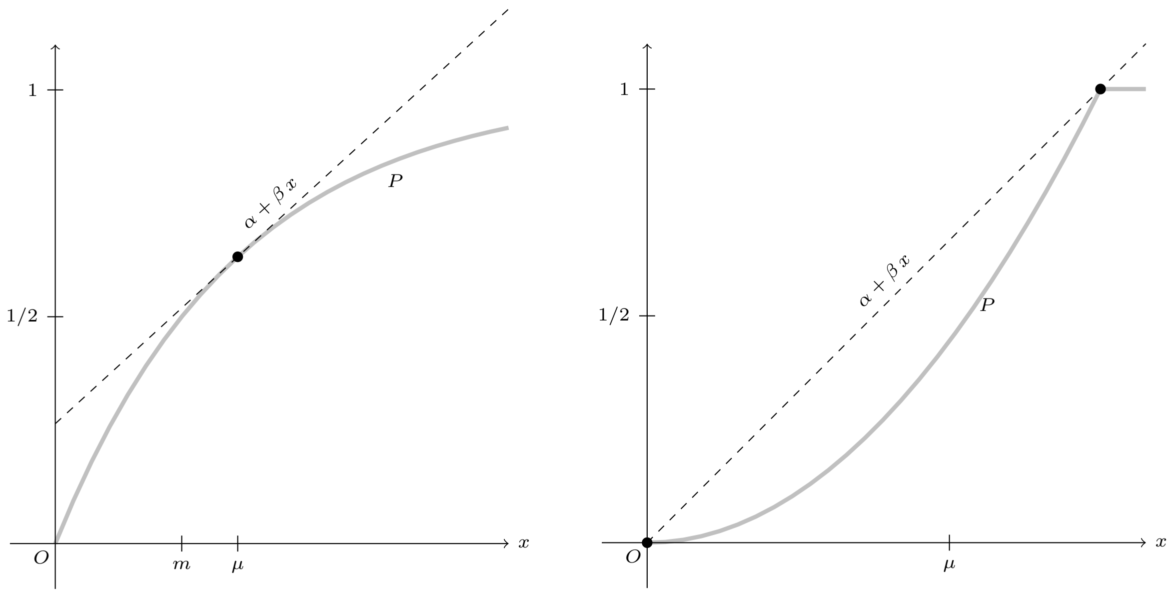
Suppose that the contestant's capacity, $\mu$, equals the mean of $P$. In this case, we can interpret our problem as one picking a challenge distribution, $F$, with the best chance of besting a fixed distribution, $P$, with the same mean. One possible solution is to set $F = P$ which would yield a probability of winning equal to 1/2. Is it possible to do better, i.e., can a contestant "best" $P$ by garnering a probability of winning exceeding 1/2?

First, consider the case where $P$ is strictly concave. Consider the optimal challenge distribution against a given concave fixed distribution detailed above. This challenge distribution places all the probability weight on $\mu$ and, hence, yields a probability of winning equal to $P(\mu)$. Since $\mu$ is also the mean of the random variable, $Y$, whose distribution, $P$, is strictly concave, by Jensen's inequality, $P(\mu) > \mathbb{E}[P(Y)] = 1/2$. Thus, "playing safe" always bests a strictly concave CDF with the same mean. This result is illustrated in Panel A of Figure 1.

Now consider the case where $P$ is strictly convex with support $[0, \zeta]$. As shown earlier, the optimal challenge distribution against $P$ involves placing all the probability mass on the points 0 and $\zeta$. For the challenge distribution to have expectation $\mu$, the weight on $\zeta$, which also equals the probability of the challenge distribution winning, must equal $\mu/\zeta$. Since $P$ is strictly convex and supported by $[0, \zeta]$, $P$ first-order stochastically dominates the uniform distribution over $[0, \zeta]$. Hence, the mean of $P$ exceeds the mean of the uniform distribution, i.e., $\mu > \zeta/2$. Thus, the probability of winning using the optimal challenge distribution, $\mu/\zeta$, exceeds 1/2. Therefore, a strictly convex CDF can always be bested by "taking a gamble," i.e., choosing a distribution with the same mean that places all the weight on the two endpoints of the support of the convex CDF. This result is illustrated in Panel B of Figure 1.

However, a strictly convex CDF implies an increasing probability density function (PDF). This is not a common property for "textbook" distributions. Symmetric unimodal distribution functions are far more commonly encountered in the economics and statistics literature. Is it possible to best these distributions? In fact, the optimal challenge distribution against such distributions is a "win-small/lose-big" distribution. Suppose that $P$ is a symmetric strictly unimodal distribution, with the lower bound of its support being 0. Define $x^*$ as the maximizer of $P(x)/x$ over the support of $P$. Note that $x^*$ exists, is unique, and exceeds the mode of the distribution.[13] Since the distribution is

[13]Since $P$ is convex below the mode, $\lim_{x \downarrow 0} P(x)/x$ exists and is finite. Extending the definition of $P(x)/x$ when $x = 0$ by using this limit, we see that $P(x)/x$ is a continuous function defined on a compact




Panel A. Concave case

Panel B. Convex case

Figure 1: Besting concave and convex CDFs. The graphs illustrate optimal challenge distributions played against fixed distributions whose mean equals the mean of the challenge distribution. The upper support lines are denoted by dashed lines. The fixed distributions are denoted by thick grey lines. The expectations and medians of the fixed distributions are denoted by $\mu$ and $m$, respectively.

symmetric, $x^*$ exceeds the mean of the distribution as well. Consider the line connecting the origin to the point $(x^*, P(x^*))$. Note that this line is the only upper support line for $P$ that satisfies the condition in Remark 1. Consider a distribution that places weight of $\mu/x^*$ on $x^*$ and weight of $(1-\mu/x^*)$ on 0. Then the expectation of this distribution is $\mu$ and the probability of winning is

$$\frac{\mu}{x^*} P(x^*). \tag{9}$$

Note that, by the definition of $x^*$, over the support of $P$,

$$\frac{P(x^*)}{x^*} > \frac{P(x)}{x} \quad \forall\, x \neq x^*,\, x \neq 0.$$

Thus, over the support of $P$,

$$P(x^*)\, x > x^*\, P(x) \quad \forall\, x \neq x^*,\, x \neq 0.$$

support and, hence, a maximizer must exist. Unimodality implies that $P(x)$ is convex below the mode and, hence, $P(x)/x$ is increasing below the mode. Thus, any maximizer of $P(x)/x$ must be greater than the mode of $P$. Unimodality also implies that $P(x)$ is strictly concave (on its support) above its mode. Thus, $P(x)/x$ is strictly quasi-concave. Strict quasi-concavity implies that the maximizer of $P(x)/x$ is unique.

Integrating both sides over $P$ yields

$$P(x^*)\,\mu = P(x^*)\int x\,dP(x) > x^*\int P(x)dP(x) = x^*\,\frac{1}{2}. \tag{10}$$

Combining (9) and (10) shows that a symmetric strictly unimodal fixed distribution can always be bested by a win-small/lose-big challenge distribution. The construction of the optimal challenge distribution is illustrated by Figure 2. In fact, the win-small/lose-big distribution bests asymmetric unimodal distributions whenever the fixed distribution's mean, $\mu$, is less than $x^*$. A sufficient condition for $\mu < x^*$ is for the mean-median-mode inequality to hold. If the opposite inequality holds, the mode-median-mean inequality, then $1/2 = P(\text{median}) < P(\mu)$ and thus a play-safe distribution bests the unimodal distribution. The "typical" case for unimodal distributions is for either the mean-median-mode or the mode-median-mean inequality to hold, and there are a number of results in the statistical literature identifying sufficient conditions for one of these inequalities to hold.[14] Thus, typically, a unimodal distribution can be bested by a simple distribution, either a win-small/lose-big or a play-safe distribution.

These specific results can be extended to find challenge distributions that best many other classes of fixed distributions; however, it is impossible to enumerate all of them. But we can answer one more general question. Is there a distribution that cannot be bested? The best reply to such a distribution would have to yield a probability of winning equal to $1/2$. One such distribution is the uniform distribution over $[0, 2\mu]$. To see this, note that the probability of winning against a uniform distribution with mean $\mu$ when the challenge performance equals $x$ is $P(x) = \min[x/(2\mu), 1]$. Clearly, an optimal challenge distribution must not place any weight on the points outside the support of $P$, and, for any challenge distribution, $F$, over the support of $P$ with mean $\mu$,

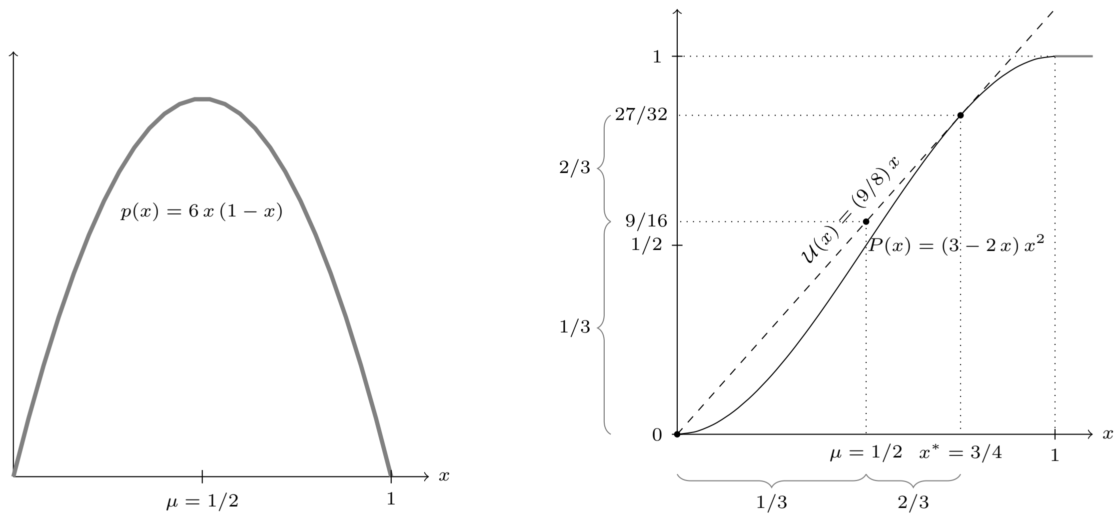
$$\int P(x)dF(x) = \int \frac{x}{2\mu}dF(x) = \frac{1}{2\mu}\int x\,dF(x) = \frac{1}{2}.$$

The following proposition shows that the uniform distribution is in fact the only distribution that cannot be bested.

**Proposition 1.** *If a distribution $P$ has the property that the maximum probability of winning against $P$ for all challenge distributions with the same mean equals 1/2, then $P$ is a uniform distribution.*

As we have shown how to best concave and convex CDFs, Proposition 1 is not difficult to understand. If a distribution $P$ exhibits local convexity or concavity somewhere on its support, it will be bested by a challenge distribution that bests $P$ locally

[14] See Sato (1997) for further discussion of the mode-median-mean inequality.




Panel A. PDF

Panel B. CDF

Figure 2: Panel A illustrates the PDF, $p$, for a symmetric unimodal distribution — the Beta distribution with shape parameters $(2,2)$. Panel B illustrates a distribution which bests the Beta distribution: The CDF, $P$, of the Beta distribution is denoted by the grey line and the upper support line, $\mathscr{U}$, for the CDF by the dashed line. The distribution which bests this Beta distribution places all its weight on two points, $x = 0$ and $x = x^* = 3/4$, where the support line meets the distribution function. The probabilities of these points are set so that the expected value of the challenge distribution equals the expected value of the Beta distribution. Thus, the probability of $x = 0$ equals 1/3 and the probability of $x = x^* = 3/4$ equals 2/3. The probability that the challenge distribution will win is given by the probability that $x = x^* = 3/4$ times the probability of winning when $x = x^*$, $P(x^*) = 27/32$, which is $(2/3)(27/32) = 9/16 > 1/2$.

and mimics $P$ everywhere else. Thus, to avoid being bested, $P$ has to be linear over its support. Proposition 1 implies that, in a contest game where two equally matched contestants compete for one prize, there exists a unique equilibrium in which both contestants choose a uniform distribution, since playing a uniform distribution is the only way of preventing a contestant from being bested by his strategic competitor.[15]

# 3 Contests with no capacity uncertainty

In Section 2, we started with a decision problem—how to best a fixed distribution—and we ended with an equilibrium solution for the simple contest with two strategic contestants and one prize. In this section, we examine a more general contest in which $n$ strategic contestants compete for $m$ identical prizes, where $1 \leq m \leq n-1$. A contestant can win at most one prize. The model consists of two stages: in stage one,

[15] Strictly speaking, this implication, for the time being, is based on the assumption that $P$ is continuous. However, we show in the next section that $P$ must be continuous if it is chosen by a strategic competitor.

each contestant simultaneously and effortlessly chooses a probability distribution for his random performance; in stage two, each contestant's realized performance is independently drawn from the distribution he chose and the $m$ prizes are given to the $m$ contestants with the highest realized performances. Ties are broken randomly.

The two conditions imposed on admissible distributional choices of a strategic contestant are the same as those in Section 2: (i) the support of the distribution must be on the nonnegative real line, and (ii) the expected performance cannot exceed the contestant's capacity. In this section, we assume that all the contestants have the same capacity equal to $\mu > 0$. We relax this assumption in the next section.

We focus on symmetric Nash equilibria throughout the paper, in which contestants with the same capacity play the same distribution. In this section, this means that all the contestants play the same distribution as they have the same capacity. While we restrict attention to symmetric equilibria, we show in the Supplementary material that there are no asymmetric equilibria when all the contestants have the same capacity. Thus, the symmetric equilibrium derived later on in this section is in fact the unique equilibrium.

## 3.1 Probability of winning

Since the contestants' choice sets are convex, we do not need to consider mixed strategies. Contestant $i$'s performance is a random variable, denoted by $X_i$, and his realized performance is $x_i$, drawn from the CDF, $F_i(\cdot)$, he chose. Since we concentrate on symmetric equilibria and all the contestants are homogeneous, for notational convenience, we suppress the index of a contestant's identity. We denote a contestant's probability of winning when his realized performance equals $x$ by $P(x)$, where $P$ is produced by his competitors' strategies.

Note that, in a symmetric equilibrium, no contestant's performance distribution places positive mass on a single performance level. This follows from symmetry: if one contestant placed positive weight on a given point, then all contestants would place positive weight on this point. If all contestants placed weight on a given point, say $x^o$, then, with positive probability, all contestants would have performance equal to $x^o$ and, hence, would tie at $x^o$. In this case, a performance level greater than $x^o$ by an infinitesimal amount would break the tie and, hence, would generate, with positive probability, a non-infinitesimal upward jump in a contestant's payoff. This contradicts the optimality of performance distributions placing mass on $x^o$. This "no-mass" result implies that the probability of winning function, $P$, is continuous and intersects the origin.

**Lemma 1.** *In a symmetric equilibrium, the probability of winning function, $P$, is continuous and $P(0) = 0$.*

Lemma 1 implies that $P$ satisfies the continuity property of the "fixed distribution" in Section 2. In essence, $P$ is the CDF that a contestant plays against. Since no one can do better than others in a symmetric equilibrium, a contestant's best reply must not best $P$. As suggested by Proposition 1, $P$ then has to be a uniform distribution. The next proposition confirms this result and completely characterizes $P$.

**Proposition 2.** *In a symmetric equilibrium, the probability of winning function, P, is given by*

$$P(x) = \begin{cases} \frac{m}{n\mu} x & \text{if } 0 \leq x \leq \frac{n\mu}{m} \\ 1 & \text{if } x > \frac{n\mu}{m} \end{cases}. \tag{11}$$

## 3.2 Individual contestant strategies

Based on Proposition 2, to solve for the equilibrium strategy, $F$, we only need to determine the relation between $P$ and $F$.

Consider a contestant with a realized performance level, $x$. He has $n-1$ competitors, each choosing an independent and identical distribution, $F$. Since no one places point mass, there is no chance of a tie. His probability of winning given performance $x$, $P(x)$, equals the probability that his realized performance $x$ exceeds the $m$th highest performance of the remaining $n-1$ contestants. Since the $m$th highest performance out of $n-1$ is the $(n-m)$th lowest performance,

$$P(x) = \mathbb{P}\Big\{X_{n-m:n-1} \leq x\Big\},$$

where $X_{n-m:n-1}$ represents the $(n-m)$th order statistic for distribution $F$. Thus, $P$ equals the distribution of this order statistic, given by (see Lemma 1.3.1 in Reiss (1980))

$$P(x) = F_{n-m:n-1}(x) = \sum_{i=n-m}^{n-1} \binom{n-1}{i} F(x)^i (1-F(x))^{(n-1)-i}. \tag{12}$$

Inserting (12) into (11) gives the equilibrium performance distribution.

**Proposition 3.** *In a symmetric equilibrium, the performance distribution, F, satisfies the following equation on its support* $[0, n\mu/m]$*:*

$$\sum_{i=n-m}^{n-1} \binom{n-1}{i} F(x)^i (1-F(x))^{(n-1)-i} = \frac{m}{n\mu} x.$$

By Proposition 3, when $n=2$ and $m=1$, $F$ is a uniform distribution, confirming our finding in Proposition 1. When $n>2$ and $m=1$, $F$ is a power-function distribution.

In general, $F$ is a stretched *Complementary Beta distribution*, which coincides with some other types of distributions in some special cases (Jones (2002)). It is obtained by swapping the roles of the CDF and the quantile function of the Beta distribution. We provide its definition below.[16]

**Definition 1.** *If $U$ is a random variable on the support $[0, 1]$ whose CDF, $F_U$, satisfies*

$$\sum_{j=a}^{a+b-1} \binom{a+b-1}{j} F_U(u)^j (1-F_U(u))^{a+b-1-j} = u,$$

*then $U$ has a Complementary Beta distribution with shape parameters being a and b, denoted by $U \sim CB(a,b)$.*

The next result follows from Proposition 3 and Definition 1.

**Proposition 4.** *In a symmetric equilibrium, contestant random performance, $X$, satisfies $\frac{m}{n\mu}X \sim CB(n-m,m)$.*

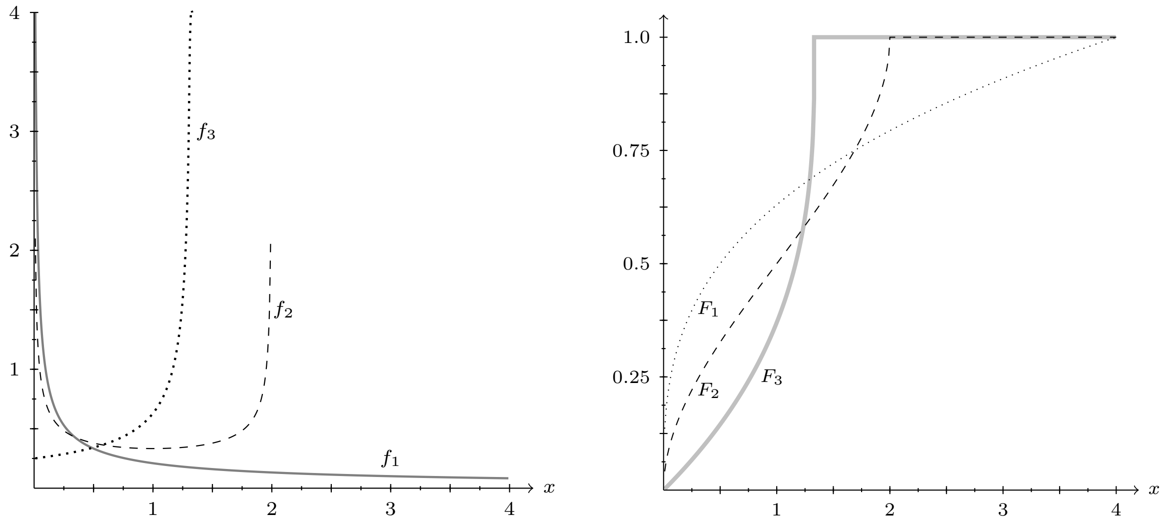
The two shape parameters of the equilibrium distribution, $n-m$ and $m$, coincide with the number of losers and the number of winners, respectively. Jones (2002) shows that the PDF of the Complementary Beta distribution is U-shaped when the two shape parameters are strictly greater than 1, i.e., when $1 < m < n-1$ in our case. The equilibrium distribution is never strictly unimodal and, except in the one-winner/one-loser case, even never weakly unimodal. The literature on endogenous risk taking in contests usually assumes unimodality and symmetry of admissible distributions. Under this assumption, the literature finds that, if risk taking is costless, the equilibrium level of risk taking is always extremal (Klette and de Meza (1986), Hvide (2002), Gaba, Tsetlin, and Winkler (2004), Goel and Thakor (2008), and Kräkel (2008)). In contrast, our result suggests that the exogenously specified symmetric unimodal distributions are the antithesis of the equilibrium distributional choices, and the equilibrium level of risk taking is never extremal: contestants choose non-atomic performance measures and, in this sense, their strategies are "locally dispersed," but, the global dispersion of performance levels is limited.

## 3.3 Contest selectivity

Contest selectivity increases when the number of contestants, $n$, increases or the number of prizes, $m$, decreases. Making the contest more selective increases the right skewness

[16] Jones (2002) defines the Complementary Beta distribution in a more general way where the two shape parameters characterizing the distribution can be any positive numbers. As we only need to look at the case where the two shape parameters are positive integers, we use a confined definition.

of the equilibrium distribution. The intuition for this result is easiest to understand if we restrict attention to the contest with a fixed number of contestants and we increase selectivity by decreasing the number of prizes. As the number of prizes falls, for any fixed distribution selected by the contestants, the probability of a given contestant winning over the high performance level range relative to the low performance level range increases.[17] As shown in Proposition 2, in equilibrium, the marginal incentives must be the same at all performance levels in the support of the equilibrium distribution. Thus, the equilibrium distribution function's slope at the high end must decrease relative to the low end to compensate, i.e., skewness must increase. This result is illustrated in Figure 3 and demonstrated in Proposition 5.




Panel A. PDF

Panel B. CDF

Figure 3: *Effects of increasing contest selectivity on the equilibrium distribution.* The figure plots equilibrium distributions for contests with four contestants. Contestant capacity is normalized to 1. The equilibrium PDF (CDF) when $m$ winners are selected is denoted by $f_m$ ($F_m$). Increasing selectivity increases the right skewness (or decreases the left skewness) and the dispersion of the equilibrium distribution.

Using the properties of the Complementary Beta distribution, in Proposition 5, we characterize the effects of selectivity on performance dispersion and skewness, as measured, respectively, by the L-scale and the L-skewness of the equilibrium distribution.[18] We use L-moments instead of conventional moments for computational convenience.

[17]This effect of selectivity on the relative likelihood of winning over the high versus low range follows since the distribution of marginal winning bid in the more selective contest first-order stochastically dominates that in the less selective contest (Nanda and Shaked (2001)).

[18]The L-scale is the second L-moment of a distribution. The second L-moment equals half the expected difference between the highest and lowest of two random draws from the distribution. It is a measure of dispersion analogous to standard deviation and satisfies the conditions specified in Oja (1981) for dispersion measures. The L-skewness is the third L-moment ratio, calculated by dividing the third L-moment by the L-scale. The third L-moment equals one third of the expected difference of the differences

**Proposition 5.** *i. The L-scale of the equilibrium distribution is*

$$\lambda_{2,F} = \frac{(n-m)\mu}{n+1}. \tag{13}$$

*ii. The L-skewness of the equilibrium distribution is*

$$\tau_{3,F} = \frac{n-2m}{n+2}. \tag{14}$$

*iii. The L-scale and the L-skewness of the equilibrium distribution are both strictly increasing in n and strictly decreasing in m.*

*iv. The equilibrium PDF is symmetric about its mean if $m/n = 1/2$, right skewed if $m/n < 1/2$, and left skewed if $m/n > 1/2$.*

Proposition 5 shows that the equilibrium distribution is right skewed when the contest is selective, left skewed when the contest is inclusive, and symmetric when exactly one half of the contestants receive a prize. This result can be contrasted with Gaba, Tsetlin, and Winkler (2004) who find that, when contestants are restricted to symmetric distributions, each contestant maximizes performance variance when the contest is selective, minimizes performance variance when the contest is inclusive, and is indifferent between all levels of variance when one half of them receive a prize. Their result is driven by the symmetric distribution assumption that forces contestants with skewness preference to exhibit variance preference.

Proposition 5 also implies that increasing contest selectivity increases the dispersion of the equilibrium distribution, which suggests that contestants play riskier strategies when contest selectivity increases. In the next proposition, we confirm this point by showing that an increase in selectivity induces a *simple mean-preserving spread (s-MPS)* of the equilibrium performance distribution in the sense of Diamond and Stiglitz (1974): for two CDFs, $F$ and $G$, $F$ is a s-MPS of $G$ if $F$ and $G$ have the same mean and there exists $x'$ such that $F(x) - G(x) \leq (\geq) 0$ when $x \geq (\leq) x'$. The s-MPS relation between more and less selective contests is illustrated by Panel B in Figure 3.

**Proposition 6.** *Increasing contest selectivity, either by increasing the number of contestants or by decreasing the number of prizes, induces a simple mean-preserving spread of contestant performance.*

between the highest draw and the middle draw and the middle draw and the lowest draw of three random draws from the distribution. The L-skewness is a measure of the asymmetry of a distribution analogous to conventional skewness measures (see Hosking (1990)) and satisfies the conditions specified in Oja (1981) for skewness measures.

In contrast to the result in Gaba, Tsetlin, and Winkler (2004) that increasing selectivity either has no or an extreme effect on the riskiness of contestant performance when contestants are restricted to symmetric distributions, Proposition 6 implies that the change in riskiness is incremental in the change in selectivity when the symmetric distribution restriction is removed.

## 3.4 Contest size

Contest size increases when $n$ increases in proportion to $m$. In contrast to increasing selectivity, increasing contest size does not affect the support of the equilibrium distribution. However, increasing contest size affects the shape of the distribution. Using simple calculations based on the properties of the Complementary Beta distribution, we show, in Proposition 7, that increasing contest size increases both the absolute value of skewness and the dispersion of the equilibrium distribution.

**Proposition 7.** *When n and m increase by the same proportion,*

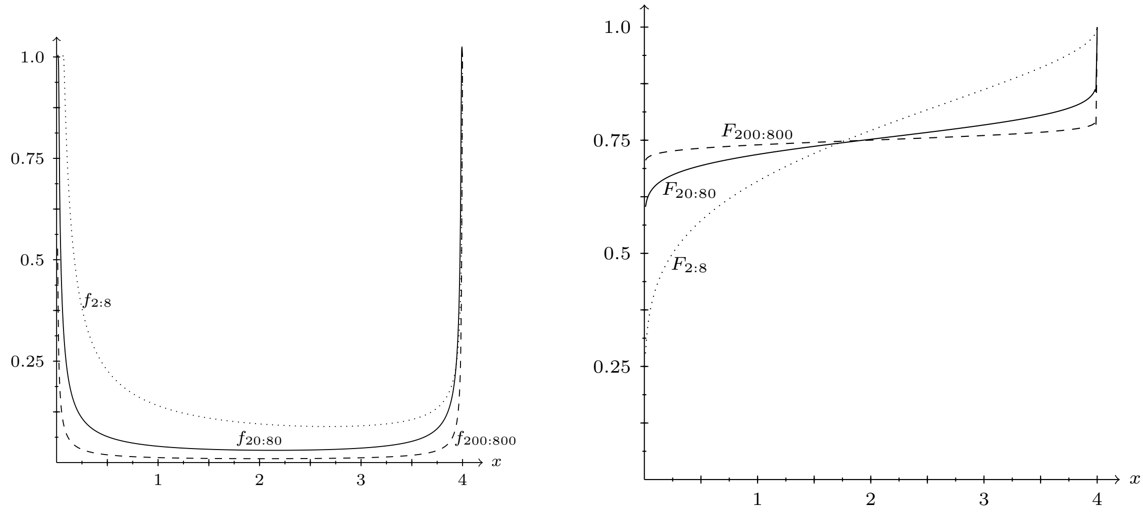
*i. the support of the equilibrium distribution and the direction of skewness remain constant;*

*ii. both the L-scale and the absolute value of the L-skewness increase;*

*iii. the performance distribution undergoes a simple mean-preserving spread.*

The effects identified in Proposition 7 are depicted in Figure 4. In all the three cases depicted in Figure 4, one fourth of contestants in the contest win a prize. The PDFs and CDFs plotted vary by the number of contestants, which varies in multiples of ten between eight and eight hundred.

Figure 4 illustrates that (a) increasing contest size always increases the dispersion of the equilibrium performance distribution but (b) the increase in dispersion is quite modest: increasing dispersion does not change the support of the performance distribution and even a 10-fold scale increase in contest size has only a modest effect on the graph. Characterizations (a) and (b) can both be understood by considering the effect of increasing contest size holding the performance distribution constant and then considering how the performance distribution must adjust to continue to satisfy the equilibrium conditions. Given a fixed performance distribution, as contest size increases, the probability of winning function converges asymptotically at rate $1/\sqrt{n}$ to a Normal distribution (See Theorem 4.2.3 in Reiss (1980)). The Normal distribution is unimodal. From the results in Section 2, we know that the best response to a symmetric unimodal distribution is a win-small/lose-big distribution. Such distributions have two-point supports. Since two-point support distributions are discontinuous, they cannot be played in




Panel A. PDF Panel B. CDF

Figure 4: *Effects of increasing contest size on the equilibrium distribution.* The figure plots the equilibrium distributions for three contest sizes; $f_{m:n}$ ($F_{m:n}$) represents the equilibrium PDF (CDF) when $m$ out of $n$ contestants win. In each of the cases graphed, winner proportion is fixed at 1/4. Contestant capacity is normalized to 1.

equilibrium. Thus, the performance distribution must adjust to counter the emergence of a unimodal probability of winning function. Unimodality is countered by shifting weight in the performance distribution toward the endpoints. The required shift is fairly modest since the rate of asymptotic convergence is fairly slow, $O(1/\sqrt{n})$. This effect is absent in Gaba, Tsetlin, and Winkler (2004) who find that an increase in contest size has no effect on risk-taking strategies. Note that Figure 4 seems to indicate convergence to a limiting distribution as contest size increases without bound. This conjecture is indeed correct, as the next proposition shows.

**Proposition 8.** *Fixing the proportion of winners at $\rho$, i.e., $m/n = \rho$, when $n \to \infty$, the equilibrium performance distribution converges weakly to the limiting distribution, $F_\infty$, defined by*

$$F_\infty(x) = \begin{cases} 1-\rho & \text{if} \quad 0 \le x < \mu/\rho \\ 1 & \text{if} \quad x \ge \mu/\rho \end{cases}. \tag{15}$$

The limiting distribution, $F_\infty$, is Bernoulli, placing all its weight on the extreme points of the common support for the sequence of equilibrium distribution functions. The logic behind Proposition 8 is fairly straightforward: holding the proportion of winners constant while increasing the number of contestants makes the performance level required to win a prize more predictable. To counter this effect, the equilibrium distribution must become more unpredictable. Since contest size has no effect on the equilibrium range of performance levels, reduced predictability can only be produced by moving probability weight toward the extreme points of the support. In the limit, all weight is placed on these extreme points.

In contrast to the result in Gaba, Tsetlin, and Winkler (2004) that when contestants are restricted to symmetric distributions, they play Bernoulli distributions when the contest is selective, our result shows that, when contestants can choose asymmetric distributions, they never play Bernoulli distributions but that the Bernoulli distribution is always the limiting distribution as contest size increases to infinity regardless of whether the contest is selective or inclusive.

# 4 Contests with capacity uncertainty

Contests are frequently used for selection, e.g., universities admit students based on the results of public examinations, firms choose between rival executive candidates based on on-the-job performance in similar tasks, and investors choose portfolio managers based on managers' relative performance. In these cases, the gain from selecting the contest winner over the loser does not arise per se from performance in the contests. Rather, winning is a signal of underlying ability or, in our terminology, capacity. Thus, the efficiency of contest-based selection depends on the strength of the performance–capacity relation. Since such selection mechanisms are only rational when capacity is not directly observable, contests in which contestants' capacities are private information are a natural focus of analysis. Moreover, the natural question to pose to such an analysis is how strong is the link between equilibrium performance and capacity.

To develop this analysis, we introduce capacity uncertainty into the model. We assume that each contestant has probability $\theta$ of being strong (S), with capacity equal to $\mu_S$, and probability $(1-\theta)$ of being weak (W), with capacity equal to $\mu_W$, where $0 < \mu_W < \mu_S$. A contestant's type is private information and independent of the types of the other contestants. Except for this uncertainty with respect to contestants' capacities, the contest remains the same as the one defined at the beginning of Section 3.

## 4.1 Probability of winning

In a symmetric equilibrium, the probability of winning function, $P$, faced by all the contestants is the same. Each contestant must play a best reply to this function conditional on his type, and the set of best-reply distributions is the same for all the contestants conditional on contestant type. Thus, we will discuss the best reply and optimal strategy for type $t \in \{S, W\}$, recognizing that we are in fact referring to the best reply for any contestant whose type is $t$.

To initiate the analysis of capacity uncertainty, first note that, for essentially the same reason as advanced in the capacity-certainty case, the conclusion of Lemma 1—that the equilibrium performance distributions are continuous—also holds in the capacity-uncertainty case. This result is recorded below.

**Lemma 2.** *In a symmetric equilibrium, the probability of winning function, P, is continuous and* $P(0) = 0$*.*

Next, note that, by (8), $P$ satisfies the following condition: for each type $t \in \{S, W\}$, there exist nonnegative scalars, $\alpha_t$ and $\beta_t$, such that

$$P(x) \leq \alpha_t + \beta_t x \quad \forall\, x \geq 0 \tag{16}$$

$$\text{Supp}_t \subset \{x \geq 0 : P(x) = \alpha_t + \beta_t\, x\}, \tag{17}$$

where $\text{Supp}_t$ denotes the support of $F_t$, the distribution selected by type $t$. Define $\psi$ as the concave lower envelope of the two upper support lines, $\{\alpha_t + \beta_t\, x\}_{t=S,W}$, associated with the two types, i.e.,

$$\psi(x) = \min[\alpha_S + \beta_S\, x,\ \alpha_W + \beta_W\, x]. \tag{18}$$

Equation (16) and the definition of the concave lower envelope, $\psi$, imply that

$$\forall t \in \{S, W\}, \quad \alpha_t + \beta_t\, x \geq \psi(x) \geq P(x), \tag{19}$$

i.e., $\psi$ lies between the support lines and $P$. Equations (17) and (19) imply that $\text{Supp}_S$ and $\text{Supp}_W$ are contained in the region where $\psi$ meets $P$. In fact, in a symmetric equilibrium, $P$ must trace out $\psi$ until $P$ reaches 1. The intuition behind the proof is that $P$ can only grow at points in $\text{Supp}_S$ and $\text{Supp}_W$. Since $\text{Supp}_S$ and $\text{Supp}_W$ rest on the points at which the concave lower envelope, $\psi$, meets $P$, to stay on the envelope, $P$ can never increase at a rate in excess of the envelope's rate of increase. As soon as $P$ breaks contact with the envelope, by equation (16), $P$ must stay below the envelope and, given that $P$ cannot jump up (see Lemma 2), $P$ cannot ever increase again. Admittedly, this argument is a bit loose, but it captures the essence of the formal proof.

**Proposition 9.** *There exist nonnegative constants,* $\alpha_S, \alpha_W, \beta_S, \beta_W$*, and* $\hat{x} > 0$*, such that, in a symmetric equilibrium, the probability of winning function, P, satisfies*

$$P(x) = \begin{cases} \min[\alpha_S + \beta_S\, x,\ \alpha_W + \beta_W\, x] & \text{if } 0 \leq x \leq \hat{x} \\ 1 & \text{if } x > \hat{x} \end{cases},$$

*where $\hat{x}$ is defined by*

$$\min[\alpha_S + \beta_S \hat{x},\ \alpha_W + \beta_W \hat{x}] = 1. \tag{20}$$

In essence, the probability of winning function, $P$, is the distribution a contestant plays against. Since $P$ is weakly concave and, by the analysis in Section 2, playing safe is a best reply to weakly concave continuous distributions, the type-$t$ contestant's expected probability of winning in equilibrium equals $P(\mu_t)$. Thus, the weak concavity of $P$ implies the weak concavity of the value of capacity.

**Corollary 1.** *The value of capacity is weakly concave.*

Although we assumed two possible types of contestants for analytical convenience, Corollary 1 holds true for any number of possible types, since no matter how many types there are, $P$ always traces out the concave lower envelope of the support lines of all the types until $P$ reaches 1. Thus, $P$ is always weakly concave, which implies the weak concavity of the value of capacity.

Since the support of $P$ equals the union of the supports of two types' equilibrium distributions, the following corollary is evident from Proposition 9.

**Corollary 2.** $Supp_S \bigcup Supp_W = [0, \hat{x}]$, *where $\hat{x}$ is defined by* (20).

Note that, for $t \in \{S, W\}$, $\text{Supp}_t$ must be contained in the set of points where $P$ meets the $t$-support line. For all values of $x$ at which the $S$-support line lies above the $W$-support line, the $S$-support line must lie above $P$. Thus, these values cannot be in $\text{Supp}_S$. Analogously, all values of $x$ at which the $W$-support line lies above the $S$-support line cannot be in $\text{Supp}_W$. The next lemma thus follows from Corollary 2.

**Lemma 3.** *In any symmetric equilibrium, $Supp_S$ and $Supp_W$ satisfy*

$$\begin{array}{c} Supp_S \subset \{x \in [0, \hat{x}] : \alpha_S + \beta_S x \leq \alpha_W + \beta_W x\} \\ Supp_W \subset \{x \in [0, \hat{x}] : \alpha_S + \beta_S x \geq \alpha_W + \beta_W x\} \end{array},$$

*where $\hat{x}$ is defined by* (20) *and $\alpha_S, \alpha_W, \beta_S$, and $\beta_W$ are the optimal dual variables.*

Next, in Lemma 4, we provide a characterization of the optimal dual variables in the maximization problems that define the support lines. These variables are generated by the dual problem (4) of the primal problem (2) for each type. Our results show that the $W$-support line always intersects the origin and that the slope of the $W$-support line is always weakly greater than the slope of the $S$-support line.

**Lemma 4.** *The optimal dual variables satisfy the following conditions: either (i) $\alpha_S = \alpha_W = 0$ and $0 < \beta_S = \beta_W$ or (ii) $\alpha_S > \alpha_W = 0$ and $0 < \beta_S < \beta_W$.*

## 4.2 Equilibrium configuration

Corollary 2 and Lemmas 3 and 4 provide a complete characterization of the supports of the equilibrium distributions of the two types. They suggest that, in equilibrium, there are only two candidate configurations for the probability of winning function, $P$. These configurations are illustrated in Figures 5 and 6. Figure 5 illustrates the case where $\alpha_S > \alpha_W$ and $\beta_S < \beta_W$. In this configuration, the interior of $\text{Supp}_W$ lies strictly below $\text{Supp}_S$, which implies that weak contestants concede to strong ones and concentrate their capacity on beating other weak contestants. We call equilibria with this configuration, "concession equilibria." Figure 6 illustrates the contrasting case where $\alpha_W = \alpha_S$ and $\beta_S = \beta_W$. In Figure 6, the union of $\text{Supp}_S$ and $\text{Supp}_W$ equals $[0, \hat{x}]$, but type distributions cannot be uniquely identified, since many different combinations of $F_S$ and $F_W$ can produce the probability of winning function that is linear over the union of both types' supports. In this configuration, the upper bound of $\text{Supp}_W$ lies strictly above the lower bound of $\text{Supp}_S$, which implies that a weak contestant's performance sometimes tops a strong contestant's. We thus call equilibria with this configuration, "challenge equilibria."[19]

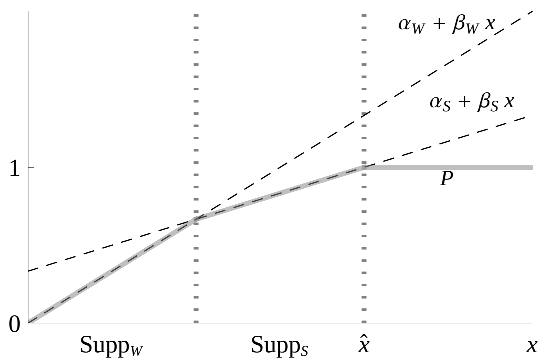



Figure 5: The probability of winning function, $P$, in concession equilibria

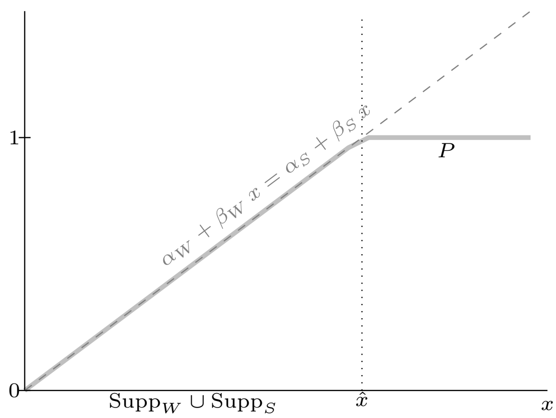



Figure 6: The probability of winning function, $P$, in challenge equilibria

Before we embark on systematic investigation of equilibrium strategies, it will be useful to illustrate our result by a simple example. In the example, there are two contestants and one prize, i.e., $n = 2$ and $m = 1$. Ex ante, each contestant is equally likely to be strong or weak, i.e., $\theta = 1/2$. Since, by Lemma 2, no one places point mass, the probability that a given contestant wins with a realized performance level, $x$, equals the

[19] There is also a non-generic borderline case in which the conditions on the optimal dual variables for challenge equilibria are satisfied but the upper bound of $\text{Supp}_W$ coincides with the lower bound of $\text{Supp}_S$. Since, in this borderline case, weak contestants still concede to strong ones, we categorize this case as a concession configuration.

probability that $x$ is no less than the realized performance level of the other contestant. Since the other contestant is equally likely to be weak or strong,

$$P(x) = \frac{F_S(x) + F_W(x)}{2}. \tag{21}$$

In challenge equilibria, $P$ is the CDF of a uniform distribution. Thus, (21) requires that the average of $F_S$ and $F_W$ equals a uniform distribution. This case is illustrated with $\mu_W = 1$ and $\mu_S = 2$ in Figure 7. Since $P$ is a uniform CDF and the average capacity of the two types is $3/2$,

$$\frac{x}{3} = \frac{F_S(x) + F_W(x)}{2} \quad \forall\, x \in [0, 3]. \tag{22}$$

Note that $F_W$ and $F_S$ are not unique. All that is required for $F_W$ and $F_S$ to be equilibrium distributions for weak and strong types, respectively, is the satisfaction of (22). Figure 7 presents a particular choice of distribution functions satisfying (22). In Figure 7, the distributions chosen by weak and strong contestants are as follows: the weak type plays the uniform distribution on $[0, 3/2]$ with probability 5/6 and plays the uniform distribution on $[3/2, 3]$ with probability 1/6; the strong type plays the uniform distribution on $[0, 3/2]$ with probability 1/6 and plays the uniform distribution on $[3/2, 3]$ with probability 5/6. Since $P$ is a uniform distribution, all the distributions with the same mean whose support is enclosed in $P$'s produce the same payoff when used against $P$. Thus, we can evaluate the probability of winning of the weak and of the strong by evaluating $P$ at $\mu_W$ and $\mu_S$, respectively. This yields a probability of winning equal to 1/3 for the weak type and 2/3 for the strong type.

It is not always possible to satisfy the conditions for a challenge equilibrium. For example, consider the case where $\mu_W = 1$ and $\mu_S = 5$. Given these parameters, challenge equilibria require that

$$\frac{x}{6} = \frac{F_S(x) + F_W(x)}{2} \quad \forall\, x \in [0, 6]. \tag{23}$$

Equation (23) implies that, over the support of $F_W$,

$$\frac{x}{6} \geq \frac{F_W(x)}{2}. \tag{24}$$

Integrating both sides of (24) over $F_W$ yields

$$\int \frac{x}{6} dF_W(x) \geq \int \frac{F_W(x)}{2} dF_W(x). \tag{25}$$

Given that $\mu_W = 1$, the left hand side of (25) equals 1/6. However, the right hand side of (25) equals $1/4$. Thus, (25) cannot hold. Thus, no challenge equilibria exist. In fact, quite generally, when the weak type's capacity is too small relative to the strong type's, it is never possible to generate a probability of winning function which is linear over the union of both types' supports. In this case, only "concession equilibria" exist. A concession equilibrium is illustrated in Figure 8. In this symmetric equilibrium, weak contestants use a uniform distribution over [0, 2] while strong contestants use a uniform distribution over [2, 8]. A weak contestant can win only when the other contestant is also weak, which occurs one half of the time. Moreover, if both contestants are of the same type, both are using the same distribution and thus have an equal probability of winning. Hence, a weak contestant's probability of winning is $(1/2)\times(1/2)=1/4$, and similarly, a strong contestant's probability of winning is $(1/2)+(1/2)\times(1/2)=3/4$.

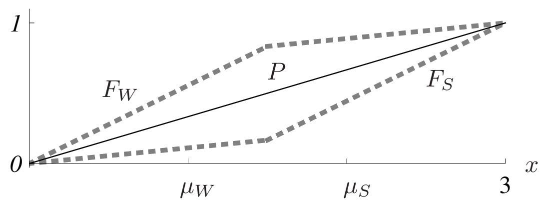



Figure 7: An example of a challenge equilibrium. In the example, $\mu_W=1$, $\mu_S=2$, the number of contestants, $n$, equals 2, and the number of prizes, $m$, equals 1. A contestant's probability of being strong, $\theta$, equals 1/2.

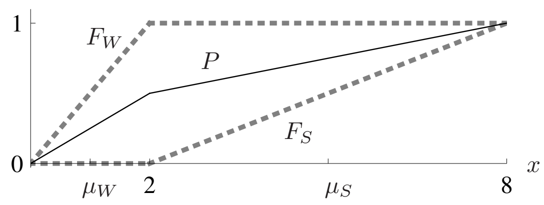



Figure 8: An example of a concession equilibrium. In the example, $\mu_W=1$, $\mu_S=5$, the number of contestants, $n$, equals 2, and the number of prizes, $m$, equals 1. A contestant's probability of being strong, $\theta$, equals 1/2.

Now consider the general case with $1\le m\le n-1$, $0<\theta<1$, and $0<\mu_W<\mu_S$. Recall that there are only two candidate equilibrium configurations: concession and challenge. We denote the concession configuration by $C$ and, for $t\in\{S,W\}$, denote by $p_t^C$ the type-$t$ contestant's expected probability of winning in the concession configuration. Since, in the concession configuration, weak contestants concede victory to strong contestants, the expressions for $p_S^C$ and $p_W^C$ are determined by the following prize-allocation rules: (a) strong contestants always have "absolute priority," i.e., no weak contestant wins a prize unless all strong contestants win and (b) contestants of the same type have the same chance of winning. These rules imply that $p_S^C$ and $p_W^C$ are uniquely determined by $n$, $m$, and $\theta$. Since we can analyze selection efficiency, which is the focus of this section, without presenting the expressions for $p_S^C$ and $p_W^C$, for the sake of brevity, we omit developing them.

We denote the challenge configuration by $G$ and, for $t\in\{S,W\}$, denote by $p_t^G$ the type-$t$ contestant's expected probability of winning in the challenge configuration.

Since, in the challenge configuration, both types' support lines overlap and intersect the origin, we must have $p_W^G : p_S^G = \mu_W : \mu_S$. By symmetry, ex ante each contestant has a chance of winning equal to $m/n$, so $\theta p_S^G + (1-\theta) p_W^G = m/n$. Hence,

$$p_S^G = \frac{m\mu_S}{n\left[\theta\mu_S + (1-\theta)\mu_W\right]} \quad \text{and} \quad p_W^G = \frac{m\mu_W}{n\left[\theta\mu_S + (1-\theta)\mu_W\right]}.$$

The next lemma shows that, which configuration emerges, for a given parameterization of the model, is determined by which configuration favors the weak type.

**Lemma 5.** *If $p_W^C$ and $p_W^G$ denote, respectively, the weak type's probability of winning under the concession and challenge configurations, concession equilibria exist if and only if $p_W^C \geq p_W^G$; challenge equilibria exist if and only if $p_W^C < p_W^G$.*

The intuition of Lemma 5 is not difficult to understand. Enabling contestants to choose performance distributions offers weak contestants the possibility of outperforming strong ones. However, challenging the strong may not always be optimal for a given weak contestant, since he does not know his competitors' types and, to challenge, he has to prolong the right tail of his performance distribution, which, through the capacity constraint, increases the probability of low performance, thus reducing his chance of winning when competing against weak competitors. Therefore, a weak contestant challenges the strong only when challenging benefits him.

## 4.3 Selection efficiency

Lemma 5 helps us examine selection efficiency in contests. Before we analyze selection efficiency, we need to settle on its definition. Selection efficiency of a mechanism is a characteristic of the mechanism but not the quality of contestants *per se*. The most efficient selection mechanism is a mechanism under which strong contestants have absolute priority. For a given contest, *maximum selection efficiency* is the probability that a selected contestant will be strong under the most efficient selection mechanism. So, for example, if there are two contestants and one prize, and the probability that a given contestant is strong equals 1/2, then the most efficient mechanism will select a strong contestant whenever at least one of the two contestants is strong. Since 3/4 of the time at least one contestant is strong, maximum selection efficiency equals 3/4. We denote maximum selection efficiency by $\Pi^*$. We compare maximum selection efficiency with *actual selection efficiency*, denoted by $\Pi$. Actual selection efficiency is the equilibrium probability of a selected contestant being strong in a symmetric equilibrium. Define $\Delta\Pi = \Pi^* - \Pi$, where $\Delta\Pi$ represents *selection efficiency loss*. A contest is said to be *efficient* if and only if $\Delta\Pi = 0$. While actual selection efficiency measures the quality of prize winners, selection efficiency loss measures the reduction in winner quality caused by a given contest mechanism. A contest mechanism has poor selection properties if it produces large selection efficiency losses.[20]

Denote by $\Pi^C$ and $\Pi^G$ the actual selection efficiency in concession and challenge equilibria, respectively. In concession equilibria, strong contestants have absolute priority, so $\Pi^C = \Pi^*$. In challenge equilibria, each contestant's probability of winning is proportional to his capacity, so by Bayes' Rule,

$$\Pi^G = \frac{\theta r}{(1-\theta)+\theta r}, \tag{26}$$

where $r = \mu_S/\mu_W > 1$, the ratio of the strong type's capacity to the weak type's. We interpret $r$ as a measure of strength asymmetry. By Lemma 5, the equilibrium configuration is always the one that favors the weak type, so

$$\Pi = \min\left[\Pi^G, \Pi^C\right], \tag{27}$$

where $\Pi^G$ is given by (26) and $\Pi^C = \Pi^*$. After deriving the expression for $\Pi^*$, the next proposition follows from (27) and the definition of $\Delta\Pi$.

**Proposition 10.** *Maximum selection efficiency,* $\Pi^*$*, actual selection efficiency,* $\Pi$*, and selection efficiency loss,* $\Delta\Pi$*, are given as follows:*

$$\Pi^* = \sum_{i=0}^{n} \binom{n}{i} \min\left[\frac{i}{m}, 1\right] \theta^i (1-\theta)^{n-i}; \tag{28}$$

$$\Pi = \min\left[\frac{\theta r}{(1-\theta)+\theta r}, \Pi^*\right]; \tag{29}$$

$$\Delta\Pi = \max\left[\Pi^* - \frac{\theta r}{(1-\theta)+\theta r}, 0\right]. \tag{30}$$

Using Proposition 10, we perform comparative statics on selection efficiency.

**Corollary 3.** *The comparative static results on* $\Pi^*$*,* $\Pi$*, and* $\Delta\Pi$ *are as follows:*

i. *For fixed* $\theta$*, n and m, maximum selection efficiency,* $\Pi^*$*, is constant in r, the strength asymmetry, actual selection efficiency,* $\Pi$*, is weakly increasing in r, and selection efficiency loss,* $\Delta\Pi$*, is weakly decreasing in r.*

[20] Our approach of using two metrics, actual selection efficiency and selection efficiency loss, contrasts with Hvide and Kristiansen (2003) who only use actual selection efficiency as the metric of efficiency. Like our paper, Ryvkin and Ortmann (2008) and Ryvkin (2010) also take into account both actual selection efficiency and selection efficiency loss, but in a somewhat different way: in these two papers, actual selection efficiency can be thought of as represented by the expected ability of the winner while selection efficiency loss by the expected ability rank of the winner.

*ii.* *For fixed r, n, and m, both maximum selection efficiency,* $\Pi^*$*, and actual selection efficiency,* $\Pi$*, are strictly increasing in* $\theta$*, the probability of a contestant being strong, whereas selection efficiency loss,* $\Delta\Pi$*, is nonmonotonic in* $\theta$*.*

*iii.* *For fixed* $\theta$*, r, and m, maximum selection efficiency,* $\Pi^*$*, is strictly increasing in n, the number of contestants, and both actual selection efficiency,* $\Pi$*, and selection efficiency loss,* $\Delta\Pi$*, are weakly increasing in n.*

*iv.* *For fixed* $\theta$*, r, and n, maximum selection efficiency,* $\Pi^*$*, is strictly decreasing in m, the number of prizes, and both actual selection efficiency,* $\Pi$*, and selection efficiency loss,* $\Delta\Pi$*, are weakly decreasing in m.*

*v.* *Fixing* $\theta$ *and r while increasing contest size by multiplying both n and m by a common integer factor, k, strictly increases maximum selection efficiency,* $\Pi^*$*, and weakly increases actual selection efficiency,* $\Pi$*, and selection efficiency loss,* $\Delta\Pi$.

Shifts in model parameters affect both selection efficiency loss, $\Delta\Pi$, and maximum selection efficiency, $\Pi^*$. Selection efficiency loss is minimized when weak contestants concede to strong ones. The likelihood that a weak contestant will concede is increased by (i) an increase in the degree of strength asymmetry, measured by $r$, (ii) a decrease in the number of contestants, $n$, and (iii) an increase in the number of prizes, $m$. Selection efficiency loss is not monotonic in contestant quality, measured by $\theta$. Increasing $\theta$ increases a weak contestant's benefit from challenging the strong. Once this benefit exceeds a threshold, only challenge equilibria can be supported. This effect raises $\Delta\Pi$ above zero. However, as $\theta$ continues to increase toward 1, the probability that a contestant is weak decreases to 0, in which case $\Delta\Pi$ approaches 0 again.

The effect of a parameter shift on maximum selection efficiency, $\Pi^*$, is solely determined by the effect of the shift on the distribution of contestant quality. Since maximum selection efficiency depends on the probability that winners are strong but not on the winners' absolute strength, changes in absolute strength, $\mu_S$ and $\mu_W$, or relative strength, $r$, do not affect maximum selection efficiency. In contrast, an increase in the number of contestants, $n$, a decrease in the number of prizes, $m$, and an increase in each contestant's probability of being strong, $\theta$, all increase maximum selection efficiency. Although, the effect is a bit more subtle, increasing contest size also increases maximum selection efficiency: when there are few contestants, the realized proportion of strong contestants is more likely to deviate from its expected value, $\theta$. If more strong contestants are drawn than prizes, the excess of strong contestants has no positive effect on winner quality, but if less are drawn, it has a negative effect. Thus, random variation in contestant quality lowers maximum selection efficiency. By the weak law of large numbers, scaling up the contest reduces random variation in the realized fraction of strong contestants and, hence, increases maximum selection efficiency.

The effect of a parameter shift on maximum selection efficiency is mechanical; the effect on selection efficiency loss is strategic. A change in parameters, such as a change in $n$ or $m$, frequently has the same directional effect on maximum selection efficiency and selection efficiency loss, increasing or decreasing both. Since actual selection efficiency is the difference between maximum selection efficiency and selection efficiency loss, this directional alignment could potentially lead to ambiguous comparative statics. However, our results imply that, whenever this happens, the mechanical effect is always weakly dominant. Thus, the effect of a change in parameters on actual selection efficiency is always weakly monotonic. This result can be contrasted with Hvide and Kristiansen (2003), who find that an increase in $n$ or $\theta$ can sometimes decrease actual selection efficiency. Their result is driven by the assumption that contestants can only choose between a constant and a Bernoulli-distributed random variable. This assumption prevents strong contestants from playing win-small/lose-big strategies to better accommodate the challenge brought by weak contestants and, hence, amplifies the negative strategic effect on actual selection efficiency when $n$ or $\theta$ increases.

# 5 Selection efficiency of modified contest mechanisms

In Section 4, we studied selection efficiency of a simple contest mechanism in which contestants compete for several identical prizes with no restriction imposed on distributional choices apart from nonnegativity and capacity. In this section, we treat this simple mechanism as the benchmark and study whether a principal who seeks to maximize selection efficiency could do better with some simple modifications of the benchmark mechanism. As Corollary 3 has already performed comparative statics on selection efficiency with respect to the number of contestants, the number of prizes, the quality of contestants, and strength asymmetry, to avoid repetition of analysis, in this section, we assume that all these contest parameters are fixed for the principal. Thus, maximum selection efficiency is fixed, in which case the two metrics—actual selection efficiency and selection efficiency loss—used for efficiency evaluation are equivalent. Thus, without loss of generality, in what follows, we use actual selection efficiency as the base of our analysis. We consider three alternative contest arrangements that are close to the benchmark mechanism and commonly observed in practice.

## 5.1 Scoring caps

Suppose that performance is capped by a certain level, $\bar{x}$, so that contestants are restricted to distributions satisfying $F(\bar{x}) = 1$. The use of a scoring cap imposes an upper bound on contestants' performance levels. Many real-life contests have a scoring cap. For example, in examinations and many sports games, contestants' performance levels are bounded by full scores. A principal can change this upper bound by changing the difficulty of reaching a full score. A scoring cap can be imposed if the principal can credibly specify that all performance levels no less than $\bar{x}$ will be treated the same for the purpose of determining contest winners. Under this specification, contestants have no incentive to assign probability weight to performance levels exceeding $\bar{x}$.

When $\bar{x} \in [\mu_W, \mu_S)$, strong contestants cannot fully utilize their capacity while weak contestants can. Thus, a scoring cap between $\mu_W$ and $\mu_S$ handicaps strong contestants. It is obvious that imposing a scoring cap less than $\mu_W$ also handicaps strong contestants as it forces them to utilize the same amount of capacity as weak ones. Hence, imposing a scoring cap strictly less than $\mu_S$ never improves selection efficiency. Thus, in what follows, we focus on the case where $\bar{x} \geq \mu_S$.

A performance level equal to the scoring cap cannot be topped. Thus, in equilibrium, point mass on the scoring cap can possibly occur, which implies that the continuity of the probability of winning function, $P$, may break at the scoring cap.

**Lemma 6.** *In a symmetric equilibrium, the probability of winning function, $P$, intersects the origin and is continuous on $[0, \bar{x})$, where $\bar{x}$ denotes the scoring cap.*

Although a discontinuity of $P$ can occur at $\bar{x}$, by Lemma 6 and the fact that $P$ is bounded above by $P(\bar{x})$, $P$ is upper semicontinuous. Given that $P$ is nondecreasing and bounded, upper semicontinuity of $P$ guarantees the existence of a best reply to $P$ and makes the analysis in Section 2.1 applicable here. Hence, conditions (16) and (17) still hold, which implies that the support of contestants' performance distributions must still fall in the range where the probability of winning function, $P$, meets $\psi$, the concave lower envelope of the two upper support lines. In this case, Lemma 4 still holds, which implies that there are still only two candidate equilibrium configurations: concession and challenge. Since both types' configuration-conditioned payoffs are determined by the configuration-conditioned prize allocation rules presented in Section 4.2, and also since imposing a scoring cap $\bar{x} \in [\mu_S, \infty)$ does not change any prize-allocation rule, both types' configuration-conditioned payoffs are unaffected by the scoring cap. Since the equilibrium configuration is the one with the prize allocation rules favoring the weak type, which result is given by Lemma 5, whose proof is unaffected by the existence of a scoring cap, the equilibrium configuration must be unaffected by the scoring cap $\bar{x} \in [\mu_S, \infty)$. Thus, we obtain the next result.

**Proposition 11.** *Imposing a scoring cap, weakly greater than the strong type's capacity, does not affect any type's probability of winning or selection efficiency.*

Although the use of a scoring cap greater than $\mu_S$ does not affect selection efficiency, it changes the equilibrium distribution: any weight that is originally placed above $\bar{x}$ is now moved to $\bar{x}$ and, to balance its effect on the mean, some weight that is originally placed over a neighborhood range below $\bar{x}$ is also moved to $\bar{x}$. If we lower $\bar{x}$, more and more probability weight will be placed on $\bar{x}$ rather than spread over the neighborhood of $\bar{x}$, which results in a simple mean-preserving contraction of contestant performance, the opposite operation of a simple mean-preserving spread. Hence, a principal who is averse to contestant performance riskiness prefers a tighter scoring cap, which, by Proposition 11, has no side effect on selection efficiency if the cap exceeds $\mu_S$.

**Proposition 12.** *A principal who is averse to contestant performance riskiness weakly prefers a tighter scoring cap, if the scoring cap is weakly greater than* $\mu_S$.

## 5.2 Penalty triggers

In the previous sections, contestants were not penalized for bad performance. In this section, we consider whether selection efficiency can be improved by the use of "penalty triggers," i.e., penalties for performance levels below $\underline{x} > 0$. We consider the case where $\underline{x} \leq \mu_W$ so that it is possible for weak contestants to completely avoid the penalty. We assume that the penalty is sufficiently large relative to the prize for winning so that no contestant has an incentive to place any weight below $\underline{x}$. Thus, the use of penalty triggers imposes a lower bound, $\underline{x}$, on the performance levels chosen by contestants.

When penalty triggers are used, for $t \in \{S,W\}$, we can think of type-$t$'s performance as the sum of two parts: a safe performance level equal to $\underline{x}$, submitted to avoid the penalty, and a random performance level whose expected value equals $\mu_t - \underline{x}$, submitted to maximize the chance of winning. Thus, the contest is essentially the same as the one studied in Section 4 except that now each type-$t$ contestant's "manipulable" capacity is reduced to $\mu_t - \underline{x}$. An increase in $\underline{x}$ increases $(\mu_S - \underline{x})/(\mu_W - \underline{x})$, the ratio of $S$'s manipulable capacity to $W$'s. The next proposition thus follows from (i) in Corollary 3.

**Proposition 13.** *Increasing the threshold that triggers the penalty weakly improves selection efficiency.*

Proposition 13 suggests that, if a principal wants to use contests to promote the most able employees, penalty triggers will be beneficial to selection efficiency. Proposition 13 thus provides an explanation for why mixed incentive systems with both "carrots" and "sticks" are often found in practice.

Since penalty triggers can improve selection efficiency, which is the major concern of the principal in our model, we only briefly talk about their effect on performance riskiness. The next result implies that the effect is ambiguous.

**Proposition 14.** *If the contest without penalty triggers has a concession equilibrium, then imposing penalty triggers will induce a mean-preserving contraction of the weak type's performance while a mean-preserving spread of the strong type's.*

Since the unconditional performance distribution is a convex combination of the two types' distributions, Proposition 14 implies that the effect of penalty triggers on performance riskiness is ambiguous. Thus, in contrast to the use of scoring caps, which reduces performance riskiness but never improves selection efficiency, the use of penalty triggers increases selection efficiency but has an ambiguous effect on performance riskiness. Thus, scoring caps and penalty triggers are complements rather than substitutes.

## 5.3 Localizing contests

In this subsection, we study the effect on selection efficiency of dividing the original grand contest involving $n$ contestants and $m$ prizes into multiple smaller local contests with contestants in the same local contest competing for prizes allocated to that local contest. An admissible division requires the number of contestants in each local contest summing up to $n$ and the number of prizes summing up to $m$. The next proposition shows that localizing a contest weakly lowers selection efficiency.

**Proposition 15.** *Localizing a contest never improves selection efficiency. The effect is neutral if and only if every local contest has challenge equilibria.*

The intuition of Proposition 15 is easiest to understand if we look at the effect of the opposite operation — grouping local contests together. Grouping local contests acts in a similar way as scaling up those local contests, which by (v) in Corollary 3, weakly increases selection efficiency. When every local contest has challenge equilibria, implying that, in every local contest, the competition is so intense that weak contestants cannot afford to concede to strong ones, grouping local contests into a bigger contest will never bring down the intensity of competition, so the bigger contest will also induce challenge equilibria. Since, in challenge equilibria, actual selection efficiency is given by (26), which is independent of contest size, selection efficiency remains unchanged when grouping local contests with challenge equilibria.

# 6 Applications

Our framework can be applied to study capacity-constrained contests where contestants have great freedom in manipulating performance distributions. We illustrate the application to mutual fund tournaments in Section 6.1. In Section 6.2, we study R&D contests in which prize value is not fixed but dependent on winner performance. Although winning small and winning big are no longer payoff equivalent there, we show that contestants still have win-small/lose-big preferences and these preferences lead to risk-taking strategies safer than social optimum. In Section 6.3, we show how our framework can be applied to study stochastic contests in which each contestant decides when to stop a privately observed stochastic process. This application adds a new interpretation of our framework.

## 6.1 Mutual fund tournaments

Many studies identify a convex relation between a mutual fund's ranking and its capital inflows (Chevalier and Ellison (1997), Sirri and Tufano (1998), and Huang, Wei, and Yan (2007)). Since fees charged by mutual funds are linked to assets under management, increased capital inflows lead to higher profits. Thus, funds have strong incentives to be "top performers." At the same time, the capital inflows generated by a mediocre and a bottom ranking are not very different, since, in either case, the fund is unlikely to attract ranking-motivated investors. For this reason, we illustrate the implications of our results in the context of mutual fund tournaments, in which $n$ managers compete for the top-$m$ ranks with all the winners receiving the same positive payoff and all the losers receiving zero. When $m$ is small relative to $n$, this payoff structure approximates the "convexity" of fund flows documented in the literature.

Let $\mathscr{N} = \{1, \cdots, n\}$ be the set of $n$ managers. Following Berk and Green (2004), we express the return of manager $i \in \mathscr{N}$, in excess of the passive benchmark, as

$$R_i = \mu_i + \varepsilon_i,$$

where $\mu_i$ is a constant, determined by manager $i$'s capacity, and $\varepsilon_i$ is fund $i$'s idiosyncratic risk. In contrast to Berk and Green (2004) in which the distribution of $\varepsilon$ is exogenously specified, we assume that, by using dynamic trading strategies, manager $i \in \mathscr{N}$ can choose the distribution of $\varepsilon_i$, subject to $E[\varepsilon_i] \leq 0$.[21] To use the analysis developed earlier, which assumed that the support of a performance distribution is bounded below by a constant, we assume that, manager $i \in \mathscr{N}$ will be penalized, such as being fired,

[21] Theorem 4 in Cover (1974) shows that any return distribution on the nonnegative real line subject to a mean constraint can be produced by a sequential gambling scheme on fair coin tosses. Rubinstein and Leland (1981) find that a fund manager can create an option position by using dynamic trading strategies even if options do not exist in the market. These results suggest that, even with a small number of return distributions directly provided by the assets in the market, fund managers are able to have access to a much larger set of return distributions by using dynamic trading strategies.

if $R_i$ falls below a certain threshold, say $\underline{R}$, where $\underline{R} < \min\{\mu_i\}_{i\in\mathcal{N}}$, and this penalty is sufficiently large such that manager $i$ has no incentive to place probability weight on excess return levels below $\underline{R}$. Under these assumptions, each manager $i \in \mathcal{N}$ independently chooses the distribution of $R_i$ on $[\underline{R}, \infty)$, subject to his capacity constraint, $E[R_i] \leq \mu_i$.

When all the managers have the same fixed $\mu$, these assumptions map mutual fund tournaments into the certain-capacity contest model studied in Section 3. Based on the results we obtained there, we make three predictions about fund managers' risk-taking strategies. First, by Proposition 3, neither the safest nor the riskiest strategy is played in equilibrium. This result is in sharp contrast to Chen, Hughson, and Stoughton (2012) who study fund managers' risk-taking strategies in mutual fund tournaments and find that equilibrium strategies are always extremal. Their result is largely driven by their symmetric return distribution assumption. Empirical evidence seems to favor our prediction of limited risk taking. For example, as shown by Falkenstein (1996), mutual funds shun both high and low volatility stocks. Second, given that mutual fund tournaments are selective, i.e., "star" funds are a minority of the total population of competing funds, Proposition 5 predicts positive skewness of the idiosyncratic risk of funds. This prediction coincides with the empirical evidence from Wagner and Winter (2013). Third, by Propositions 5 and 6, we predict that, when the competition between funds becomes more intense (maybe because of cooling down of market hotness, which requires funds to be exceptionally outstanding to attract a higher volume of capital inflows), funds take more idiosyncratic risk and both the dispersion and the skewness of the idiosyncratic risk increase.

Almost all mutual funds belong to a mutual fund family. Kempf and Ruenzi (2008a) and Kempf and Ruenzi (2008b) find evidence for intrafirm competition among fund managers within the same mutual-fund family for marketing expenses. When managers' capacities, $\mu$, vary and managers are privately informed about their capacity, the problem for the top management of a mutual fund family of ensuring efficient allocation of marketing resources to managers with highest capacities becomes salient. Given that the top management typically directs marketing resources toward the best-performing funds within the family, our results in the capacity-uncertainty case can be applied. Proposition 13 suggests that, to improve allocation efficiency, the top management can raise up the penalty trigger $\underline{R}$. Propositions 11 and 12 suggest that the top management can reduce funds' risk and maintain allocation efficiency by using a scoring cap, i.e., committing to offer all the managers with excess returns hitting or exceeding the cap the same chance of obtaining the marketing resources.

## 6.2 R&D contests

Does market competition bias firms against risky R&D strategies? This question was first addressed by Dasgupta and Stiglitz (1980) with a model that can be thought of as a variation of our earlier setup. In their model, several identical contestants (research units) compete for one prize. Each contestant independently chooses a distribution of discovery times for an innovation subject to the same mean-discovery time constraint. The contestant with the earliest realized discovery time wins the competition. However, in contrast to our setup, the winner's prize is not fixed but discounted continuously at a fixed rate. Dasgupta and Stiglitz (1980) restrict contestants' risk choices to a sequence of distributions that differ from one another by a mean-preserving spread and find that the equilibrium strategy is safer than social optimum. However, as pointed out by Klette and de Meza (1986), the derivation in Dasgupta and Stiglitz (1980) is erroneous: Dasgupta and Stiglitz assume both costless risk taking and an interior solution to risk taking in the contest, which are often inconsistent assumptions. Following Dasgupta and Stiglitz (1980), Klette and de Meza (1986) further impose a symmetric distribution assumption on admissible risk choices, and, under this symmetry assumption, show that contestants play riskiest strategies in equilibrium and thus the market equilibrium cannot be safer than social optimum.[22] In an independent work, Bhattacharya and Mookherjee (1986) find results consistent with Klette and de Meza (1986). Both Bhattacharya and Mookherjee (1986) and Klette and de Meza (1986) acknowledge that the symmetric distribution assumption plays a crucial role in their analysis.

We reexamine the question originally raised by Dasgupta and Stiglitz (1980) by following the model structure of Dasgupta and Stiglitz (1980) and Klette and de Meza (1986) while relaxing the distributional restrictions imposed on risk choices. To simplify the mathematical expressions, we assume that there are just two contestants. Extension to the multi-contestant case does not change the qualitative nature of the result. Following Klette and de Meza (1986), we assume that there exists a time by which it is certain that a contestant discovers the innovation. Denote this time by $T^*$. Under this assumption, the support of an admissible distribution is contained in $[0, T^*]$ and the riskiest strategy is the Bernoulli distribution with all weight placed on 0 and $T^*$. Society only benefits from the first innovation, which produces a nominal social value equal to $V_s$. If the first innovation takes place at time $t$, the discounted social value is $e^{-rt}V_s$, where $r$ represents the discount rate. Given this social payoff function, society

[22] Klette and de Meza (1986) further argue that, although, both society and contestants prefer the same risk-maximizing strategy, the contest effect makes the contestants' gain from risk-taking greater than society's. Thus, if a cost of risk-taking were introduced into the analysis, contestants might choose excessively risky strategies.

prefers both contestants to play the riskiest strategy (Klette and de Meza (1986) and Bhattacharya and Mookherjee (1986)). This is because of the "option-effect": fixing one contestant's realized performance, social payoff is convex in the other contestant's performance, so society is risk-seeking and prefers the other contestant to play the riskiest strategy. Applying this argument to both contestants shows that the unique social optimum is to have both contestants play the riskiest strategy.

Now consider contestant strategies. Note that, in Dasgupta and Stiglitz (1980) and Klette and de Meza (1986), performance is measured by the discovery time, $t$. In this case, the smaller the measure, the better the performance. As this is not convenient for us to apply our earlier framework, we make one cosmetic change by measuring performance as the time saved from a quicker discovery, interpreted as innovation speed. Denote this new performance measure by $x$, where $x = T^* - t$. With this change, the larger the measure, the better the performance. The contest is between two equally matched contestants, both independently choosing a performance distribution on the closed interval $[0, T^*]$ with the mean no greater than $\mu$, where $\mu \in (0, T^*)$ denotes a contestant's capacity. The one with better performance wins and ties are broken randomly. The winner's payoff when his realized performance equals $x$ is $e^{-r(T^*-x)}V_c$, where $V_c$ represents the nominal value of the winner's prize. The loser's payoff is 0. The following proposition presents the strategies played in a symmetric equilibrium.

**Proposition 16.** *In the R&D contest, the symmetric equilibrium is given as follows:*

*i. if* $\hat{x} \leq \min[T^* e^{r(\hat{x}-T^*)}, 1/r]$, *where* $\hat{x}$ *is implicitly defined by*

$$\hat{x} + \frac{1}{r} + \frac{1 - e^{r\hat{x}}}{r^2 \hat{x}} = \mu, \tag{31}$$

*there exists a symmetric equilibrium in which both contestants play*

$$F(x) = \begin{cases} \frac{x}{\hat{x}} e^{r(\hat{x}-x)} & \text{if } 0 \leq x \leq \hat{x} \\ 1 & \text{if } x \geq \hat{x} \end{cases}; \tag{32}$$

*ii. otherwise, in a symmetric equilibrium, both contestants play*

$$F(x) = \begin{cases} \frac{x}{2T^* - x' e^{r(T^*-x')}} e^{r(T^*-x)} & \text{if } 0 \leq x \leq x' \\ \frac{x'}{2T^* e^{r(x'-T^*)} - x'} & \text{if } x' \leq x < T^* \\ 1 & \text{if } x \geq T^* \end{cases}, \tag{33}$$

*where* $x'$ *is determined by making the capacity constraint bind.*

When $r$ goes to 0, prize value is not discounted and the R&D contest becomes the contest with a scoring cap studied in Section 5.1. In this case, the condition in (i) of Proposition 16 reduces to $\hat{x} \leq T^*$. When the cap constraint is not binding, i.e., when $\hat{x} \leq T^*$, both contestants submit uniformly distributed random performance, which is implied by (32) with $r \to 0$. When the cap constraint is binding, both contestants mix between a uniform distribution over a common interval and a degenerate random performance at the scoring cap, $T^*$, which is implied by (33) with $r \to 0$.

Introducing the time discounting of prize value has two effects on contestant strategies. First, if the scoring cap is not binding, a convex prize-value function leads to a right-skewed equilibrium strategy, even when there are just two contestants. This is because an equilibrium requires a constant marginal incentive for a contestant over his distribution support, and thus, to produce this constant marginal incentive, the convex prize-value function must be counterbalanced by a concave probability of winning function over a contestant's distribution support. Since, in equilibrium, no one places point mass on $[0, T^*)$ (see Lemma 6), if the scoring cap is not binding, a concave probability of winning function implies a concave CDF played by the competitor, which further implies the right-skewness of the equilibrium strategy. The larger the discount rate, $r$, the more right-skewed the equilibrium strategy. By Proposition 5, this right-skewness will be further enhanced if there are more than two contestants.

Second, an increase in the convexity of the prize-value function, produced by an increase in $r$, makes it more likely that the scoring cap is binding and contestants place point mass on the scoring cap. There are basically two reasons for that. First, increasing the convexity of the prize-value function increases the right-skewness of the equilibrium distribution, which prolongs the right tail of the distribution and makes it more likely that the right tail hits the scoring cap. Second, increasing the convexity of the prize-value function by an increase in the discount rate, $r$, does not affect prize value at the maximum performance level permitted by the scoring cap (immediate innovation) but reduces it at other performance levels. Thus, increasing prize-value discount gives contestants more incentives to put weight on immediate innovation.

Note that, although in the R&D contest, prize value depends on winner performance, the rank-dependent reward feature and the jump in contestant payoff created by winning persist, which generates a preference for win-small/lose-big strategies. This preference, through strategic interaction, concavifies the performance-payoff relation and, in contrast to Bhattacharya and Mookherjee (1986) and Klette and de Meza (1986), leads to equilibrium strategies safer than social optimum. Thus, our result suggests that the implication of Dasgupta and Stiglitz (1980) holds if contestants have access to a wider set of risk-taking strategies that are not restricted by symmetry or the mean-preserving-spread constraints. Moreover, when the condition in (i) of Proposition 16 is satisfied, which requires $T^*$ being neither too small nor too large and $r$ being sufficiently small, the equilibrium strategy is much safer than social optimum; no one gambles on the immediate innovation, and every contestant randomizes performance levels over a short range. In this case, society cannot have a quick innovation and much of the "option-value" of multiple independent research units is lost.

## 6.3 Stochastic contests

In Section 6.1, we applied our framework to mutual fund tournaments. In fact, funding competition also takes place in alternative investment sector. For example, private equity (PE) funds whose performance is "top quartile" among the PE funds started in the same vintage year receive special attention from investors. Investors select PE funds and benchmark performance based on top-quartile results. As a result, PE managers seek a top-quartile ranking, which benefits the marketing of new funds (Harris, Jenkinson, and Stucke (2012)).

In contrast to mutual funds, which are usually open-end, PE funds typically have a limited partnership structure with a finite lifespan, usually 10-13 years. During the life of a PE fund, the PE managers make investments and eventually exit from the investments to realize capital gains (Cumming and Johan (2009)). PE managers have significant discretion over the timing of exit (Robinson and Sensoy (2013)). The realized returns after exit are, in most cases, a PE fund's final performance, because, typically, PE managers are contractually not allowed to reinvest the proceeds from earlier investments (Cumming and Johan (2009)).

Here, we model the competition among PE managers as a stochastic contest in which each manager decides when to exit investments to realize capital gains. Although the return of the investments follows a pre-specified stochastic process, a manager can choose among different return distributions by choosing among different exit timing strategies. We show that our framework can be applied to study this contest when the stochastic process is privately observable.

Suppose there are $n$ managers competing for $m$ identical prizes. Manager $i \in \{1, \cdots, n\}$ *privately* observes the continuous-time realization of his stochastic process $X^i = (X^i_t)_{t \in \mathscr{R}^+}$ with $X^i_0 = \mu_i$, where $\mu_i > 0$ is a constant. The private information assumption here is plausible, because PE funds historically, in most legal domiciles, have not been required to submit public audited, periodic return reports, and they normally provide their investors with periodic reports under confidentiality agreements that prevent information sharing with the public (Harris, Jenkinson, and Stucke (2012)). We assume that the processes $X^i$ are independent. Manager $i$ decides when to stop his own process $X^i$, based on his observation of $X^i$. We assume that stopped values are private information until all the managers stop their processes, and every manager has to stop before a deadline $T < \infty$. We assume that $T$ is sufficiently large. The $m$ managers with the highest stopped values each win a prize. This assumption is consistent with the practice that money multiple is commonly used as a measure of PE performance (Harris, Jenkinson, and Stucke (2012)). Ties are broken randomly.

Seel and Strack (2013) are the first to study this kind of stochastic contests. They show that the problem of finding the optimal stopping strategy can be reduced to the problem of finding the optimal performance distribution. This is because a contestant's stopping strategy affects the contestant's payoff only via its impact on the contestant's performance distribution. Thus, to solve for the equilibrium, one can first find the equilibrium performance distribution and then verify that there exists a stopping strategy that induces the equilibrium performance distribution. The second step is known, in the probability literature, as Skorokhod embedding problem (Skorokhod (1965)), which studies whether a distribution is feasible by stopping a stochastic process.

We consider two types of stochastic processes here. One is Brownian motion absorbed at zero (a manager has to stop if his Brownian Motion hits zero). The other is geometric Brownian motion. The existing literature, e.g., Ankirchner, Hobson, and Strack (2014), shows that any distribution that can be induced by stopping these two stochastic processes in bounded time must lie on the nonnegative real line and have its mean weakly less than the initial value of the process. Thus, the set of feasible distributions in these stochastic contests is included in the set of feasible distributions in our earlier framework, with the initial value of a stochastic process representing a contestant's capacity. When contestants are ex ante homogeneous, the certain and uncertain capacity cases are analyzed in Sections 3 and 4 respectively, with contestants' equilibrium performance distributions being Complementary Beta. Since a Complementary Beta distribution is absolutely continuous and has a compact support and positive density everywhere on its support, Proposition 5 in Ankirchner, Hobson, and Strack (2014) implies that Complementary Beta distributions can be induced by stopping these two stochastic processes in bounded time.[23] Thus, the equilibrium and comparative static results presented in Sections 3 and 4 still hold under the stochastic contest framework.

To study managers' equilibrium exit timing strategies, we need to construct a stopping strategy that induces the equilibrium performance distribution. However, such construction is not unique. Given that PE funds have a finite lifespan, it is natural to

[23]Proposition 5 in Ankirchner, Hobson, and Strack (2014) is presented in Result 2 in the proof of Lemma 7.

examine the *minimal* stopping time, which is, loosely speaking, the quickest way to induce a given distribution. This minimality concept was first introduced by Monroe (1972) and defined as follows.

**Definition 2.** *A stopping time $\tau$ for the process $X$ is* minimal *if whenever $\iota \leq \tau$ is a stopping time such that $X_\iota$ and $X_\tau$ have the same distribution, then $\iota = \tau$ almost surely.*

Note that the stopping time can depend on the progress of the stochastic process, so it is a random variable. The next lemma implies that an increase in risk taking coincides with an increase in the expected minimal stopping time.

**Lemma 7.** *Suppose a probability distribution $F'$ is a mean-preserving spread of $F$. Let $\tau'$ and $\tau$ be a minimal and integrable stopping time for inducing $F'$ and $F$, respectively, by stopping a stochastic process $X$. Suppose $X$ is either a Brownian motion absorbed at zero or a geometric Brownian motion, with its initial value equal to the mean of $F'$ and $F$. Suppose $\tau'$ and $\tau$ exist. Then $E[\tau'] > E[\tau]$.*

As shown in Propositions 6 and 7, when contestants have the same fixed capacity, an increase in contest selectivity or contest size leads to a mean-preserving spread of the equilibrium distribution. Thus, the next proposition follows from Lemma 7.

**Proposition 17.** *Suppose the stochastic process is a Brownian motion absorbed at zero or a geometric Brownian motion. When $\mu_i = \mu$ for all $i \in \{1, \cdots, n\}$, an increase in contest selectivity (an increase in $n$ or a decrease in $m$) or an increase in contest size (an increase in $n$ and $m$ by the same proportion) increases the expected minimal stopping time to induce the equilibrium performance distribution.*

Given that PE managers compete for a top-quartile ranking, Proposition 17 implies that, as the number of competing PE funds increases, and thus the PE funding contest is scaled up, the mean period between investment and realization will increase.

# 7 Conclusion

In this paper, we studied contestants' risk-taking strategies in contests in which contestants are free to choose performance distributions subject only to a capacity constraint on mean performance. We showed that the rank-dependent reward feature of contests gives capacity-constrained contestants win-small/lose-big preferences, which, contrary to the results in most of the literature, always leads them to eschew risk-maximizing strategies. In the case of symmetric known capacity, we derived closed-form solutions for equilibrium performance distributions and analyzed the effects of contest structure on equilibrium behavior. We found that contestants prefer positive (negative) skewness when the contest is selective (inclusive) and increasing contest selectivity increases both the variance and the skewness of contestant performance. We then extended the analysis to the case where contestants are unaware of each other's capacities. In this setting, we characterized equilibria and analyzed the effects of changing contest parameters on strategies, payoffs, and overall contest efficiency. We showed that, contrary to the risk-taking-and-ruin intuition, weaker contestants do not always gamble on high-risk strategies and that, when the capacities of weak and strong contestants are sufficiently different, the contest mechanism produces perfect selection efficiency. We then considered the effects of various modifications of the contest mechanism and applied our results to mutual fund tournaments, R&D contests, and stochastic contests.

As well as contributing to the literature on endogenous risk taking in contests, our paper also indirectly contributes to the solutions of Colonel Blotto games. In the simplest form of these games, two contestants simultaneously decide on how to assign their "use-it-or-lose-it" resources to different battlefields; the one who assigns more resources to a battlefield wins that battlefield and each contestant's objective is to win as many battlefields as possible. Typically, these games have no pure strategy equilibrium and the construction of a mixed strategy equilibrium can be technically challenging.[24] Colonel Blotto games have a relaxed version, called General Lotto games, in which the resource constraint only has to be satisfied on average (Hart (2008)). While game theorists have focused primarily on Colonel Blotto models, more applied modelling of contests, such as electoral contests, in economics and political science (Myerson (1993), Lizzeri (1999), Lizzeri and Persico (2001), and Sahuguet and Persico (2006)) has typically been developed in a General Lotto framework. Because our model can be viewed as a General Lotto model, our results extend the analysis of General Lotto games from the two-contestant/one-prize, complete information, setting to a multi-contestant/multi-prize setting under both complete and incomplete information.[25] As pointed out by Hart (2008), solving General Lotto games is a first step toward solving the associated Colonel Blotto games.[26]

Given the simplicity of our underlying model structure and the algorithm we devel-

[24]See Roberson (2006) for a complete analysis of the two-contestant Colonel Blotto game with complete information. See Kvasov (2007) and Barelli, Govindan, and Wilson (2014), respectively, for discussions of a non-zero-sum and a majority-rule version of Colonel Blotto games.

[25]In an earlier version of the paper, we included a General Lotto type of application of our framework to student examinations where students, competing for college admissions, decide on how to allocate their limited human capital to different potential exam questions when they prepare for the exam. This application is available from the authors upon request.

[26]The second step is to construct a joint resource distribution that satisfies the overall resource constraint and has its marginal resource distribution for each battlefield equal to the distributional choice derived from the associated General Lotto game.

oped for finding equilibrium solutions, there is significant room to extend the analysis without losing tractability. As shown by our analysis of R&D contests in Section 6.2, our framework can be easily modified to permit the value of the prize to be dependent on the performance levels selected by winning contestants. *Mutatis mutandis*, our analytical machinery can be targeted at any sort of contest in which (i) contestants' payoffs are rank-dependent and (ii) their strategies are capacity-constrained but otherwise flexible.

Thus, the model is extensible in a number of directions. First, our analysis could be extended to the case where risk taking is costly. The cost of riskiness could be based on performance variance, skewness, or entropy.[27] This extension could investigate the effects of the cost of riskiness on contestants' risk-taking behavior and selection efficiency. Second, we could endogenize capacity through a two-stage model, with capacity building through costly effort bidding in the first stage followed by capacity-constrained distributional choice in the second stage. By taking into account the interaction between effort choice and risk choice, this extension could evaluate contest designs from a wider perspective, including effort incentive provision, risk taking, and selection efficiency. Third, we could extend the analysis by considering asymmetric contests where heterogeneous contestants know each other's capacity. This extension could represent contests between socially connected contestants with intimate knowledge of each other's abilities, e.g., insider contests for CEO succession. This extension could also lead to a fully rationalized contest success function that maps the vector of capacities into probability of winning.[28] Finally, we could explicitly model the preferences of the contest designer and examine how these preferences affect contest design parameters. This extension could address issues such as the dynamic consistency of the designer's ex ante preference for contest-based selection with the designer's ex post preference conditional on observed contestant actions.

[27]Entropy has been adopted as a measure of uncertainty that serves as a basis for a cost function in the information economics literature. See, for example, Gentzkow and Kamenica (2014).

[28]Contest success functions (CSFs), especially the Tullock CSF, have been widely used in the contest literature. Although axiomatic derivations of CSFs, based on welfare and symmetry conditions, have been formulated (Skaperdas, 1996), less attention has been focused on how CSFs might be rationalized as equilibrium outcomes of strategic interaction.

# Appendix: Proofs of results

We establish the following technical lemma, which we will refer to in later proofs.

**Lemma 8.** *Suppose $dP$ is a continuous finite measure over $[0, c)$, where $c > 0$ and $c$ can be infinite. Let $P$ be the nondecreasing function associated with $dP$. Let $\bar{P}$ be a nondecreasing absolutely continuous function defined on $[0, c)$ and let $d\bar{P}$ be the associated measure. Suppose that*

i. *for all $x \in [0, c)$, $P(x) \leq \bar{P}(x)$,*
ii. *for some $x' \in [0, c)$, $P(x') < \bar{P}(x')$,*
iii. *for some $x'' \in [0, c)$, $P(x'') = \bar{P}(x'')$,*
iv. *and $dP\{x \in [0, c) : P(x) < \bar{P}(x)\} = 0$.*

*Then $P$ is absolutely continuous and there exists $\hat{x} \in [0, c)$ such that*

$$P(x) = \begin{cases} \bar{P}(x) & \text{if } x \in [0, \hat{x}] \\ \bar{P}(\hat{x}) & \text{if } x \in [\hat{x}, c) \end{cases}. \tag{34}$$

*Proof of Lemma 8.* Suppose that, at $x^o \in [0, c)$,

$$P(x^o) < \bar{P}(x^o). \tag{35}$$

Let

$$U^o = \{u \in [x^o, c) : \forall x \in [x^o, u), P(x) < \bar{P}(x)\}.$$

By hypothesis (35) and the continuity of $P$ and $\bar{P}$, $U^o$ is a non-degenerate interval. We claim that $\sup U^o = c$. We show this by way of contradiction. Suppose $\sup U^o = u^o < c$. By the continuity of $P$ and $\bar{P}$, $P(u^o) \geq \bar{P}(u^o)$, which implies, by condition i, that

$$P(u^o) = \bar{P}(u^o). \tag{36}$$

Since $U^o$ is a non-degenerate interval whose supremum is $u^o$, there must exist a sequence $(u_n)_n$ such that $u_n < u^o$, $u_n \in U^o$, and $u_n \uparrow u^o$. Since $u_n \in U^o$ and $u_n < u^o$, $P(x) < \bar{P}(x)$, when $x \in [x^o, u_n]$ for all $n$. Thus, condition iv implies that, for all $n$,

$$P(u_n) = P(x^o). \tag{37}$$

The continuity of $P$ implies that $\lim_{n \to \infty} P(u_n) = P(u^o)$, which, by (37), further implies

$$P(u^o) = P(x^o). \tag{38}$$

Since $\bar{P}$ is nondecreasing,

$$\bar{P}(x^o) \leq \bar{P}(u^o). \tag{39}$$

Equations (35), (38), and (39) imply that $P(u^o) < \bar{P}(u^o)$, which contradicts equation (36). This contradiction establishes that $\sup U^o = c$. Thus, if $P(x^o) < \bar{P}(x^o)$, then for all $x \in [x^o, c)$, $P(x) < \bar{P}(x)$. This fact and condition iv imply that

$$P(x^o) < \bar{P}(x^o) \Rightarrow \forall x \in [x^o, c), P(x) \text{ is constant.} \tag{40}$$

We can think of $P(x^o) < \bar{P}(x^o)$ as representing $P$ breaking contact with $\bar{P}$ at some $x$ less than $x^o$. Under this interpretation, equation (40) can be thought of as asserting that once $P$ "breaks contact" with $\bar{P}$, which, by condition i, can only happen if $P$ falls below $\bar{P}$, $P$ "never grows" and thus stays below $\bar{P}$. Now, let

$$\hat{x} = \inf\{x \in [0, c) : P(x) < \bar{P}(x)\}. \tag{41}$$

Condition ii ensures $\hat{x} < c$. By (40) and (S-9) and the continuity of $P$, $P(x) = P(\hat{x})$ if $x \in [\hat{x}, c)$. Condition iii implies $P(0) = \bar{P}(0)$ (otherwise $P$ never meets $\bar{P}$). Thus, by (S-9), condition i, and the continuity of $P$ and $\bar{P}$, $P(x) = \bar{P}(x)$ if $x \in [0, \hat{x}]$. Thus, (34) is established. The absolute continuity of $P$ follows from the absolute continuity of $\bar{P}$. □

*Proof of Proposition 1.* If the maximum for problem (3) equals 1/2, then $P$ itself must be a maximizer for problem (3), since playing $P$ against $P$ always yields 1/2. This implies that the necessary conditions for the optimal challenge distribution, expressed by (8), are satisfied with $F = P$, i.e.,

$$P(x) \leq \alpha + \beta x \quad \forall x \geq 0$$
$$dP\{x \geq 0 : P(x) < \alpha + \beta x\} = 0.$$

By Lemma 8, this implies that there exists $\hat{x} < \infty$ such that

$$P(x) = \begin{cases} \alpha + \beta x & \text{if } 0 \leq x \leq \hat{x} \\ \alpha + \beta \hat{x} & \text{if } x > \hat{x} \end{cases}. \tag{42}$$

The continuity of $P$ at $x = 0$ implies that $\alpha = 0$. The facts that $\alpha = 0$ and $P$ is a CDF in the form of (42) imply that $P$ is a uniform distribution. □

*Proof of Lemma 1.* Suppose, contrary to the lemma, that there exists a point $x^o \geq 0$ such that at least one contestant places point mass on $x^o$. Then symmetry implies that all the contestants place point mass on $x^o$. In this case, because of the random resolution of ties, a contestant is always better off transferring mass from $x^o$ to $x^o+\varepsilon$, for $\varepsilon>0$ sufficiently small. Such a transfer's effect on the capacity constraint can be made arbitrarily small by shrinking $\varepsilon$ to zero while, for all positive $\varepsilon$, no matter how small, the transfer generates a gain that is bounded below by a strictly positive number. Thus, no one places point mass in a symmetric equilibrium, which implies that $P$ is continuous and $P(0)=0$. □

*Proof of Proposition 2.* By Lemma 1, $P$ satisfies the continuity property of the "fixed distribution" in Section 2. Thus, the distribution $F$ played against $P$, together with the two nonnegative optimal dual variables, $\alpha$ and $\beta$, must satisfy (8). Since $dF=0$ implies that $dP=0$, $P$ satisfies the conditions in Lemma 8 with $\bar{P}(x)$ replaced by $\alpha+\beta x$. Applying Lemma 8 shows that $P(x)=\alpha+\beta x$ over its support. By Lemma 1, $P(0)=0$. Thus, $\alpha=0$. Hence, a contestant's winning probability, expressed by (7), equals $\beta\mu$. By symmetry, contestants have the same winning probability. Thus, $\beta\mu=m/n$, which implies that $\beta=m/(n\mu)$. Thus, $P(x)=(mx)/(n\mu)$ over its support. Since $P$ is bounded above by 1, the upper bound of its support must equal $(n\mu)/m$. □

*Proof of Proposition 3.* Inserting (12) into (11) gives the result. □

*Proof of Proposition 4.* Follows immediately from Proposition 3 and Definition 1. □

*Proof of Proposition 5.* This proposition is straightforward from the following result:

**Result 1** (Jones (2002), Section 2.7)**.** *The L-scale of $CB(a,b)$ is*

$$\lambda_{2,CB}=\frac{ab}{(a+b)(a+b+1)}. \tag{43}$$

*The L-skewness of $CB(a,b)$ is*

$$\tau_{3,CB}=\frac{a-b}{a+b+2}. \tag{44}$$

i Since, by Proposition 4, $\frac{m}{n\mu}X\sim CB(n-m,m)$, substituting $(n-m)$ for $a$ and $m$ for $b$ in equation (43) and multiplying the result by $(n\mu)/m$ yield (13).

ii Since the L-skewness is scale invariant, substituting $(n-m)$ for $a$ and $m$ for $b$ in equation (44) yields (14).

iii Follows immediately from the partial derivatives.

iv The sign of the L-skewness is determined by the sign of $n-2m$ and the result on right- and left-skewness follows. As shown by Jones (2002), the Complementary Beta distribution is symmetric when the two shape parameters equal, i.e., when $n-m=m$ in our case. Thus, when $n=2m$, the distribution is symmetric. □

*Proof of Proposition 6.* Since capacity is always used up in equilibrium, to show the s-MPS result, it suffices to show the single-crossing property. Let $F_{m:n}$ be the equilibrium CDF when $n$ contestants compete for $m$ prizes.

First, we show that increasing $n$ induces a s-MPS. Suppose $n' > n$. Note that, $F_{m:n'}$ and $F_{m:n}$ are Complementary Beta, so they are smooth and have positive derivatives on the interior of their supports. Thus, the inverse functions, $F^{-1}_{m:n'}$ and $F^{-1}_{m:n}$, are smooth and have positive derivatives over the open interval $(0,1)$. Thus, the single-crossing condition for a s-MPS can be expressed in terms of the quantile functions: there exists $\hat{q} \in (0,1)$ such that $F^{-1}_{m:n'}(q) - F^{-1}_{m:n}(q) \geq (\leq) 0$, when $q \geq (\leq) \hat{q}$. We prove this below.

By Proposition 3,

$$
\begin{aligned}
F^{-1}_{m:n'}(q) - F^{-1}_{m:n}(q) &= \frac{n'\mu}{m} \sum_{i=n'-m}^{n'-1} \binom{n'-1}{i} q^i (1-q)^{(n'-1)-i} \\
&\quad - \frac{n\mu}{m} \sum_{i=n-m}^{n-1} \binom{n-1}{i} q^i (1-q)^{(n-1)-i}. \qquad (45)
\end{aligned}
$$

Differentiate (45) with respect to $q$, apply the result that $(i+1)\binom{n-1}{i+1} = (n-1-i)\binom{n-1}{i}$ to cancel the common terms, and combine the common factors. This yields

$$
\frac{d(F^{-1}_{m:n'}(q) - F^{-1}_{m:n}(q))}{dq} = \frac{\mu q^{n-1-m}(1-q)^{m-1}}{m} K(q), \qquad (46)
$$

where $K(q) = n'(n'-m)\binom{n'-1}{m-1} q^{n'-n} - n(n-m)\binom{n-1}{m-1}$. When $q \in (0,1)$, the sign of (46) is determined by the sign of $K$. Note that $K < 0$ when $q = 0$, $K > 0$ when $q = 1$, and $K$ is continuous and strictly increasing for $q \geq 0$. Thus, there exists $q^* \in (0,1)$ such that $K$ single crosses the horizontal axis from below at $q = q^*$. This implies, by (46), that $F^{-1}_{m:n'} - F^{-1}_{m:n}$ is strictly decreasing for $q \in (0, q^*)$ and strictly increasing for $q \in (q^*, 1)$. Since $F^{-1}_{m:n'}(0) = F^{-1}_{m:n}(0) = 0$, it follows that $F^{-1}_{m:n'}(q) - F^{-1}_{m:n}(q) < 0$ for $q \in (0, q^*]$. This result, together with the facts that $F^{-1}_{m:n'}(1) - F^{-1}_{m:n}(1) = (n'-n)\mu/m > 0$ and $F^{-1}_{m:n'} - F^{-1}_{m:n}$ is continuous and strictly increasing for $q \in (q^*, 1)$, implies a single crossing. Thus, $F_{m:n'}$ is a s-MPS of $F_{m:n}$.

Next, we show, by a similar argument, that decreasing $m$ induces a s-MPS. Suppose $m' < m$. By Proposition 3,

$$
\begin{aligned}
F^{-1}_{m':n}(q) - F^{-1}_{m:n}(q) &= \frac{n\mu}{m'} \sum_{i=n-m'}^{n-1} \binom{n-1}{i} q^i (1-q)^{(n-1)-i} \\
&\quad - \frac{n\mu}{m} \sum_{i=n-m}^{n-1} \binom{n-1}{i} q^i (1-q)^{(n-1)-i}. \qquad (47)
\end{aligned}
$$

Differentiate (47) with respect to $q$, apply the result that $(i+1)\binom{n-1}{i+1} = (n-1-i)\binom{n-1}{i}$ to cancel the common terms, and combine the common factors. This yields

$$\frac{d(F_{m':n}^{-1}(q) - F_{m:n}^{-1}(q))}{dq} = n\mu q^{n-1-m}(1-q)^{m'-1} J(q), \tag{48}$$

where $J(q) = \frac{n-m'}{m'}\binom{n-1}{n-m'}q^{m-m'} - \frac{n-m}{m}\binom{n-1}{n-m}(1-q)^{m-m'}$. When $q \in (0,1)$, the sign of (48) is determined by the sign of $J$. Note that $J < 0$ when $q = 0$, $J > 0$ when $q = 1$, and $J$ is continuous and strictly increasing for $q \in [0,1]$. Thus, there exists $q^o \in (0,1)$ such that $J$ single crosses the horizontal axis from below at $q = q^o$. This implies, by (48), that $F_{m':n}^{-1} - F_{m:n}^{-1}$ is strictly decreasing for $q \in (0,q^o)$ and strictly increasing for $q \in (q^o,1)$. Since $F_{m':n}^{-1}(0) = F_{m:n}^{-1}(0) = 0$, it follows that $F_{m':n}^{-1}(q) - F_{m:n}^{-1}(q) < 0$ for $q \in (0,q^o]$. This result, together with the facts that $F_{m':n}^{-1}(1) - F_{m:n}^{-1}(1) = n\mu(1/m' - 1/m) > 0$ and $F_{m':n}^{-1} - F_{m:n}^{-1}$ is continuous and strictly increasing for $q \in (q^o,1)$, implies a single crossing. Thus, $F_{m':n}$ is a s-MPS of $F_{m:n}$. □

*Proof of Proposition 7.* i and ii. Follows immediately from Propositions 3 and 5.

iii. Multiply contest size by $\rho > 1$ such that $\rho n$ and $\rho m$ are both integers. Let $F_{m:n}$ and $F_{\rho m:\rho n}$ be the equilibrium distributions before and after an increase in contest size, respectively. Similar to the proof of Proposition 6, we prove the s-MPS result by showing that $F_{\rho m:\rho n}$ and $F_{m:n}$ satisfy a single-crossing property. Consider the horizontal difference between the two distributions, expressed as

$$\begin{aligned} F_{\rho m:\rho n}^{-1}(q) - F_{m:n}^{-1}(q) &= \frac{n\mu}{m}\sum_{i=\rho n-\rho m}^{\rho n-1}\binom{\rho n-1}{i}q^i(1-q)^{(\rho n-1)-i} \\ &\quad - \frac{n\mu}{m}\sum_{i=n-m}^{n-1}\binom{n-1}{i}q^i(1-q)^{(n-1)-i}. \end{aligned} \tag{49}$$

Differentiate equation (49) with respect to $q$, apply the results that $(i+1)\binom{\rho n-1}{i+1} = (\rho n-1-i)\binom{\rho n-1}{i}$ and $(i+1)\binom{n-1}{i+1} = (n-1-i)\binom{n-1}{i}$ to cancel the common terms, and combine the common factors. This yields

$$\frac{d(F_{\rho m:\rho n}^{-1}(q) - F_{m:n}^{-1}(q))}{dq} = \frac{n(n-m)\mu q^{n-m-1}(1-q)^{m-1}}{m}H(q), \tag{50}$$

where $H(q) = \binom{\rho n-1}{\rho m-1}\rho q^{(\rho-1)(n-m)}(1-q)^{(\rho-1)m} - \binom{n-1}{m-1}$. When $q \in (0,1)$, the sign of (50) is determined by the sign of $H$. Differentiating $H$ with respect to $q$ shows that $H$ is maximized at $q = (n-m)/n$, strictly increasing for $q \in (0, (n-m)/n)$, and strictly decreasing for $q \in ((n-m)/n, 1)$. Thus, $H$ has an inverse-U shape on its domain.

This implies, since $H(0) < 0$ and $H(1) < 0$, that, if $H$ is strictly positive somewhere on $(0,1)$, $H$ must cross the horizontal axis twice on its domain. We claim that $H$ is strictly positive somewhere on its domain, which we prove later. Thus, there must exist two distinct points $q_1$ and $q_2$, where $0 < q_1 < (n-m)/n < q_2 < 1$, such that $H(q) < 0$ for $q \in (0,q_1)\bigcup(q_2,1)$ and $H(q) > 0$ for $q \in (q_1,q_2)$. This implies that $F^{-1}_{\rho m:\rho n} - F^{-1}_{m:n}$ is strictly decreasing for $q \in (0,q_1)$, strictly increasing for $q \in (q_1,q_2)$, and strictly decreasing again for $q \in (q_2,1)$. Since $F^{-1}_{\rho m:\rho n} - F^{-1}_{m:n}$ equals 0 for $q = 0$ and for $q = 1$, it follows that $F^{-1}_{\rho m:\rho n} - F^{-1}_{m:n}$ is strictly negative for $q \in (0,q_1]$, strictly positive for $q \in [q_2,1)$, and equals zero at a unique $q^* \in (q_1,q_2)$, implying a single crossing.

What remains to show is the claim we made that $H$ must be strictly positive somewhere on its domain. Suppose not; so $H$ is weakly negative everywhere. By (50), this implies, since it is obvious that $H$ cannot equal 0 everywhere, the horizontal difference, $F^{-1}_{\rho m:\rho n} - F^{-1}_{m:n}$, when $q = 1$ must be strictly less than that when $q = 0$, which contradicts that the horizontal difference is 0 for both $q = 0$ and $q = 1$. This contradiction establishes our claim and completes the proof. □

*Proof of Proposition 8.* Define $\Phi(w,n,m) = \sum_{i=n-m}^{n-1} \binom{n-1}{i} w^i (1-w)^{n-1-i}$, for all $w \in [0,1]$. Note that $\Phi(w,n,m)$ is the distribution of the $n-m : n-1$ order statistic for a uniform distribution. Let $\big((n_j,m_j)\big)$ be a sequence of natural numbers such that for all $j$, $m_j = \rho n_j$, where $0 < \rho < 1$. Then $\Phi(w,n_j,m_j)$ is the distribution of the $(1-\rho)n_j : n_j - 1$ order statistic for a uniform distribution. The weak convergence of order statistics to their associated quantiles (see Lemma 1.2.9 in Reiss (1980)) implies that the sequence of distribution functions $\big(\Phi(\cdot,n_j,m_j)\big)$ converges weakly to the distribution function of a degenerate random variable concentrated at $1-\rho$. Thus,

$$\lim_{j\to\infty} \Phi(w,n_j,m_j) = \begin{cases} 0 & \text{if } w < 1-\rho \\ 1 & \text{if } w > 1-\rho \end{cases}, \quad \forall\, w \in [0,1]. \tag{51}$$

Let $F^j_Z$ be a sequence of equilibrium distributions for the normalized performance levels, $\tilde{Z} = \tilde{X}/(\mu/\rho)$. Using the normalized performance levels, we can express the equilibrium condition given by Proposition 3 as

$$z = \Phi(F^j_Z(z),n_j,m_j). \tag{52}$$

Suppose the proposition is incorrect, then equation (15) is not satisfied. Thus, some sequence $\left(F_Z^j\right)$ does not converge to the distribution

$$F_Z^\infty(z) = \begin{cases} 1-\rho & \text{if } z < 1 \\ 1 & \text{if } z \geq 1 \end{cases}.$$

This implies, passing to a subsequence if necessary, using Helly's selection theorem (See Billingsley (1985) Theorem 25.9), that there exists a limit distribution for the sequence, $F_Z'$ unequal to $F_Z^\infty$. This implies that, at some continuity point of $F_Z'$, $z' \in (0,1)$,

$$F_Z^j(z') \to F_Z'(z') \neq 1-\rho.$$

Exploiting the continuity of $\Phi$, the fact that $z'$ is a continuity point of $F_Z'$, (52), and (51), we can take the limit of (52). This yields

$$z' = \begin{cases} 0 & \text{if } F_Z'(z') < 1-\rho \\ 1 & \text{if } F_Z'(z') > 1-\rho \end{cases}. \tag{53}$$

Since, by hypothesis, the left hand side of (53), $z'$, lies in the interval (0, 1) and the function represented by the right hand side of (53) only takes on the values 0 and 1, (53) cannot be satisfied. This contradiction establishes the result. □

*Proof of Lemma 2.* Suppose there exists a point $x^o \geq 0$ such that at least one type-$t$ contestant places point mass on it. Then symmetry implies that all type-$t$ contestants place point mass on $x^o$. In this case, by the same argument as in the proof of Lemma 1, a type-$t$ contestant is always better off transferring mass from $x^o$ to $x^o + \varepsilon$, for $\varepsilon > 0$ sufficiently small. The contradiction implies that no type places point mass. Thus, $P$ must be continuous and, hence, $P(0) = 0$. □

*Proof of Proposition 9.* Note that $\psi$, defined by (18), is an increasing concave function defined over the nonnegative real line. Since $P$ is bounded above by the two support lines, $P(x) \leq \psi(x)$ for all $x \geq 0$. Since the support of $P$ is contained by the support of the best replies of types $W$ and $S$, the support of $P$ is contained within $\{x \geq 0 : P(x) = \psi(x)\}$. Thus, the conditions of Lemma 8 are satisfied and the result follows. □

*Proof of Corollary 1.* See the discussion right before Corollary 1 in the main text. □

*Proof of Corollary 2.* Follows immediately from Proposition 9 and the fact that the support of $P$ equals the union of the supports of two types' equilibrium distributions. □

*Proof of Lemma 3.* See the discussion right before Lemma 3 in the main text. □

*Proof of Lemma 4.* First, we show that $\alpha_S \geq \alpha_W$ by way of contradiction. Suppose that $\alpha_S < \alpha_W$. Then we must have $\beta_S > \beta_W$, since otherwise, by Lemma 3, $\text{Supp}_W$ would be an empty set. However, if $\alpha_S < \alpha_W$ and $\beta_S > \beta_W$, Lemma 3 implies that all the points in $\text{Supp}_S$ are weakly smaller than all the points in $\text{Supp}_W$, which contradicts that $\mu_S > \mu_W$. Thus, we must have $\alpha_S \geq \alpha_W$ and, by Proposition 9 and the fact that $P(0) = 0$, we must have $\alpha_W = 0$. Thus, $\alpha_S \geq \alpha_W = 0$.

The above analysis implies that there are only two cases: either (i) $\alpha_S = \alpha_W = 0$ or (ii) $\alpha_S > \alpha_W = 0$. In case (i), we must have $\beta_S = \beta_W$, since otherwise, by Lemma 3, $\text{Supp}_S$ and $\text{Supp}_W$ cannot both contain strictly positive values of $x$. In case (ii), we must have $\beta_S < \beta_W$, since otherwise, by Lemma 3, $\text{Supp}_S$ would be an empty set. Finally, note that no strong contestant has sufficient capacity to guarantee winning and, hence, $\beta_S$, which measures the marginal value of $S$'s capacity, must be strictly positive. □

*Proof of Lemma 5.* For the concession configuration to sustain a symmetric equilibrium, $W$'s payoff, $p_W^C$, must be at least the payoff from playing the alternative strategy of placing all the mass on 0 with probability $1 - \mu_W/\mu_S$ and mimicking $S$'s strategy with probability $\mu_W/\mu_S$. By playing this alternative strategy, $W$'s payoff is $p_S^C \mu_W/\mu_S$. Thus, the condition can be expressed as $p_W^C \geq p_S^C \mu_W/\mu_S$, which, given that $\mu_W/\mu_S = p_W^G/p_S^G$ and $\theta p_S^C + (1-\theta) p_W^C = \theta p_S^G + (1-\theta) p_W^G = m/n$, is equivalent to $p_W^C \geq p_W^G$.

In the challenge configuration, $W$'s payoff, $p_W^G$, includes the probability of outperforming $S$. Thus, for this configuration to sustain a symmetric equilibrium, $p_W^G$ must be strictly greater than $p_W^C$. The latter is $W$'s payoff when $W$ concedes to $S$.[29] □

*Proof of Proposition 10.* With probability $\binom{n}{i}\theta^i(1-\theta)^{n-1-i}$, exactly $i$ out of $n$ contestants are strong, in which case selection efficiency is maximized when the number of selected strong contestants equals $\min[i, m]$. Summing up the expected numbers from $i = 0$ to $i = n$ and dividing the sum by $m$ yield equation (28). The rest of the proof is clear from equation (27) and the definition of $\Delta\Pi$. □

*Proof of Corollary 3.* i. The result follows from (29), (30), and the facts that $\Pi^*$ is constant in $r$ whereas $\frac{\theta r}{(1-\theta)+\theta r}$ is strictly increasing in $r$ for all $r > 1$.

ii. Since $\sum_{i=0}^{n}\binom{n}{i}\theta^i(1-\theta)^{n-i} = 1$, $\Pi^*$, expressed in (28), can be rewritten as

$$\Pi^* = 1 - \sum_{i=0}^{m-1} \frac{m-i}{m}\binom{n}{i}\theta^i(1-\theta)^{n-i}. \tag{54}$$

[29] Given the condition for a concession equilibrium presented by Lemma 5, the construction of a concession equilibrium is obvious. Given the condition for a challenge equilibrium presented by Lemma 5, a challenge equilibrium can be constructed by following the procedure illustrated by the example in Section 4.2.

Differentiate (54) with respect to $\theta$, apply the result that $i\binom{n}{i} = (n-i+1)\binom{n}{i-1}$ to cancel the common terms, and combine the common factors. This yields

$$\frac{\partial \Pi^*}{\partial \theta} = \sum_{i=0}^{m-1} \frac{n-i}{m} \binom{n}{i} \theta^i (1-\theta)^{n-i-1},$$

which is positive and goes to 0 when $\theta \to 1$. Note that

$$\frac{\partial \left(\frac{\theta r}{(1-\theta)+\theta r}\right)}{\partial \theta} = \frac{r}{((1-\theta)+\theta r)^2},$$

which is positive and goes to $1/r$ when $\theta \to 1$. Thus, both $\Pi^*$ and $\frac{\theta r}{(1-\theta)+\theta r}$ are strictly increasing in $\theta$ and the rate of increase is smaller for $\Pi^*$ than for $\frac{\theta r}{(1-\theta)+\theta r}$ when $\theta \to 1$. These results imply two facts: (1) $\Pi$, given by (29), must be strictly increasing in $\theta$, and (2) since, when $\theta = 1$, $\Pi^* = \frac{\theta r}{(1-\theta)+\theta r} = 1$, we must have, for $\theta$ sufficiently close to 1, $\Pi^* > \frac{\theta r}{(1-\theta)+\theta r}$ and thus $\Delta\Pi > 0$, which implies, since $\Delta\Pi$ is continuous in $\theta$ and $\Delta\Pi = 0$ for both $\theta = 0$ and $\theta = 1$, that $\Delta\Pi$ must be nonmonotonic in $\theta$.

iii. For any realization of contestant types, adding an additional contestant will never reduce but sometimes increase the number of strong contestants. Thus, increasing the number of contestants will strictly increase the expected number of strong contestants, which implies, since $\Pi^*$ requires giving strong contestants absolute priority, that $\Pi^*$ must be strictly increasing in $n$. The result then follows from (29), (30), and the fact that $\frac{\theta r}{(1-\theta)+\theta r}$ is constant in $n$.

iv. It is clear from (28) that $\Pi^*$ is strictly decreasing in $m$. The result then follows from (29), (30), and the fact that $\frac{\theta r}{(1-\theta)+\theta r}$ is constant in $m$.

v. First, we show that increasing contest size strictly increases $\Pi^*$. Let $k$ be the scale factor, where $k > 1$ is an integer. After a $k$-fold scaling of the game, the contest involves $k$ bunches of contestants, with each bunch containing $n$ contestants. Let $S_i^k$ be the number of strong contestants in the $i$th bunch, for all $i \in \{1, \cdots, k\}$. Let

$$\bar{S}^k = \frac{1}{k} \sum_{i=1}^{k} S_i^k$$

be the average number of strong contestants *per bunch* produced by a $k$-fold scaling of the game. Note that $\Pi^*$ is strictly increasing in $k$ if and only if $k \mapsto \mathbb{E}\left[\min[\bar{S}^k, m]\right]$ is strictly increasing.

By the concavity of $x \mapsto \min[x, m]$ and the conditional form of Jensen's inequality, the following inequality holds with probability 1:

$$\mathbb{E}\left[\min[\bar{S}^{k-1}, m]\middle|\bar{S}^k\right] \leq \min\left[\mathbb{E}[\bar{S}^{k-1}\middle|\bar{S}^k], m\right].$$

Since $\mathbb{E}[\bar{S}^{k-1}\big|\bar{S}^k] = \bar{S}^k$, we know that, with probability 1,

$$\min[\bar{S}^k, m] \geq \mathbb{E}\left[\min[\bar{S}^{k-1}, m]\middle|\bar{S}^k\right]. \tag{55}$$

Next, note that

$$\mathbb{P}\left\{\min[\bar{S}^k, m] > \mathbb{E}\left[\min[\bar{S}^{k-1}, m]\middle|\bar{S}^k\right]\right\} > 0. \tag{56}$$

To see this, consider the event where

$$S_1^k = m+1,\, S_2^k = m-1, \text{ and } S_i^k = m \quad \forall i \neq 1, 2.$$

Since $1 \leq m < n$ and since contestant types are independent, this event has positive probability. When this event occurs,

$$\bar{S}^k = m \text{ and } \mathbb{P}\left\{\bar{S}^{k-1} < m\middle|\bar{S}^k\right\} > 0.$$

Thus, when this event occurs,

$$\min[\bar{S}^k, m] = m > \mathbb{E}\left[\min[\bar{S}^{k-1}, m]\middle|\bar{S}^k\right]. \tag{57}$$

Thus, (57) implies (56). By taking the unconditional expectation of (55) and noting (56), we have

$$\mathbb{E}\left[\min[\bar{S}^k, m]\right] > \mathbb{E}\left[\min[\bar{S}^{k-1}, m]\right].$$

Thus, $\Pi^*$ is strictly increasing in $k$.

It is clear by equations (29) and (30) that both $\Pi$ and $\Delta\Pi$ are weakly increasing in $\Pi^*$ and constant in $\frac{\theta r}{(1-\theta)+\theta r}$. Thus, by equations (29) and (30) and the result that $\Pi^*$ is increasing in $k$, both $\Pi$ and $\Delta\Pi$ must be weakly increasing in $k$. □

*Proof of Lemma 6.* Applying the argument in the proof of Lemma 2 to all $x \in [0, \bar{x})$ establishes the result. □

*Proof of Proposition 11.* See the discussion right before Proposition 11 in the main text. □

*Proof of Proposition 12.* First, we show that the optimal dual variables, $\alpha_t$ and $\beta_t$, for

$t \in \{S, W\}$, are constant in $\bar{x} \in [\mu_S, \infty)$.[30]

Note that $W$'s payoff, expressed by $\alpha_W + \beta_W \mu_W$, is, by Proposition 11, constant in $\bar{x} \in [\mu_S, \infty)$. Since, by Lemma 4, $\alpha_W = 0$, $\alpha_W$ and $\beta_W$ must both be contestant in $\bar{x} \in [\mu_S, \infty)$. Now consider $\alpha_S$ and $\beta_S$. Since, as argued in the proof of Proposition 11, a scoring cap $\bar{x} \in [\mu_S, \infty)$ does not affect the equilibrium configuration, to show that $\alpha_S$ and $\beta_S$ are constant in $\bar{x} \in [\mu_S, \infty)$, it suffices to show that the configuration-conditioned multipliers, $\alpha_S$ and $\beta_S$, are constant in $\bar{x} \in [\mu_S, \infty)$. Note that, in the challenge configuration, $\alpha_S = \alpha_W$ and $\beta_S = \beta_W$. In this case, since $\alpha_W$ and $\beta_W$ are constant in $\bar{x} \in [\mu_S, \infty)$, so are $\alpha_S$ and $\beta_S$. In the concession configuration, $\alpha_S$ and $\beta_S$ are determined by the following two equations: (1) $\alpha_S + \beta_S \mu_S = p_S^C$, where $p_S^C$ is $S$'s payoff if $S$ has absolute priority, and (2) $\alpha_S + \beta_S \check{x} = \check{p}$, where $\check{x}$ is the intersection of the two support lines in the concession configuration, determined by $\alpha_S + \beta_S \check{x} = \alpha_W + \beta_W \check{x}$, and $\check{p}$ is a contestant's probability of winning if he concedes to strong contestants but outperforms all weak ones. Since $\bar{x} \in [\mu_S, \infty)$ does not affect $p_S^C$, $\check{p}$, $\alpha_W$, or $\beta_W$, it does not affect either equation. Thus, $\alpha_S$ and $\beta_S$ must both be constant in $\bar{x} \in [\mu_S, \infty)$ in the concession configuration.

The above analysis implies that the concave lower envelope, $\psi$, defined by (18), is constant in $\bar{x} \in [\mu_S, \infty)$. Thus, when the scoring cap is not binding, i.e., when $\bar{x} \geq \hat{x}$, where $\hat{x}$ is defined in Proposition 9, equilibrium distributions are unaffected.

When the scoring cap is binding, i.e., when $\bar{x} \in [\mu_S, \hat{x})$, since $P$ is bounded above by $\psi$ and since $\psi$ is strictly increasing, we must have $P(\bar{x}) \leq \psi(\bar{x}) < \psi(\hat{x}) = 1$, which implies point mass on $x = \bar{x}$, and, by the random resolution of ties, $P$ is discontinuous at $x = \bar{x}$. Moreover, since probability weight is placed only on points where $P$ meets $\psi$, we must have $P(\bar{x}) = \psi(\bar{x})$. By Lemma 6, $P$ is continuous on $[0, \bar{x})$. Applying Lemma 8 over the interval $[0, \bar{x})$ and using the result that, given $\bar{x} \in [\mu_S, \hat{x})$, $P(\bar{x}) = \psi(\bar{x})$ yield

$$P(x) = \begin{cases} \psi(x) & \text{if} \quad x \in [0, \tilde{x}] \\ \psi(\tilde{x}) & \text{if} \quad x \in [\tilde{x}, \bar{x}) \\ \psi(\bar{x}) & \text{if} \quad x = \bar{x} \end{cases}, \tag{58}$$

where $\tilde{x} \in (0, \bar{x})$ is a constant and $\bar{x} \in [\mu_S, \hat{x})$. Thus, imposing a scoring cap $\bar{x} \in [\mu_S, \hat{x})$ induces each contestant to transfer the mass over $(\tilde{x}, \hat{x}]$ to the point mass on $\bar{x}$, leaving the mean of the distribution and the distribution over $[0, \tilde{x}]$ unchanged. Such a change in distribution represents a simple mean-preserving contraction (s-MPC).

[30] When $\bar{x} = \mu_S$, the optimal dual variables are not unique. However, $S$'s strategy is unique: $S$ places all the mass on $\mu_S$. Thus, without loss of generality, we redefine the values of $\alpha_S$ and $\beta_S$ when $\bar{x} = \mu_S$ by their limiting values when $\bar{x} \downarrow \mu_S$.

Moreover, given that $\bar{x} \in [\mu_S, \hat{x})$, decreasing $\bar{x}$ decreases $\tilde{x}$. To see this, note that, given $\bar{x} \geq \mu_S$, $\psi$ is fixed and strictly increasing. Thus, decreasing $\bar{x}$, provided that $\bar{x}$ is still weakly greater than $\mu_S$, decreases $\psi(\bar{x})$ and, by (58), decreases $P(\bar{x})$. This implies that, when the scoring cap becomes tighter, each contestant places more point mass on the cap. In this case, $\lim_{x \uparrow \tilde{x}} P(x)$ must decrease. Thus, by (58), $\psi(\tilde{x})$ must decrease. Since, when $\bar{x} \geq \mu_S$, $\psi$ is fixed and strictly increasing, $\tilde{x}$ must decrease. Such a change in $P$ implies a s-MPC of contestant performance.

The proposition follows from the s-MPC result and Diamond and Stiglitz (1974) that a s-MPC increases the utility of a risk-averse expected utility maximizer. □

*Proof of Proposition 13.* See the discussion right before Proposition 13 in the main text. □

*Proof of Proposition 14.* If, without a penalty trigger, $W$ concedes to $S$, then as implied by Proposition 13, after a penalty trigger, $\underline{x} > 0$, is imposed, $W$ will still concede to $S$. Let $P^*$ and $P$ be the probability of winning function with and without a penalty trigger, respectively. For $t \in \{S, W\}$, let $\text{Supp}_t^*$ and $\text{Supp}_t$ be the support of type-$t$'s strategy with and without a penalty trigger, respectively, and let $\bar{x}_t^*$ and $\bar{x}_t$ be the upper bound of the support of type-$t$'s strategy with and without a penalty trigger, respectively.

Note that the following three conditions must hold:

i. $P^*(\bar{x}_W^*) = P(\bar{x}_W)$,

ii. $\bar{x}_W^* < \bar{x}_W$,

iii. and $\bar{x}_S^* > \bar{x}_S$.

The reason for (i) is that, in the concession configuration, both $P(\bar{x}_W)$ and $P^*(\bar{x}_W^*)$ equal the probability of winning if a contestant concedes to strong contestants but outperforms all weak ones. Condition (i) implies (ii), since otherwise, given that $P$ and $P^*$ are straight lines over $\text{Supp}_W$ and $\text{Supp}_W^*$, respectively, and given that $S$ places no weight over $W$'s support in concession equilibria, $W$'s strategy with a penalty trigger would first-order stochastically dominate that without a penalty trigger, a contradiction. Condition (ii) further implies (iii), since otherwise, given that $P$ and $P^*$ are straight lines over $\text{Supp}_S$ and $\text{Supp}_S^*$, respectively, and given that $W$ places no weight over $S$'s support in concession equilibria, $S$'s strategy with a penalty trigger would be first-order stochastically dominated by that without a penalty trigger, a contradiction.

Since, in concession equilibria, the lower bound of $S$'s strategy meets the upper bound of $W$'s, conditions (i), (ii), and (iii) imply that $W$'s strategy with a penalty trigger single crosses that without a penalty trigger from below and that $S$'s strategy with a penalty trigger single crosses that without a penalty trigger from above, and the proposition follows from these single-crossing results. □

*Proof of Proposition 15.* Let $k$ be the number of local contests and $n_i$ and $m_i$ the number of contestants and the number of prizes in the $i$th local contest, respectively, such that $\sum_{i=1}^{k} n_i = n$, $\sum_{i=1}^{k} m_i = m$, and $n_i > m_i \geq 1$ for all $i \in \{1, \cdots, k\}$. Let $\Pi^g$ and $\Pi^l$ be actual selection efficiency of the grand and of the local contest structure, respectively.

First, we show that $\Pi^g \geq \Pi^l$ and thus localizing a contest never improves efficiency. If the grand contest has concession equilibria, its selection is efficient and thus $\Pi^g \geq \Pi^l$. If the grand contest has challenge equilibria, $\Pi^g = \frac{\theta r}{(1-\theta)+\theta r}$. By Proposition 10, the efficiency of every local contest is bounded above by $\frac{\theta r}{(1-\theta)+\theta r}$. Thus, $\Pi^l \leq \frac{\theta r}{(1-\theta)+\theta r} = \Pi^g$. Therefore, $\Pi^g \geq \Pi^l$ regardless of the equilibrium configuration of the grand contest.

Next, we prove that having challenge equilibria in every local contest is sufficient for $\Pi^l = \Pi^g$. If every local contest has challenge equilibria, $\Pi^l = \frac{\theta r}{(1-\theta)+\theta r}$. Thus, by Proposition 10, $\Pi^g \leq \frac{\theta r}{(1-\theta)+\theta r} = \Pi^l$, which implies, given the result we proved above that $\Pi^g \geq \Pi^l$, that $\Pi^g = \Pi^l$.

Finally, we prove that having challenge equilibria in every local contest is necessary for $\Pi^l = \Pi^g$. If there exists at least one local contest with concession equilibria, we must have $\Pi^l < \frac{\theta r}{(1-\theta)+\theta r}$. In this case, if the grand contest has challenge equilibria, we must have $\Pi^g = \frac{\theta r}{(1-\theta)+\theta r} > \Pi^l$; if the grand contest has concession equilibria, $\Pi^g = \Pi^*(n,m)$, in which case $\Pi^g > \Pi^l$ follows from the following result:

$$\Pi^*(n,m) > \sum_{i=1}^{k} \left(\frac{m_i}{m}\right) \Pi^*(n_i,m_i) \geq \sum_{i=1}^{k} \left(\frac{m_i}{m}\right) \min\left[\frac{\theta r}{(1-\theta)+\theta r}, \Pi^*(n_i,m_i)\right] = \Pi^l. \tag{59}$$

What remains to show is (59). In (59), the last equality comes from Proposition 10, and the second inequality is obvious. To prove the first inequality, it suffices to show that a local contest mechanism is not efficient even if all the local contests have concession equilibria. This result holds for the following reason. Since $m < n$, under a local contest structure, there must exist a local contest, say the $j$th, in which $m_j < n_j$. Since every local contest has at least one prize, we must have $m_j < m$. With positive probability, exactly $m_j + 1$ out of $n$ contestants are strong. In this case, since $m_j + 1 \leq m$, an efficient selection requires selecting all the $m_j + 1$ strong contestants. However, under a local contest structure, with positive probability, the $m_j + 1$ strong contestants are all assigned to the $j$th local contest, in which case it is impossible to select all of them. This establishes the first inequality in (59) and completes the proof of necessity. □

*Proof of Proposition 16.* Let $P$ be a contestant's probability of winning function. A contestant's problem is to choose a distribution, $F$, with support included in $[0, T^*]$, subject to the capacity constraint, to maximize $\int_0^{T^*} k e^{rx} P(x) dF(x)$, where $k = e^{-rT^*} V_c$ is a constant.

By the similar argument developed in Section 2.1, it is clear that the optimal measure, $dF$, and the associated dual variables, $\alpha$ and $\beta$, must satisfy the following:

$$\begin{aligned} k e^{rx} P(x) &\leq \alpha + \beta x \quad \forall x \in [0, T^*], \\ dF\{x \in [0, T^*] : k e^{rx} P(x) &< \alpha + \beta x\} = 0. \end{aligned} \tag{60}$$

Following the proof of Lemma 1, it is evident that, in equilibrium, no one places point mass on $[0, T^*)$. Thus, when there are just two contestants, in a symmetric equilibrium,

$$P(x) = F(x) \quad \forall x \in [0, T^*). \tag{61}$$

To derive the equilibrium $F$, we first establish the following lemma.

**Lemma 9.** *If Supp$\{F\}$ represents the support of the equilibrium distribution, then Supp$\{F\}\backslash\{T^*\}$ is an interval containing 0, and F is absolutely continuous on* $[0, T^*)$.

*Proof.* The function $\bar{P} : [0, T^*) \mapsto [0, \infty)$, defined by $\bar{P} = \frac{1}{k} e^{-rx} (\alpha + \beta x)$ is absolutely continuous. By equations (60) and (61), with respect to $\bar{P}$, $F$ satisfies conditions (i) and (iv) of Lemma 8 on $[0, T^*)$. Moreover, since no one places point mass on $[0, T^*)$, $F$ must be continuous on $[0, T^*)$ and thus the result follows from a straightforward modification of the proof of Lemma 8. □

Next, we show that $\alpha = 0$ and $\beta > 0$.

**Lemma 10.** *The optimal dual variables satisfy:* $\alpha = 0$ *and* $\beta > 0$.

*Proof.* Since $\beta$ measures the marginal value of capacity and since no contestant has sufficient capacity to guarantee winning, $\beta$ must be strictly positive. By Lemma 9, $F(0) = \alpha$. Since no one places point mass on 0, $F(0) = 0$. Thus, $\alpha = 0$. □

Lemmas 9 and 10 imply that there are only two candidate equilibrium configurations:

A. Supp$\{F\} = [0, \hat{x}]$, where $\hat{x} \leq T^*$, in which case $F$ has no point mass and, hence, $P = F$ everywhere;

B. Supp$\{F\} = [0, x'] \bigcup \{T^*\}$, where $x' < T^*$, in which case, $F$ has point mass on $T^*$ but not elsewhere and, hence, $P(x) = F(x)$ for all $x < T^*$ and, by the random resolution of ties, $P(T^*) = (1 + F(x'))/2$.

We first construct a symmetric equilibrium according to configuration (A). In (A), the payoff function must meet the support line everywhere over $[0, \hat{x}]$. Thus, $ke^{rx}P(x) =$

$\beta x$ for all $x \in [0, \hat{x}]$, which implies, since $P = F$ everywhere in (A), that $ke^{rx}F(x) = \beta x$ for all $x \in [0, \hat{x}]$. Since $F(\hat{x}) = 1$ in (A), it must be that

$$\beta = k\,e^{r\hat{x}}/\,\hat{x}. \tag{62}$$

Thus, $F$ must equal (32), which sustains an equilibrium if and only if the following conditions are satisfied:

(A.i) $F$ is nondecreasing on the nonnegative real line;

(A.ii) the mean of $F$ equals $\mu$;

(A.iii) $T^* \geq \hat{x}$, where $\hat{x}$ is the upper bound of the support of $F$;

(A.iv) the payoff function, $k\,e^{rx}P(x)$, is bounded above by the support line everywhere on $[0, T^*]$.

Differentiating (32) with respect to $x$ shows that $F'(x) \geq 0$ if and only if either $0 \leq x \leq \min[\hat{x}, 1/r]$ or $x \geq \hat{x}$. Thus, (A.i) holds if and only if $\hat{x} \leq 1/r$.

Note that the mean of $F$ is expressed by the left hand side of (31), yielded from (32) by integrating $x$ over $F$. Thus, (A.ii) holds if and only if $\hat{x}$ satisfies (31).

Since the payoff function meets the support line on $[0,\hat{x}]$, (A.iv) is satisfied if and only if the payoff function is bounded above by the support line on $(\hat{x}, T^*]$. Note that, in configuration (A), $P(x) = 1$ for all $x \geq \hat{x}$, so the payoff function when $x \geq \hat{x}$ is $ke^{rx}$. Since this function is convex and increasing, it is bounded above by the support line on $(\hat{x}, T^*]$ if and only if it is bounded above by the support line at $T^*$, i.e., if and only if

$$k\,e^{rT^*} \leq \alpha + \beta\,T^*, \tag{63}$$

where $\beta$ is given by (62) and, by Lemma 10, $\alpha = 0$. Thus, (63) is equivalent to

$$\frac{e^{rT^*}}{T^*} \leq \frac{e^{r\hat{x}}}{\hat{x}}. \tag{64}$$

Note that, on the positive real line, $e^{rx}/x$ is a continuous function that is strictly decreasing on $(0, 1/r)$ and strictly increasing on $(1/r, \infty)$. Thus, $\hat{x} \leq T^*$ is implied by $\hat{x} \leq 1/r$ and (64). This means that (A.iii) is implied by (A.i) and (A.iv). Rearranging (64) yields that $\hat{x} \leq T^*\,e^{r(\hat{x}-T^*)}$.

The above analysis shows that configuration (A) sustains an equilibrium if and only if $\hat{x}$ satisfies (31) and $\hat{x} \leq \min\left[T^*\,e^{r(\hat{x}-T^*)},\, 1/r\right]$. If any of these conditions is violated, a symmetric equilibrium must have configuration (B), which we examine below.

In (B), the payoff function must meet the support line at $x = T^*$ and everywhere over $[0, x']$. Thus, $ke^{rx}P(x) = \beta x$ for all $x \in [0,x']$ and for $x = T^*$. This implies, given $P(T^*) = (1+F(x'))/2$ in (B), that

$$\beta = \frac{ke^{rT^*}(1+F(x'))}{2T^*},$$

and further implies, given $P(x) = F(x)$ for all $x \in [0,T^*)$ in (B), that

$$F(x) = \frac{(1+F(x'))x}{2T^*}e^{r(T^*-x)} \quad \forall\, x \in [0,x']. \tag{65}$$

Evaluating (65) at $x = x'$ establishes the relation between $F(x')$ and $x'$, expressed as

$$F(x') = \frac{x'}{2T^* e^{r(x'-T^*)} - x'}. \tag{66}$$

Substituting (66) into (65) and applying the results that, in configuration (B), $F(x) = F(x')$ for all $x \in [x', T^*)$ and $F(T^*) = 1$ yield the expression of $F$ given by (33), with $x'$ chosen to make the capacity constraint bind. □

*Proof of Lemma 7.* We first present a result in Ankirchner, Hobson, and Strack (2014). Then, based on this result, we give two corollaries that we will use to prove this lemma and the next proposition.

**Result 2** (Ankirchner, Hobson, and Strack (2014), Theorem 3, Proposition 4, and Proposition 5)**.** *Consider the time-homogeneous local martingale diffusion $X$, where $X$ solves*

$$dX_t = \eta(X_t)dW_t, \quad \text{with} \quad X_0 = \mu; \tag{67}$$

*here $W_t$ is a Wiener process, $\mu \in \mathbb{R}^+$, and $\eta : \mathbb{R} \mapsto \mathbb{R}^+$ is Borel-measurable. Define*

$$q(x) = \int_\mu^x dy \int_\mu^y \frac{2}{\eta^2(z)}dz. \tag{68}$$

*Suppose the target distribution $F$ satisfies $\int x dF(x) = \mu$. There exists an integrable stopping time that induces $F$ by stopping $X$ if and only if $q$ is integrable with respect to $F$. In this case, any minimal and integrable stopping time $\tau$ satisfies*

$$E[\tau] = \int q(x)dF(x).$$

*Moreover, if $F$ is absolutely continuous with a compact support and both its density and $\eta$ are bounded away from zero on its support, then there exists a bounded time* *embedding of $F$ into $X$.*

The following two corollaries are straightforward from Result 2.

**Corollary 4.** *Suppose $X$ is a Brownian motion absorbed at zero, i.e., $X$ is given by* (67) *with $\eta(X_t) = \delta > 0$ if $X_t > 0$ and $\eta(X_t) = 0$ if $X_t = 0$. Suppose $F$ satisfies $\int x dF(x) = \mu$. There exists an integrable stopping time that induces $F$ by stopping $X$ if and only if $F$ has its support on the nonnegative real line and $x^2$ is integrable with respect to $F$. In this case, any minimal and integrable stopping time $\tau$ satisfies*

$$E[\tau] = \int \frac{(x-\mu)^2}{\delta^2} dF(x).$$

*Proof.* Given that $\eta(X_t) = \delta$ if $X_t > 0$ and $\eta(X_t) = 0$ if $X_t = 0$, function $q$, expressed by (68), satisfies $q(x) = \frac{(x-\mu)^2}{\delta^2}$ if $x \geq 0$ and $q(x) = +\infty$ if $x < 0$. The corollary then follows immediately from Result 2. □

**Corollary 5.** *Suppose $X$ is a geometric Brownian motion, i.e., $X$ is given by* (67) *with $\eta(X_t) = \upsilon X_t$, where $\upsilon > 0$. Suppose $F$ satisfies $\int x dF(x) = \mu$. There exists an integrable stopping time that induces $F$ by stopping $X$ if and only if $F$ has its support on the nonnegative real line and* $\ln(x)$ *is integrable with respect to $F$. In this case, any minimal and integrable stopping time $\tau$ satisfies*

$$E[\tau] = -\int \frac{2}{\upsilon^2}(\ln(x) - \ln\mu) dF(x).$$

*Proof.* Given that $\eta(X_t) = \upsilon X_t$, function $q$, expressed by (68), satisfies $q(x) = -\frac{2}{\upsilon^2}(\ln(x) - \ln\mu - \frac{x}{\mu} + 1)$. The corollary then follows immediately from Result 2. □

By Corollaries 4 and 5, Lemma 7 follows immediately from the convexity of $x \mapsto \frac{1}{\delta^2}(x-\mu)^2$ and of $x \mapsto -\frac{2}{\upsilon^2}\ln(x)$ and the fact that a mean-preserving spread increases the expected value of a convex function (this is analogous to the well-known result that a mean-preserving spread increases the expected utility of an individual with a convex utility function). □

*Proof of Proposition 17.* Given Propositions 6 and 7 and Lemma 7, to show Proposition 17, it suffices to show the existence of an integrable stopping time. By Corollaries 4 and 5, this is equivalent to showing that both $x^2$ and $\ln(x)$ are integrable with respect to the equilibrium distribution presented in Proposition 3. Since this distribution is Complementary Beta, which has a finite variance, integrability condition holds for $x^2$.

Now we show that integrability condition also holds for $\ln(x)$. Integration by parts yields

$$\int \ln(x)dF(x) = \ln(x)F(x)|_0^{\frac{n\mu}{m}} - \int_0^{\frac{n\mu}{m}} \frac{1}{x}F(x)dx.$$

To show that the left hand side of the above equation is finite, it suffices to show that, on the right hand side,

i. $\lim_{x\to 0}|\ln(x)F(x)|$ is finite,

ii. and $\int_0^{\frac{n\mu}{m}} |\frac{1}{x}F(x)|dx$ is finite.

Note that the left hand side of the equation in Proposition 3 is weakly greater than $F(x)^{n-1}$. Thus, $F(x)^{n-1}$ must be weakly less than the right hand side of that equation, which implies that

$$0 \leq F(x) \leq kx^{\frac{1}{n-1}} \quad \forall x \in [0, \frac{n\mu}{m}], \tag{69}$$

where $k = \left(\frac{m}{n\mu}\right)^{\frac{1}{n-1}}$ is a constant. By (69), $\lim_{x\to 0}|\ln(x)F(x)|$ is weakly less than $\lim_{x\to 0}|\ln(x)kx^{\frac{1}{n-1}}|$, which, by L'Hôpital's rule, equals $\lim_{x\to 0}|-\frac{k}{n-1}x^{\frac{1}{n-1}}| = 0$. Thus, (i) holds. By (69), $\int_0^{\frac{n\mu}{m}} |\frac{1}{x}F(x)|dx$ is weakly less than $\int_0^{\frac{n\mu}{m}} |\frac{k}{x}x^{\frac{1}{n-1}}|dx$, which equals $k(n-1)\left(\frac{n\mu}{m}\right)^{\frac{1}{n-1}}$, a finite constant. Thus, (ii) also holds. Hence, integrability condition holds for $\ln(x)$, which completes the proof. □

# References


AKERLOF, R. J., AND R. T. HOLDEN (2012): "The Nature of Tournaments," *Economic Theory*, 51, 289–313.

ANKIRCHNER, S., D. HOBSON, AND P. STRACK (2014): "Finite, Integrable and Bounded Time Embeddings for Diffusions," *Bernoulli*, Forthcoming.

BARBERIS, N., AND M. HUANG (2008): "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, 98, 2066–2100.

BARELLI, P., S. GOVINDAN, AND R. WILSON (2014): "Competition for a Majority," *Econometrica*, 82, 271–314.

BAYE, M. R., D. KOVENOCK, AND C. G. DE VRIES (1993): "Rigging the Lobbying Process: An Application of the All-Pay Auction," *American Economic Review*, 83, 289–294.

——— (1996): "The All-Pay Auction with Complete Information," *Economic Theory*, 8, 291–305.

BERK, J. B., AND R. C. GREEN (2004): "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112, 1269–1295.

BHATTACHARYA, S., AND D. MOOKHERJEE (1986): "Portfolio Choice in Research and Development," *RAND Journal of Economics*, 17, 594–605.

BILLINGSLEY, P. (1985): *Probability and Measure*. John Wiley and Sons.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2012): "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics*, 127, 1243–1285.

BOYD, S., AND L. VANDENBERGHE (2004): *Convex Optimization*. Cambridge University Press.

BRUNNERMEIER, M. K., C. GOLLIER, AND J. A. PARKER (2007): "Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns," *American Economic Review*, 97, 159–165.

CHE, Y.-K., AND I. L. GALE (1998): "Caps on Political Lobbying," *American Economic Review*, 88, 643–651.

CHEN, J., E. HUGHSON, AND N. STOUGHTON (2012): "Strategic Mutual Fund Tournaments," *Mimeo*.

CHEVALIER, J., AND G. ELLISON (1997): "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105, 1167–1200.

CLARK, D. J., AND C. RIIS (1998): "Competition over More than One Prize," *American Economic Review*, 88, 276–289.

——— (2001): "Rank-Order Tournaments and Selection," *Journal of Economics*, 2, 167–191.

COVER, T. M. (1974): *Universal Gambling Schemes and the Complexity Measures of Kolmogorov and Chaitin*, Technical Report No. 12. Stanford Statistics Department.

CUMMING, D. J., AND S. A. JOHAN (2009): *Venture Capital and Private Equity Contracting: An International Perspective*. Elsevier Inc.

DASGUPTA, P., AND J. STIGLITZ (1980): "Uncertainty, Industrial Structure, and the Speed of R&D," *Bell Journal of Economics*, 11, 1–28.

DIAMOND, P. A., AND J. E. STIGLITZ (1974): "Increases in Risk and in Risk Aversion," *Journal of Economic Theory*, 8, 337–360.

DIJK, O., M. HOLMÉN, AND M. KIRCHLER (2014): "Rank Matters — The Impact of Social Competition on Portfolio Choice," *European Economic Review*, 66, 97–110.

DUBINS, L. E., AND L. J. SAVAGE (1965): *How to Gamble if You Must: Inequalities for Stochastic Processes.* McGraw-Hill.

ELLINGSEN, T. (1991): "Strategic Buyers and the Social Cost of Monopoly," *American Economic Review*, 81, 648–657.

FALKENSTEIN, E. G. (1996): "Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings," *Journal of Finance*, 51, 111–135.

FU, Q., AND J. LU (2009): "The Beauty of "Bigness": On Optimal Design of Multi-Winner Contests," *Games and Economic Behavior*, 66, 146–161.

GABA, A., AND A. KALRA (1999): "Risk Behavior in Response to Quotas and Contests," *Marketing Science*, 18, 417–434.

GABA, A., I. TSETLIN, AND R. L. WINKLER (2004): "Modifying Variability and Correlations in Winner-Take-All Contests," *Operations Research*, 52, 384–395.

GAVIOUS, A., B. MOLDOVANU, AND A. SELA (2002): "Bid Costs and Endogenous Bid Caps," *RAND Journal of Economics*, 33, 709–722.

GENTZKOW, M., AND E. KAMENICA (2014): "Costly Persuasion," *American Economic Review*, 104, 457–462.

GILPATRIC, S. M. (2009): "Risk Taking in Contests and the Role of Carrots and Sticks," *Economic Inquiry*, 47, 266–277.

GOEL, A. M., AND A. V. THAKOR (2008): "Overconfidence, CEO Selection, and Corporate Governance," *Journal of Finance*, 63, 2737–2784.

HAN, B., D. HIRSHLEIFER, AND J. C. PERSONS (2009): "Promotion Tournaments and Capital Rationing," *Review of Financial Studies*, 22, 119–255.

HARRIS, R., T. JENKINSON, AND R. STUCKE (2012): "Are Too Many Private Equity Funds Top Quartile?," *Journal of Applied Corporate Finance*, 24, 77–89.

HART, S. (2008): "Discrete Colonel Blotto and General Lotto Games," *International Journal of Game Theory*, 36, 441–460.

HILLMAN, A. L., AND D. SAMET (1987): "Dissipation of Contestable Rents by Small Numbers of Contenders," *Public Choice*, 54, 63–82.

HOSKING, J. R. M. (1990): "L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics," *Journal of the Royal Statistical Society. Series B (Methodological)*, 52, 105–124.

HUANG, J., K. D. WEI, AND H. YAN (2007): "Participation Costs and the Sensitivity of Fund Flows to Past Performance," *Journal of Finance*, 62, 1273–1311.

HVIDE, H. K. (2002): "Tournament Rewards and Risk Taking," *Journal of Labor Economics*, 20, 877–898.

HVIDE, H. K., AND E. G. KRISTIANSEN (2003): "Risk Taking in Selection Contests," *Games and Economic Behavior*, 42, 172–179.

JONES, M. C. (2002): "The Complementary Beta Distribution," *Journal of Statistical Planning and Inference*, 104, 329–337.

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

KAWAMURA, K., AND I. MORENO DE BARREDA (2014): "Biasing Selection Contests with Ex-Ane Identical Agents," *Economics Letters*, 123, 240–243.

KEMPF, A., AND S. RUENZI (2008a): "Family Matters: Rankings Within Fund Families and Fund Inflows," *Journal of Business Finance & Accounting*, 35, 177–199.

——— (2008b): "Tournaments in Mutual-Fund Families," *Review of Financial Studies*, 21, 1013–1036.

KLETTE, T., AND D. DE MEZA (1986): "Is the Market Biased against Risky R&D?," *RAND Journal of Economics*, 17, 133–139.

KONRAD, K. A. (2007): *Strategy in Contests: An Introduction*, WZB–Markets and Politics Working Paper SP II 2007-01. Social Science Research Center Berlin (WZB), Berlin.

KRÄKEL, M. (2008): "Optimal Risk Taking in an Uneven Tournament Game with Risk Averse Players," *Journal of Mathematical Economics*, 44, 1219–1231.

KRÄKEL, M., AND D. SLIWKA (2004): "Risk Taking in Asymmetric Tournaments," *German Economic Review*, 5, 103–116.

KVASOV, D. (2007): "Contests with Limited Resources," *Journal of Economic Theory*, 136, 738–748.

LAZEAR, E. P., AND S. ROSEN (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841–864.

LIZZERI, A. (1999): "Budget Deficits and Redistributive Politics," *Review of Economic Studies*, 66, 909–928.

LIZZERI, A., AND N. PERSICO (2001): "The Provision of Public Goods under Alternative Electoral Incentives," *American Economic Review*, 91, 225–239.

MEYER, M. A. (1991): "Learning from Coarse Information: Biased Contests and Career Profiles," *Review of Economic Studies*, 58, 15–41.

MOLDOVANU, B., AND A. SELA (2006): "Contest Architecture," *Journal of Economic Theory*, 126, 70–96.

MOLDOVANU, B., A. SELA, AND X. SHI (2012): "Carrots and Sticks: Prizes and Punishments in Contests," *Economic Inquiry*, 50, 453–462.

MONROE, I. (1972): "On Embedding Right Continuous Margingales in Brownian Motion," *Annals of Mathematical Statistics*, 43, 1293–1311.

MYERSON, R. B. (1993): "Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems," *American Political Science Review*, 87, 856–869.

NANDA, A. K., AND M. SHAKED (2001): "The Hazard Rate and the Reversed Hazard Rate Orders, with Applications to Order Statistics," *Annals of the Institute of Statistical Mathematics*, 53, 853–864.

NIEKEN, P., AND D. SLIWKA (2010): "Risk-Taking Tournaments — Theory and Experimental Evidence," *Journal of Economic Psychology*, 31, 254–268.

OBŁÓJ, J. (2004): "The Skorokhod Embedding Problem and its Offspring," *Probability Surveys*, 1, 321–392.

OJA, H. (1981): "On Location, Scale, Skewness and Kurtosis of Univariate Distributions," *Scandinavian Journal of Statistics*, 8, 154–168.

PYLE, D. H., AND S. J. TURNOVSKY (1970): "Safety-First and Expected Utility Maximization in Mean-Standard Deviation Portfolio Analysis," *Review of Economics and Statistics*, 52, 75–81.

REISS, R.-D. (1980): *Approximate Distributions of Order Statistics*. Springer.

ROBERSON, B. (2006): "The Colonel Blotto Game," *Economic Theory*, 29, 1–24.

ROBINSON, D. T., AND B. A. SENSOY (2013): "Do Private Equity Fund Managers Earn Their Fees? Compensation, Ownership, and Cash Flow Performance," *Review of Financial Studies*, 26, 2760–2797.

ROSE-ACKERMAN, S. (1991): "Risk Taking and Ruin: Bankruptcy and Investment Choice," *Journal of Legal Studies*, 20, 277–310.

RUBINSTEIN, M., AND H. E. LELAND (1981): "Replicating Options with Positions in Stock and Cash," *Financial Analysts Journal*, 37, 63–72.

RYVKIN, D. (2010): "The Selection Efficiency of Tournaments," *European Journal of Operational Research*, 206, 667–675.

RYVKIN, D., AND A. ORTMANN (2008): "The Predictive Power of Three Prominent Tournament Formats," *Management Science*, 54, 492–504.

SAHUGUET, N., AND N. PERSICO (2006): “Campaign Spending Regulation in a Model of Redistributive Politics,” *Economic Theory*, 28, 95–124.

SATO, M. (1997): “Some Remarks on the Mean, Median, Mode and Skewness,” *Australian Journal of Statistics*, 39, 219–224.

SEEL, C., AND P. STRACK (2013): “Gambling in Contests,” *Journal of Economic Theory*, 148, 2033–2048.

SICULUS, D. (1718): *The Historical Library of Diodorus the Sicilian*, vol. XVII. John Peele at Mr. Lock’s Head.

SIEGEL, R. (2009): “All-Pay Contests,” *Econometrica*, 77, 71–92.

SIRRI, E. R., AND P. TUFANO (1998): “Costly Search and Mutual Fund Flows,” *Journal of Finance*, 53, 1589–1622.

SKAPERDAS, S. (1996): “Contest Success Functions,” *Economic Theory*, 7, 283–290.

SKOROKHOD, A. V. (1965): *Studies in the Theory of Random Processes*. Addison-Wesley Publishing Company.

TAYLOR, J. (2003): “Risk-Taking Behavior in Mutual Fund Tournaments,” *Journal of Economic Behavior & Organization*, 50, 373–383.

TSIANG, S. C. (1972): “The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money,” *American Economic Review*, 62, 354–371.

TULLOCK, G. (1980): *Efficient Rent Seeking. In: Buchanan, James M., Robert D. Tollison, and Gordon Tullock (eds.) Toward a Theory of the Rent-Seeking Society, pp. 3–15.* Texas A&M University Press.

VAN LONG, N. (2013): “The Theory of Contests: A Unified Model and Review of the Literature,” *European Journal of Political Economy*, 32, 161–181.

VARIAN, H. R. (1980): “A Model of Sales,” *American Economic Review*, 70, 651–659.

WAGNER, N., AND E. WINTER (2013): “A New Family of Equity Style Indices and Mutual Fund Performance: Do Liquidity and Idiosyncratic Risk Matter?,” *Journal of Empirical Finance*, 21, 69–85.

**For online publication only**

# Supplement to "Skewing the odds: Strategic risk taking in contests"

by Dawei Fang and Thomas Noe

**Abstract**

*This document provides a demonstration that the symmetric Nash equilibrium for the certain-capacity contest game studied in Section 3 is, in fact, the unique Nash equilibrium.*

In the body of the paper, we restricted our attention to symmetric equilibria. Under the assumption of symmetry, it was fairly easy to rule out equilibria in which contestants choose discontinuous performance distributions. In this online appendix, we consider asymmetric equilibria. Our goal is to show that no asymmetric equilibria exist for the contest game studied in Section 3. This result will show that the unique symmetric equilibrium identified in Proposition 3 is, in fact, the unique Nash equilibrium.

## S.1 Preliminary results

Absent the assumption of symmetry, discontinuous performance distributions are more difficult to dismiss, and thus we need to develop a few more preliminary results which, although unsurprising, are somewhat tedious to demonstrate. We first need to define the probability of winning function faced by a contestant, say contestant $i$, when some other contestants use discontinuous performance distributions that place positive mass on specific performance levels. The complication introduced by such performance distributions is the possibility of ties: two or more contestants may obtain the same performance with positive probability. We will show that such discontinuities may lead to the non-existence of best replies but that, when best replies exist, they satisfy exactly the same characterizations as developed in Section 2.1.

To take into account the issue of ties, we first define a *minimal prize-worthy performance level* for a contest with $n$ contestants and $m$ prizes. For a given vector of performance levels, $X = (x_1, \ldots x_n)$, and number of prizes, $m$, the minimal prize-worthy performance faced by $i$ given the performance levels of the other contestants, $X_{-i}$, equals

$$\text{mpw}(X_{-i}) = \min_{j \neq i} \{x_j : \#(\{k \neq i : x_k > x_j\}) < m\}.$$

A prize-worthy performance level for $i$ given $X_{-i}$ is any performance level at least equal to the minimal prize-worthy performance level. If $i$'s performance level exactly equals $\text{mpw}(X_{-i})$, then more than $m$ contestants would have prize-worthy performance. Thus, if prizes were awarded to all prize-worthy performances, more than $m$ prizes would be offered. To keep the number of prizes equal to $m$, prizes offered to "marginal contestants," those contestants who obtained the minimal prize-worthy performance level, have to be rationed. We assume that the rationing rule allocates the prizes among the marginal contestants in such a way that each has a positive probability of receiving the prize, which implies, because of the surplus of marginal contestants relative to prizes, that no rationed marginal contestant receives a price with certainty.[1]

Given that contestant $i$'s competitors submit random performance levels, $\tilde{X}_{-i}$, the minimal prize-worthy performance level for $i$ is also a random variable, $\text{mpw}(\tilde{X}_{-i})$. Let $P_i$ be the CDF of $\text{mpw}(\tilde{X}_{-i})$, i.e.,

$$P_i(x) = \text{Prob}[\text{mpw}(\tilde{X}_{-i}) \leq x]. \tag{S-1}$$

Note that $P_i$ is the probability of winning function for $i$ ignoring rationing. Because $P_i$ is a CDF, it is nonnegative, nondecreasing, right continuous, and bounded above by 1. Because of rationing, contestant $i$'s probability of winning, given his performance level $x_i$, can be lower than $P_i(x_i)$. Contestant $i$ might have performance exactly equal to the realization of $\text{mpw}(\tilde{X}_{-i})$. This event can only occur with positive probability at points on which the distribution of $\text{mpw}(\tilde{X}_{-i})$ assigns positive probability. These are the discontinuity points of $P_i$. Let $\mathbb{D}(P_i)$ denote these discontinuity points. Note that $\mathbb{D}(P_i)$ is a countable set. Let $O_i$ be the actual probability of winning function for $i$. Then we have

$$O_i(x_i) = P_i(x_i) - \ell_i(x_i),$$

where $\ell_i$ is the function of rationing penalty, i.e., the reduction in the probability of winning in the case of prize rationing, and $\ell_i$ satisfies $\ell_i(x_i) \in (0, P_i(x_i) - P_i(x_i^-))$ for all

[1] It is clear that the random tie-breaking rule assumed in the main body of the paper satisfies such a condition.

$x_i \in \mathbb{D}(P_i)$ and $\ell_i(x_i) = 0$ for all $x_i \notin \mathbb{D}(P_i)$.

Because $P_i$ is right continuous and nondecreasing and thus upper semi-continuous, the map $dF_i \mapsto \int P_i \, dF_i$ is upper semi-continuous. However, $O_i$ is generally not upper semi-continuous. Hence, under $O_i$, the problem of choosing the optimal performance distribution may not have a solution. However, the following theorem and its corollary show that, if a best reply exists for contestant $i$ to $O_i$, then this best reply is also a best reply to $P_i$. Thus, when a best-reply distribution exists, this distribution satisfies all of the characterizations of optimal performance distributions developed in Section 2.1. The exact rule for rationing prizes under ties has no effect on the character of the best reply, which is always entirely determined by the CDF $P_i$. The intuition behind this result is that there are only a countable number of discontinuities in $O_i$. Placing weight on any of them triggers a loss from rationing. Since the complement of a countable subset of the positive real line is dense, contestant $i$ can avoid the discontinuity points and approach the payoff under $P_i$ to an arbitrary degree of accuracy. The only effect of the rationing penalty, $\ell_i$, is that it may block the payoff of the sequence of distributions that approach the payoff under $P_i$ from equalling the payoff of the limit distribution. This generates a discontinuity, which, when it occurs, implies that no best reply exists for contestant $i$. This intuition is given by Figure S.1. The key insight from the figure is that the upper support lines are the same for $O_i$, the actual probability of winning function under the rationing rule, and $P_i$, the probability of winning function absent rationing. However, under $O_i$, the supremum of $i$'s payoff is never attained. Thus, no best reply exists for contestant $i$ under $O_i$.

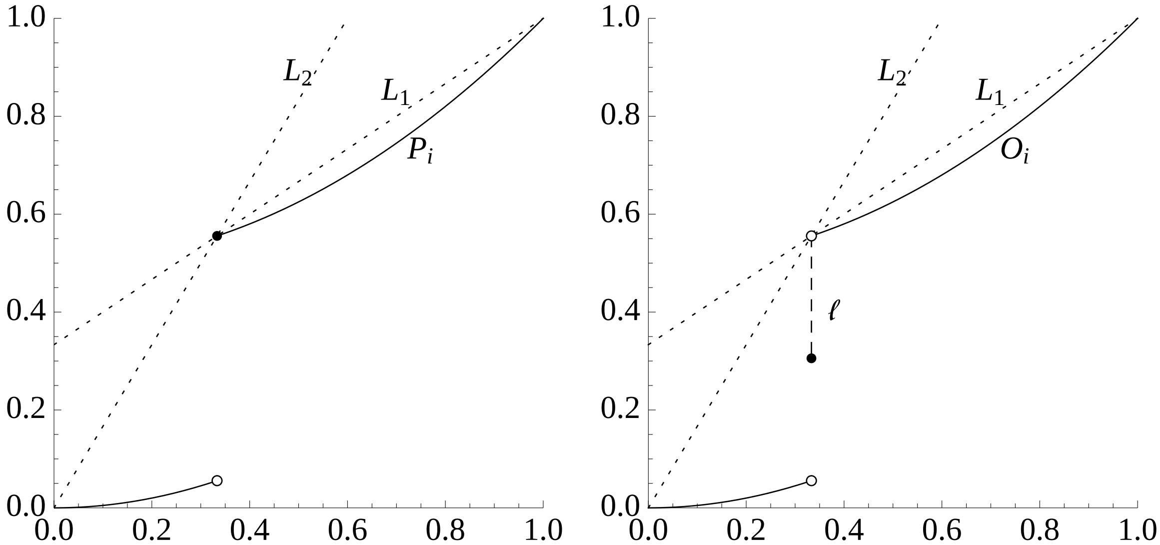



Figure S-1: An illustration of the effect of discontinuity

In a Nash equilibrium, by definition, each contestant has a best reply to the distributions chosen by the other contestants. Thus, when characterizing Nash equilibria of the contest game, we can assume without loss of generality that contestant $i$ plays best replies to the CDF $P_i$ rather than to the actual probability of winning function $O_i$. Although these assertions are by no means surprising, formally establishing them is fairly tedious. So for the details of the proof, we refer the reader to Section S.3. The key result is presented in the following theorem.

**Theorem S.1.** *Let $P$ be any cumulative probability distribution function. Let $\mathbb{D}(P)$ be the set of its discontinuity points. Let $\mathscr{C}$ be a set of measures on the nonnegative real line such that, for all $dF \in \mathscr{C}$, $dF$ satisfies*

$$\int_0^\infty x\,dF(x) \leq \mu \text{ and } \int_0^\infty dF(x) \leq 1\,, \text{ where } \mu > 0.$$

*Define $O$ by*

$$O(x) = P(x) - \ell(x),$$

*where $\ell(x) \in (0, P(x) - P(x^-))$ for all $x \in \mathbb{D}(P)$ and $\ell(x) = 0$ for all $x \notin \mathbb{D}(P)$. Then we must have*

$$\sup_{dF \in \mathscr{C}} \int_0^\infty O(x)\,dF(x) = \max_{dF \in \mathscr{C}} \int_0^\infty P(x)\,dF(x). \tag{S-2}$$

*Proof.* See Section S.3. □

**Corollary S.1.** *If a solution $dF^*$ exists to the problem*

$$\max_{dF \in \mathscr{C}} \int_0^\infty O(x)\,dF(x), \tag{S-3}$$

*then we must have*

*i.* $dF^*(\mathbb{D}(P)) = 0$;

*ii.* $\int_0^\infty O(x)dF^*(x) = \int_0^\infty P(x)dF^*(x)$;

*iii.* $dF^*$ *also solves* $\max_{dF \in \mathscr{C}} \int_0^\infty P(x)\,dF(x)$.

*Proof.* (i): If $dF^*(\mathbb{D}(P)) > 0$, then

$$\max_{dF \in \mathscr{C}} \int_0^\infty O(x)\,dF(x) = \int_0^\infty O(x)dF^*(x) < \int_0^\infty P(x)dF^*(x) \leq \max_{dF \in \mathscr{C}} \int_0^\infty P(x)\,dF(x),$$

contradicting Theorem S.1.

(ii) and (iii): (ii) follows immediately from (i). Given that $dF^*$ solves problem (S-3), (iii) follows immediately from (ii) and Theorem S.1. □

## S.2 Proof of uniqueness

In the following analysis, let $\mathscr{N}$ be the set of all the $n$ contestants. For all $i \in \mathscr{N}$, let $P_i$, which is defined by (S-1), be contestant $i$'s probability of winning function ignoring rationing and $F_i$ his distributional choice. By Corollary S.1, given the existence of equilibria, to characterize the equilibrium distribution, we only need to examine the best reply to $P$ rather than to the actual probability of winning function. This problem is expressed by (1) with $P$ interpreted as the probability of winning ignoring rationing. Given that this $P$ is a CDF, the analysis developed in Section 2.1 still holds. Let $L_i$ be contestant $i$'s upper support line, i.e., $L_i(x) = \alpha_i + \beta_i x$ for all $x \geq 0$, where $\alpha_i$ and $\beta_i$ are nonnegative optimal dual variables for contestant $i$'s dual problem (4). Let $V_i$ be contestant $i$'s expected probability of winning in equilibrium. Based on the analysis in Section 2.1 and S.1, we obtain the next result.

**Lemma S.1.** *For all $i \in \mathscr{N}$, the following results hold:*

*i.* $P_i(x) \leq L_i(x)$ *for all* $x \geq 0$,

*ii.* $dF_i\{x \geq 0 : P_i(x) < L_i(x)\} = 0$,

*iii. and* $V_i = L_i(\mu)$.

*Proof.* (i) and (ii): Follow immediately from equation (8) and Corollary S.1.

(iii): Follows immediately from equation (7) and Corollary S.1. □

**Lemma S.2.** *For all $i \in \mathscr{N}$, $\beta_i > 0$ and $L_i(x) > 0$ for all $x > 0$.*

*Proof.* Since all the contestants have the same capacity, it is obvious that $i$ cannot guarantee strictly outperforming any of his competitors. Thus, $i$ cannot guarantee winning a prize. Then by Corollary S.1, $i$ must not guarantee winning a prize even if we ignore prize rationing in the case of ties. Thus, $\beta_i$, which measures the marginal value of capacity under $P_i$, must be strictly positive. This implies that, for all $x > 0$, $L_i(x) > L_i(0) = \alpha_i \geq 0$. □

**Lemma S.3.** *No one places point mass on $(0, \infty)$ and $P_i$ is continuous on $(0, \infty)$ for all $i \in \mathscr{N}$.*

*Proof.* First, we show that no one places point mass on $(0, \infty)$ by way of contradiction. Suppose otherwise; so there exists contestant $i$ who places point mass on $x' > 0$. Then by Lemmas S.1 and S.2, we must have $P_i(x') = L_i(x') > 0$, which implies that at least $(n-m)$ of $i$'s competitors each place weight on $[0, x']$. Thus, given that $i$ places point mass on $x'$, for any $j \neq i$, at least $(n-m)$ of $j$'s competitors each place weight on $[0, x']$. Thus, $P_j(x') > 0$ for all $j \neq i$. Moreover, given that $P_j(x') > 0$, contestant $i$'s point mass on $x'$ further implies discontinuity of $P_j$ at $x'$: $P_j(x') > P_j(x'^-)$ for all $j \neq i$. Since, by Lemma S.1, $P_j \leq L_j$, and since $L_j$ is continuous, we must have $P_j(x'^-) < L_j(x'^-)$ for all $j \neq i$. Thus, given that $L_j$ is continuous and $x' > 0$, there must exist $\varepsilon > 0$ such that $P_j(x'^-) < L_j(x)$ for all $x \in [x' - \varepsilon, x')$. Thus, given that $P_j$ is nondecreasing, we must have $P_j(x) < L_j(x)$ for all $x \in [x' - \varepsilon, x')$ and $j \neq i$. This implies, by Lemma S.1, that none of $i$'s competitors places probability weight on $[x' - \varepsilon, x')$. Note also that none of $i$'s competitors places point mass on $x'$. This is because if any of them placed point mass on $x'$, $P_i$ would have discontinuity at $x'$ and, hence, $i$'s point mass on $x'$ would contradict Corollary S.1. Thus, none of $i$'s competitors places weight on $[x' - \varepsilon, x']$. Hence, $i$ can relax his capacity constraint without losing any probability of winning by transferring point mass from $x'$ to $x' - \varepsilon$. By Lemma S.2, capacity has strictly positive marginal value. Thus, such a transfer makes $i$ strictly better off, a contradiction.

Given that no one places point mass on $(0, \infty)$, $P_i$ must be continuous on $(0, \infty)$ for all $i \in \mathcal{N}$. □

**Lemma S.4.** *For any two distinct contestants $i$ and $j$, if $F_j(x) \leq F_i(x)$ for all $x \geq x'$, where $x' > \mu$, there must exist $x^o \in [0, x')$ such that $L_j(x^o) \leq L_i(x^o)$.*

*Proof.* Suppose not; so

$$L_j(x) > L_i(x) \quad \forall x \in [0, x'), \text{ where } x' > \mu. \tag{S-4}$$

First, we claim that $F_j(0) = 0$ and, hence, $F_j(0) \leq F_i(0)$. We prove this by way of contradiction. Suppose $F_j(0) > 0$, which, by Lemma S.1, implies that $P_j(0) = L_j(0)$. By hypothesis (S-4), $L_j(0) > L_i(0) \geq 0$. Thus, $P_j(0) > 0$. However, $P_j(0) > 0$ implies discontinuity of $P_j$ at $x = 0$: $P_j(0) > P_j(0^-)$. By Corollary S.1, this implies that $j$ places no mass on 0, which contradicts $F_j(0) > 0$. This contradiction establishes our claim.

Next, since $F_i$ and $F_j$ have the same mean, $F_j$ must *not* first-order stochastically dominate $F_i$. Thus, given that $0 = F_j(0) \leq F_i(0)$ and $F_j(x) \leq F_i(x)$ for all $x \geq x'$, it must be that either (i) $F_j(x) = F_i(x)$ for all $x \geq 0$ or (ii) there exists $\tilde{x} \in (0, x')$ such that $F_j(\tilde{x}) > F_i(\tilde{x})$.

If (i) occurs, we must have $V_i = V_j$ and thus, by Lemma S.1, $L_i(\mu) = L_j(\mu)$, which contradicts hypothesis (S-4). Thus, (i) is impossible.

Now consider (ii). Given that $F_i(\tilde{x}) \geq 0$, $F_j(\tilde{x}) > F_i(\tilde{x})$ implies that $F_j(\tilde{x}) > 0$, which further implies that contestant $j$ places weight on $[0, \tilde{x}]$. Thus, by Lemma S.1 and the continuity of $P_j$ and $L_j$ on $(0, \infty)$, there must exist $\bar{x} \in [0, \tilde{x}]$ such that

$$\bar{x} = \max\{x \in [0, \tilde{x}] : P_j(x) = L_j(x)\}. \tag{S-5}$$

Note that $F_j(\tilde{x}) > F_i(\tilde{x})$ implies

$$P_j(\tilde{x}) \leq P_i(\tilde{x}). \tag{S-6}$$

By hypothesis (S-4) and $\tilde{x} \in (0, x')$, we must have $L_i(\tilde{x}) < L_j(\tilde{x})$ and thus, by Lemma S.1, $P_i(\tilde{x}) \leq L_i(\tilde{x}) < L_j(\tilde{x})$. Thus, by (S-6), we must have $P_j(\tilde{x}) < L_j(\tilde{x})$. Thus, by (S-5), we must have $\bar{x} < \tilde{x}$ and $P_j(x) \neq L_j(x)$ for all $x \in (\bar{x}, \tilde{x}]$. Thus, by Lemma S.1, $j$ places no weight on $(\bar{x}, \tilde{x}]$. However, given $\bar{x} < \tilde{x} < x'$, by (S-5), (S-4), and Lemma S.1, $P_j(\bar{x}) = L_j(\bar{x}) > L_i(\bar{x}) \geq P_i(\bar{x})$. This implies, given that $j$ places no weight on $(\bar{x}, \tilde{x}]$, that $P_j(\tilde{x}) > P_i(\tilde{x})$, which contradicts equation (S-6). This completes the proof. □

In what follows, for all $i \in \mathscr{N}$, define $\mathscr{S}_i$ by

$$\mathscr{S}_i = \{x > 0 : P_i(x) = L_i(x)\}, \tag{S-7}$$

define $\hat{x}_i$ by

$$\hat{x}_i = \max \mathscr{S}_i, \tag{S-8}$$

and define $\hat{x}$ by

$$\hat{x} = \max\{\hat{x}_i\}_{i \in \mathscr{N}}. \tag{S-9}$$

**Lemma S.5.** *For all $i \in \mathscr{N}$,*

*i.* $\mathscr{S}_i \neq \emptyset$*, where* $\mathscr{S}_i$ *is defined by* (S-7)*;*

*ii.* $\hat{x}_i$*, defined by* (S-8)*, exists;*

*iii.* $\hat{x}_i = \hat{x} > \mu$*, where* $\hat{x}$ *is defined by* (S-9)*;*

*iv.* $P_i(\hat{x}) = L_i(\hat{x}) = 1$.

*Proof.* (i): Clearly, in any equilibrium, contestants must all place some weight on the strictly positive real line. This, by Lemma S.1, implies (i).

(ii): Lemma S.3, $P_i$ is continuous on $(0, \infty)$. Thus, (ii) follows immediately from (i) and the continuity of $L_i$ and $P_i$ on $(0, \infty)$.

(iii): By Lemma S.3 and the fact that the mean of $F_i$ equals $\mu$, we must have $\hat{x}_i > \mu$ for all $i \in \mathscr{N}$. Without loss of generality, assume that contestant $j$ has $\hat{x}_j = \max\{\hat{x}_i\}_{i \in \mathscr{N}} = \hat{x}$. Thus, by the definition of $\hat{x}_j$, $P_j(\hat{x}) = L_j(\hat{x})$. By the definition of $\hat{x}$ and Lemma S.1, no one places weight above $\hat{x}$. Thus,

$$P_j(\hat{x}) = L_j(\hat{x}) = 1. \tag{S-10}$$

We now prove (iii) by way of contradiction. Suppose, contrary to (iii), that there exists contestant $k \neq j$ such that $\hat{x}_k \in (\mu, \hat{x})$. This, by the definition of $\hat{x}_k$ and the fact that $L_k \geq P_k$, implies that $L_k(\hat{x}) > P_k(\hat{x})$. Since no one places weight above $\hat{x}$, $P_k(\hat{x}) = 1$.

Thus, $L_k(\hat{x}) > 1$. This, by (S-10), implies that

$$L_k(\hat{x}) > L_j(\hat{x}). \tag{S-11}$$

By the definition of $\hat{x}_k$ and Lemma S.1, $F_k(\hat{x}_k) = 1$. Thus, $F_k(\hat{x}_k) \geq F_j(\hat{x}_k)$, which implies that $P_k(\hat{x}_k) \leq P_j(\hat{x}_k)$. This further implies, by the facts that $P_k(\hat{x}_k) = L_k(\hat{x}_k)$ and $P_j \leq L_j$, that $L_k(\hat{x}_k) \leq L_j(\hat{x}_k)$. This, together with (S-11) and the fact that both $L_j$ and $L_k$ are upward sloping straight lines, implies that

$$L_j(x) > L_k(x) \quad \forall x \in [0, \hat{x}_k). \tag{S-12}$$

However, since $F_j(x) \leq 1 = F_k(x)$ for all $x \geq \hat{x}_k$, where $\hat{x}_k > \mu$, equation (S-12) contradicts Lemma S.4. This contradiction completes the proof of (iii).

(iv): Follows immediately from (iii) and equation (S-10). □

**Lemma S.6.** *All the contestants have the same upper support line.*

*Proof.* Consider any pair of distinct contestants $i$ and $j$. By Lemma S.5, $L_i(\hat{x}) = L_j(\hat{x})$, where $\hat{x} > \mu$. By the definition of $\hat{x}$, $F_i(x) = F_j(x) = 1$ for all $x \geq \hat{x}$. Thus, by Lemma S.4, neither $L_i$ nor $L_j$ can lie above the other everywhere on $[0, \hat{x})$. Since $L_i$ and $L_j$ are straight lines which meet at $\hat{x}$, they must completely overlap. □

The next lemma is a supplementary statistical result. To avoid interrupting the flow of the argument of equilibrium uniqueness, we put its proof in Section S.3.

**Lemma S.7.** *For any two distinct contestants $i$ and $j$,*

*i. if $P_i(x) > P_j(x)$, it must be that $F_i(x) < F_j(x)$;*
*ii. if $P_i(x) = P_j(x) \in (0,1)$, it must be that $F_i(x) = F_j(x)$;*
*iii. if $P_i(x) \in (0,1)$ and $P_i(x) \geq P_j(x)$, it must be that $F_i(x) \leq F_j(x)$.*

*Proof.* See Section S.3. □

**Lemma S.8.** *$\mathscr{S}_i$, defined by (S-7), is a connected set for all $i \in \mathscr{N}$.*

*Proof.* Suppose not; so there exists contestant $j$ such that $\mathscr{S}_j$ is not connected. Lemma S.3 and the continuity of $L_j$ imply that $(0, \hat{x}] \backslash \mathscr{S}_j$ is a combination of open intervals. Thus, given that $\hat{x} \in \mathscr{S}_j$, which is implied by Lemma S.5, if $\mathscr{S}_j$ is not connected, there must exist an open interval $(a,b)$, where $0 < a < b \leq \hat{x}$, such that $(a,b) \not\subset \mathscr{S}_j$ while $a, b \in \mathscr{S}_j$.

By Lemma S.1, this implies

$$P_j(a) = L_j(a) \tag{S-13}$$

$$P_j(b) = L_j(b) \tag{S-14}$$

$$P_j(x) < L_j(x) \quad \forall x \in (a,b). \tag{S-15}$$

For any contestant $k \neq j$, by Lemma S.1, $P_k(a) \leq L_k(a)$, and, by Lemma S.6, $L_j = L_k$. Thus, equation (S-13) implies

$$P_k(a) \leq P_j(a). \tag{S-16}$$

Given that $a \in (0, \hat{x})$, by equation (S-13) and Lemmas S.2 and S.5, $0 < P_j(a) < 1$. Thus, by (S-16), we must have $P_k(a) \leq P_j(a) \in (0,1)$, which, by Lemma S.7, implies

$$F_k(a) \geq F_j(a). \tag{S-17}$$

Since $(a,b) \not\subset \mathscr{S}_j$, by Lemma S.1, $j$ places no weight on $(a,b)$. Thus,

$$F_j(x) = F_j(a) \quad \forall x \in (a,b). \tag{S-18}$$

Since $F_k$ is nondecreasing, equations (S-17) and (S-18) imply that $F_k(x) \geq F_j(x)$ for all $x \in (a,b)$, which further implies that $P_k(x) \leq P_j(x)$ for all $x \in (a,b)$. Thus, by equation (S-15) and Lemma S.6, we must have $P_k(x) < L_k(x)$ for all $x \in (a,b)$. Thus, by Lemma S.1, $k$ places no weight on $(a,b)$. Thus, none of $j$'s competitor places any weight on $(a,b)$. This, together with Lemma S.3, implies that $P_j(b) = P_j(a)$. Then, by equations (S-13) and (S-14), we must have $L_j(b) = L_j(a)$, which implies that $\beta_j = 0$. This contradicts Lemma S.2. This contradiction establishes the result. □

**Lemma S.9.** *$P_i(x) = L_i(x)$ for all $i \in \mathscr{N}$ and $x \in (0, \hat{x}]$, where $\hat{x}$ is defined by* (S-9).

*Proof.* Lemma S.5 ensures the existence of $\underline{x}_i$, defined by

$$\underline{x}_i = \inf \mathscr{S}_i \quad \forall i \in \mathscr{N}, \tag{S-19}$$

where $\mathscr{S}_i$ is defined by (S-7). By Lemmas S.5 and S.8, to show Lemma S.9, it suffices to show that

$$\underline{x}_i = 0 \quad \forall i \in \mathscr{N}. \tag{S-20}$$

To establish (S-20), we first show that

$$\underline{x}_i = \underline{x} \quad \forall i \in \mathscr{N}, \text{ where } \underline{x} = \min\{\underline{x}_i\}_{i\in\mathscr{N}} \geq 0. \tag{S-21}$$

We establish (S-21) by way of contradiction. Without loss of generality, suppose that contestant $k$ has $\underline{x}_k = \min\{\underline{x}_i\}_{i\in\mathscr{N}}$ and contestant $j \neq k$ has $\underline{x}_j = \max\{\underline{x}_i\}_{i\in\mathscr{N}}$. If (S-21) does not hold, we must have

$$0 \leq \underline{x}_k < \underline{x}_j.$$

By (S-19), Lemmas S.5 and S.8, the definition of $\underline{x}_j$, and the continuity of $P_i$ at $\underline{x}_j > 0$ for all $i \in \mathscr{N}$, which is implied by Lemma S.3, we must have $P_i(x) = L_i(x)$ for all $x \in [\underline{x}_j, \hat{x}]$ and $i \in \mathscr{N}$. Hence, by Lemmas S.2 and S.5, $P_i(x) \in (0,1)$ for all $x \in [\underline{x}_j, \hat{x})$ and $i \in \mathscr{N}$. Thus, by Lemma S.6, $P_j(x) = P_k(x) \in (0,1)$ for all $x \in [\underline{x}_j, \hat{x})$, which, by Lemma S.7, implies

$$F_j(x) = F_k(x) \quad \forall x \in [\underline{x}_j, \hat{x}). \tag{S-22}$$

By Lemma S.1, $P_j \leq L_j$. Thus, by the definition of $\underline{x}_j$,

$$P_j(x) < L_j(x) \quad \forall x \in (0, \underline{x}_j). \tag{S-23}$$

By equation (S-23) and Lemmas S.1 and S.3, $j$ places no weight on $(0, \underline{x}_j]$. Thus, $F_j(x)$ is constant on $[0, \underline{x}_j]$. Thus, given that $F_k$ is nondecreasing, by (S-22), we must have

$$F_j(x) \geq F_k(x) \quad \forall x \in [0, \hat{x}). \tag{S-24}$$

Given that $\underline{x}_k < \underline{x}_j$, by Lemmas S.5 and S.8 and the definition of $\underline{x}_k$, we must have $P_k(x) = L_k(x)$ for all $x \in (\underline{x}_k, \hat{x}]$. This implies, by Lemma S.6 and equation (S-23), that $P_k(x) > P_j(x)$ for all $x \in (\underline{x}_k, \underline{x}_j)$. This further implies, given that the rule for allocating prizes is symmetric, that $F_j$ and $F_k$ cannot be the same distribution. Thus, by equation (S-24) and the fact that $F_j(\hat{x}) = F_k(\hat{x}) = 1$, $F_k$ first-order stochastically dominates $F_j$, which contradicts that $F_k$ and $F_j$ have the same mean, and (S-21) follows.

Given (S-21), to show (S-20), all we need to show is $\underline{x} = 0$. We show this by way of contradiction. Suppose $\underline{x} > 0$. Then by the definition of $\underline{x}_i$ and the fact that both $P_i$ and $L_i$ are continuous on $(0, \infty)$, we must have $P_i(\underline{x}) = L_i(\underline{x})$ for all $i \in \mathscr{N}$. Thus, by Lemma S.2, $P_i(\underline{x}) > 0$ for all $i \in \mathscr{N}$. This implies, given the fact that no one places weight on $(0, \underline{x})$, that $P_i(0) > 0$ for all $i \in \mathscr{N}$. Thus, $P_i$ is discontinuous at 0: $P_i(0) > P_i(0^-)$. This implies, by Corollary S.1, that no one places weight on 0. Thus, $P_i(0) = 0$, which contradicts $P_i(0) > 0$. This contradiction establishes equation (S-20) and completes the proof. □

**Lemma S.10.** *There are no asymmetric equilibria.*

*Proof.* For any pair of distinct contestants $i$ and $j$, Lemmas S.2, S.6, and S.9 imply that $P_i(x) = P_j(x) \in (0,1)$ for all $x \in (0,\hat{x})$. Thus, by Lemma S.7, $F_i(x) = F_j(x)$ for all $x \in (0,\hat{x})$. Thus, by the definition of $\hat{x}$ and the right-continuity of CDF functions, for all $x \geq 0$, $F_i(x) = F_j(x)$, and the result follows. $\square$

## S.3 Proofs of Theorem S.1 and Lemma S.7

*Proof of Theorem S.1:* First note that, since $O \leq P$,

$$\sup_{dF\in\mathscr{C}} \int_0^\infty O(x)\,dF(x) \leq \sup_{dF\in\mathscr{C}} \int_0^\infty P(x)\,dF(x). \tag{S-25}$$

To establish the reverse inequality, we require the following lemma.

**Lemma S.11.** *For every $\varepsilon > 0$ and measure $dF \in \mathscr{C}$, there exists a measure $d\bar{F} \in \mathscr{C}$ such that*

$$\int_0^\infty P(x)\,d\bar{F}(x) \geq \frac{\mu}{\mu+\varepsilon}\int_0^\infty P(x)\,dF(x) \text{ and } d\bar{F}(\mathbb{D}(P)) = 0.$$

*Proof.* Let $\mathbb{D}(dF)$ be the set of points on which $dF$ places point weight. Note first that we can decompose $dF$ into two measures: a measure $dF^o$ that lives outside of $\mathbb{D}(P)\cap\mathbb{D}(dF)$ and a discrete measure $dF^D$ that lives on $\mathbb{D}(P)\cap\mathbb{D}(dF)$, i.e.,

$$dF = dF^o + dF^D \text{ and } dF^D(\{x\}) = \begin{cases} dF(\{x\}) & \text{if } x \in \mathbb{D}(P)\cap\mathbb{D}(dF) \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$\int_0^\infty P(x)\,dF(x) = \int_0^\infty P(x)\,dF^o(x) + \int_0^\infty P(x)\,dF^D(x). \tag{S-26}$$

Now for each $x_i \in \mathbb{D}(P)\cap\mathbb{D}(dF)\cap[0,\infty)$, select $\bar{x}_i \in (x_i, x_i+\varepsilon]\backslash\mathbb{D}(P)$. Given that $\mathbb{D}(P)$ is a countable set, $\bar{x}_i$ exists and $\bar{x}_i > 0$. Define, for each $x_i \in \mathbb{D}(P)\cap\mathbb{D}(dF)\cap[0,\infty)$, the measure

$$d\bar{F}_i(\{x\}) = \begin{cases} \frac{\mu}{\mu+\varepsilon}\,dF^D(\{x_i\}) & \text{if } x = \bar{x}_i \\ 0 & \text{otherwise} \end{cases}.$$

Note that, since $\bar{x}_i > x_i$ and $P$ is nondecreasing, for each $x_i \in \mathbb{D}(P)\cap\mathbb{D}(dF)\cap[0,\infty)$,

$$\int_0^\infty P(x)\,d\bar{F}_i(x) \geq \frac{\mu}{\mu+\varepsilon}P(x_i)dF^D(\{x_i\}). \tag{S-27}$$

Now define the measure

$$d\bar{F} = \frac{\mu}{\mu+\varepsilon} dF^o + \sum_i d\bar{F}_i. \tag{S-28}$$

Equations (S-26), (S-27), and (S-28) imply that

$$\int_0^\infty P(x)\, d\bar{F}(x) \geq \frac{\mu}{\mu+\varepsilon} \int_0^\infty P(x)\, dF(x).$$

By construction, $d\bar{F}$ places no weight on the countable set $\mathbb{D}(P)$. By construction, it is clear that $\int_0^\infty d\bar{F}(x) \leq \int_0^\infty dF(x)$. Moreover, given the facts that $\bar{x}_i \leq x_i + \varepsilon$, $\int_0^\infty x dF(x) \leq \mu$, and $\int_0^\infty dF(x) \leq 1$, by construction, we must have

$$\begin{aligned}
\int_0^\infty x d\bar{F}(x) &\leq \frac{\mu}{\mu+\varepsilon} \int_0^\infty x dF^o(x) + \frac{\mu}{\mu+\varepsilon} \int_0^\infty (x+\varepsilon) dF^D(x) \\
&\leq \frac{\mu}{\mu+\varepsilon} \int_0^\infty (x+\varepsilon) dF(x) \\
&\leq \mu.
\end{aligned}$$

Thus, $d\bar{F} \in \mathscr{C}$. Hence, $d\bar{F}$ satisfies the hypothesis of the lemma. □

Using Lemma S.11 we next show that

$$\sup_{dF \in \mathscr{C}} \int_0^\infty O(x)\, dF(x) \geq \sup_{dF \in \mathscr{C}} \int_0^\infty P(x)\, dF(x). \tag{S-29}$$

To establish (S-29), let $(dF_n)_n$ be a sequence of measures in $\mathscr{C}$ such that

$$\int_0^\infty P(x)\, dF_n(x) \to \sup_{dF \in \mathscr{C}} \int_0^\infty P(x)\, dF(x).$$

By Lemma S.11, there exists a sequence of measures $(d\bar{F}_n)_n$ in $\mathscr{C}$ such that

$$\int_0^\infty P(x)\, d\bar{F}_n(x) \geq \frac{\mu}{\mu+\frac{1}{n}} \int_0^\infty P(x)\, dF_n(x) \text{ and } d\bar{F}_n(\mathbb{D}(P)) = 0.$$

Thus,

$$\int_0^\infty P(x)\, d\bar{F}_n(x) \to \sup_{dF \in \mathscr{C}} \int_0^\infty P(x)\, dF(x). \tag{S-30}$$

Since $d\bar{F}_n(\mathbb{D}(P)) = 0$,

$$\int_0^\infty P(x)\, d\bar{F}_n(x) = \int_0^\infty O(x)\, d\bar{F}_n(x) \leq \sup_{dF \in \mathscr{C}} \int_0^\infty O(x)\, dF(x). \tag{S-31}$$

Thus, (S-29) follows from equations (S-30) and (S-31).

Finally, the upper semi-continuity and boundedness of $P$ imply that the map $dF \mapsto$

$\int_0^\infty P\,dF$ is upper semi-continuous. Thus, a solution $dF^*$ exists to the problem

$$\max_{dF\in\mathscr{C}} \int_0^\infty P(x)\,dF(x).$$

Thus, Theorem S.1 follows from (S-25) and (S-29). $\square$

*Proof of Lemma S.7:* We first develop a statistical result that will be used to prove Lemma S.7.

**Lemma S.12.** *Consider a set of distribution functions, $\{F_1, F_2, \cdots, F_n\}$. Let $F_{m:n}$ be the distribution function for the mth highest realization of the random variables associated with these distribution functions, i.e., $F_{1:n}$ is the distribution of the highest realization and $F_{n:n}$ is the distribution of the lowest. Define $F_{m:n} = 0$ if $m \le 0$ and $F_{m:n} = 1$ if $m \ge n$. The following statement must be true: if $F_{m:n}(x) = F_{m-1:n}(x)$, it must be that either $F_{m:n}(x) = F_{m-1:n}(x) = 1$ or $F_{m:n}(x) = F_{m-1:n}(x) = 0$.*

*Proof.* We prove by an induction argument. Let $S(n)$ be the general statement in the lemma. First, we show that $S(2)$ is true. To see this, given the facts that $F_{0:2}(x) = 0$, $F_{1:2}(x) = F_1(x)F_2(x)$, $F_{2:2}(x) = 1-(1-F_1(x))(1-F_2(x))$, and $F_{3:2}(x) = 1$, simply do the calculations:

(i) $0 = F_{3:2}(x) - F_{2:2}(x) = 1 - F_{2:2}(x)$

(ii) $0 = F_{2:2}(x) - F_{1:2}(x) = F_1(x)(1-F_2(x)) + F_2(x)(1-F_1(x))$

(iii) $0 = F_{1:2}(x) - F_{0:2}(x) = F_{1:2}(x) - 0.$

It is evident that $S(2)$ is satisfied in cases (i) and (iii). For case (ii), since the value of a distribution function is bounded by 0 and 1, it is clear that $F_{2:2}(x) - F_{1:2}(x) = 0$ either (a) when $F_1(x) = F_2(x) = 0$, in which case $F_{1:2}(x) = 0$ or (b) when $F_1(x) = F_2(x) = 1$, in which case $F_{1:2}(x) = 1$. Thus, $S(2)$ holds in all the cases.

Assume, for the purpose of induction, that $S(n-1)$ is true. We prove the lemma by showing that this implies $S(n)$. Note that, for any distribution function $j$ among the $n$ distribution functions,

$$F_{m:n}(x) = F_j(x)\left(F^{-j}_{m:n-1}(x) - F^{-j}_{m-1:n-1}(x)\right) + F^{-j}_{m-1:n-1}(x) \tag{S-32}$$

$$F_{m-1:n}(x) = F_j(x)\left(F^{-j}_{m-1:n-1}(x) - F^{-j}_{m-2:n-1}(x)\right) + F^{-j}_{m-2:n-1}(x), \tag{S-33}$$

in which the superscript $-j$ denotes the extraction of $j$ from the set. Subtracting (S-33) from (S-32) yields

$$\begin{aligned} F_{m:n}(x) - F_{m-1:n}(x) = & F_j(x)\left(F^{-j}_{m:n-1}(x) - F^{-j}_{m-1:n-1}(x)\right) \\ & + (1-F_j(x))\left(F^{-j}_{m-1:n-1}(x) - F^{-j}_{m-2:n-1}(x)\right). \end{aligned} \tag{S-34}$$

Suppose that, for some $j$, $F_j(x) \in (0,1)$ (otherwise $S(n)$ must hold). Since $F^{-j}_{m:n-1}(x) - F^{-j}_{m-1:n-1}(x)$ and $F^{-j}_{m-1:n-1}(x) - F^{-j}_{m-2:n-1}(x)$ are nonnegative and $F_j(x) \in (0,1)$, by (S-34), if $F_{m:n}(x) - F_{m-1:n}(x) = 0$, then

$$F^{-j}_{m:n-1}(x) - F^{-j}_{m-1:n-1}(x) = F^{-j}_{m-1:n-1}(x) - F^{-j}_{m-2:n-1}(x) = 0. \tag{S-35}$$

Then $S(n-1)$ implies that

$$\text{either} \quad F^{-j}_{m-1:n-1}(x) = F^{-j}_{m-2:n-1}(x) = 0 \quad \text{or} \quad F^{-j}_{m-1:n-1}(x) = F^{-j}_{m-2:n-1}(x) = 1. \tag{S-36}$$

Applying the conditions, (S-35) and (S-36), to equations (S-32) and (S-33) shows that

$$\text{either} \quad F_{m-1:n}(x) = F_{m:n}(x) = 0 \quad \text{or} \quad F_{m-1:n}(x) = F_{m:n}(x) = 1,$$

and the result is established. □

For any two distinct contestants $i$ and $j$, we must have

$$P_i(x) = F_j(x)F^{-ij}_{m:n-2}(x) + (1-F_j(x))F^{-ij}_{m-1:n-2}(x) \tag{S-37}$$
$$P_j(x) = F_i(x)F^{-ij}_{m:n-2}(x) + (1-F_i(x))F^{-ij}_{m-1:n-2}(x), \tag{S-38}$$

where $F^{-ij}_{m:n-2}$ is the probability of being at least the $m$th highest of the remaining contestants (with $i$ and $j$ extracted from the set) and $F^{-ij}_{m-1:n-2}$ is the probability of being at least the $(m-1)$th highest of the remaining contestants. Define $F^{-ij}_{m:n-2} = 1$ if $m \geq n-2$, $F^{-ij}_{m-1:n-2} = 1$ if $m-1 = n-2$, and $F^{-ij}_{m-1:n-2} = 0$ if $m-1 = 0$. Equation (S-37) means that, ignoring prize rationing in the case of ties, for $i$ to be one of the $m$ top performing contestants, $i$ must either weakly beat $j$ and be one of the top $m$ when competing against the rest or lose to $j$ and be one of the top $(m-1)$ when competing against the rest. The meaning of equation (S-38) is similar.

Subtracting (S-37) from (S-38) yields

$$P_j(x) - P_i(x) = (F_i(x) - F_j(x))\left(F^{-ij}_{m:n-2}(x) - F^{-ij}_{m-1:n-2}(x)\right). \tag{S-39}$$

Since $F^{-ij}_{m:n-2}(x) - F^{-ij}_{m-1:n-2}(x) \geq 0$, it is clear, by equation (S-39), that if $P_i(x) > P_j(x)$, we must have $F_i(x) < F_j(x)$. This establishes (i) in Lemma S.7.

Next, we show (ii) by way of contradiction. Suppose $P_j(x) = P_i(x) \in (0,1)$ and suppose, contrary to (ii), that $F_i(x) - F_j(x) \neq 0$. Then by equation (S-39), we must have $F^{-ij}_{m:n-2}(x) - F^{-ij}_{m-1:n-2}(x) = 0$. However, by Lemma S.12, $F^{-ij}_{m:n-2}(x) - F^{-ij}_{m-1:n-2}(x) = 0$ implies that either $F^{-ij}_{m:n-2}(x) = F^{-ij}_{m-1:n-2}(x) = 1$ or $F^{-ij}_{m:n-2}(x) = F^{-ij}_{m-1:n-2}(x) = 0$. This, by equation (S-38), further implies that $P_j(x)$ must equal either 0 or 1. This contradicts that $P_j(x) \in (0,1)$. The contradiction establishes (ii) in Lemma S.7.

Finally, (iii) follows immediately from (i) and (ii). $\square$