Winning ways: How rank-based incentives shape risk-taking decisions*

Dawei Fang† Changxia Ke‡ Greg Kubitz§ Yang Liu¶
Thomas Noe‖ Lionel Page**

6th June, 2024

Abstract: Risk-taking spurred by rank-based contest rewards can have enormous consequences, from developing breakthrough technologies in research competitions to hedge fund collapses engendered by risky bets aimed at raising league-table rankings. We develop a theoretical model of risk-taking that permits contestants to make arbitrary, mean-preserving changes in their random contest performance and use this framework to produce determinant predictions about the effect of contest structure on the statistical properties of contestant performance—including modality, tail behavior, dispersion, and skewness. In a laboratory experiment utilizing an interactive distribution builder to elicit subjects' risk-taking strategies, we confirm our main theoretical predictions. Our results order contest structures by the amount of risk-taking they induce in terms of the dispersion and skewness of contestant performances. Increasing the real-gain inequality of contest rewards or contest size increases performance dispersion; convexifying the rank-reward relationship or adding contestants to a contest with a fixed number of identical winner rewards increases both the dispersion and skewness of performance.

Keywords: contests, risk taking, dispersion, skewness, prize inequality
JEL codes: C72, C91, D74, D81

*We thank seminar audiences at Edinburgh, Gothenburg, Korea University, Maastricht, Melbourne, Monash and various conferences for helpful comments and suggestions. Fang gratefully acknowledges financial support from Vetenskapsrådet (grant 2022-01904) and the Jan Wallander and Tom Hedelius Foundation and Tore Browaldh Foundation (grant P22-0016). Any remaining errors are our responsibility.

†Department of Economics, University of Gothenburg. Email: dawei.fang@economics.gu.se.

‡School of Economics and Finance, Queensland University of Technology. Email: changxia.ke@qut.edu.au.

§School of Economics and Finance, Queensland University of Technology. Email: kubitz@qut.edu.au.

¶Faculty of Business and Economics, University of Melbourne. Email: yang.liu@unimelb.edu.au.

‖Saïd Business School & Balliol College, University of Oxford. Email: thomas.noe@sbs.ox.ac.uk.

**School of Economics, University of Queensland. Email: lionel.page@uq.edu.au.

1 Introduction

In many contexts, economic benefits are based on performance rankings. Funds flow into a mutual fund when its Morningstar ranking increases (Del Guercio and Tkac, 2008). The profitability and growth rate of a firm increase after it wins a product design contest (Guo, 2010). Applications swell after a law school's U.S. News and World Reports (USN) ranking rises (Sauder and Lancaster, 2006). Economic logic suggests that agents should respond to these ranking systems to increase their likelihood of receiving the benefits of higher ranks. Importantly, the welfare implications of these strategic responses depend not only on how they affect mean performance but also on how they shape other economically relevant statistical properties of performance.

Frequently, the risks created by the dispersion of individual performances can have significant welfare effects. For example, in financial markets, the volatility caused by the trading activities of rank-motivated fund managers can reduce the welfare of retail investors and destabilize financial markets (Strack, 2016). In R&D contests, where innovators submit product designs and only the best design will be developed, high-risk strategies employed by innovators increase the expected quality of the best design, thereby increasing welfare. In some contexts, the skewness of individual performances also has important welfare effects. For example, stock investors seem to prefer financial assets, such as mutual fund shares, that have positive skewness ("lottery-ticket assets") (see, e.g., Conrad et al., 2014; Agarwal et al., 2022). In populations where individuals are concerned with wealth or wage ranks, overall happiness is highest for populations with negatively skewed wealth distributions (Smith et al., 1989; Brown et al., 2008).

To examine these welfare-relevant properties, in this paper we develop a parsimonious model of competitions with rank-based rewards that encompasses a broad range of contest structures and places no exogenous constraints on risk-taking strategies. This model makes determinant predictions about the effects of contest structure (i.e., the allocation of rewards across performance ranks and the number of contestants participating in the contest) on risk taking. We then test these predictions in a laboratory experiment by considering the effects of various changes of contest structure on the dispersion and skewness of contestant performance. In the experiment, we generally observe statistically significant changes in contestant risk-taking behavior in the direction predicted by the model. We never observe statistically significant changes opposite to the model's predictions.

In our model, n contestants strategically choose the distribution of their random performance to compete for n prizes of nonidentical values. We restrict only the first moment (i.e., mean) of their performance while endogenizing all higher moments. Particularly, in our model, contestants can choose *any* non-negative performance distribution subject to a fixed mean. Assuming homogeneous contestants with the same mean performance, we derive in Theorem 1 the unique symmetric equilibrium of this game.[1] We show that contestants eschew risk-maximizing or risk-minimizing strategies. Instead, they always choose a performance distribution with a bounded interval support. We present a "closed-form" representation of the quantile function for the equilibrium performance distribution. We find that the distribution is positively skewed, with a decreasing probability density function (PDF) over the support, if the prize schedule is convex, and the distribution is negatively skewed, with an increasing PDF over the support, if the prize schedule is concave. In contrast to the assumptions of symmetric and/or unimodal performance distributions frequently made in the literature, we find that such distributions arise endogenously only in special cases.

We next turn to comparing the effects of different contest structures on risk taking. We consider four aspects of contest structure: prize inequality, prize convexity, contestant entry, and contest scale. Prize inequality and convexity order contests with the same number of contestants. Contestant entry and contest scaling capture the effects of increasing the number of contestants. We relate these aspects to two properties of the equilibrium performance distribution: dispersion and skewness.

First, we consider prize inequality. We show in Theorem 2 that contestants take higher risks by choosing more dispersed performance distributions whenever there is an increase in the inequality of the real gains associated with the contest prizes. A prize's real gain is defined as the value of the prize less the value of the lowest rank prize. Fixing the lowest rank prize, a transfer of value from a lower to a higher rank prize increases the inequality of real gains, making the contest more competitive. We find however that this link between competitiveness and risk taking does not hold if competitiveness is measured by inequality of (nominal) prizes. Reducing competitiveness can sometimes paradoxically increase risk taking. For instance, reducing the reward for best performance to subsidize the worst performer can encourage risk taking, by reducing "downside penalty."

Next, we show in Theorem 5 that *convexifying* the prize schedule increases both the dispersion and skewness of performance. Convexification increases the rate at which the slope of the prize schedule increases. This finding implies that, among contests with the same number of contestants but different prize schedules, winner-take-all (WTA) contests produce the most dispersed and most positively skewed performance distribution,

[1]In Section 7 and Appendix E, we consider an extension of our model where mean performance is endogenized through contestants' effort. We show that our theoretical results remain largely robust to this extension.

whereas elimination contests—where all contestants except the worst performer receive the same prize—generate the least dispersed and most negatively skewed performance distribution. This result is consistent with Kim (2019), which documents that the reduction in the funds' flow-ranking convexity over the 2000-2008 period was accompanied by reduced fund risk-taking.

Turning to the effects of changes in contest size, we analyze how increases in the number of contestants, keeping the prize structure more or less fixed, affect risk taking. First, consider contestant entry: adding more contestants and assigning the lowest rank prize to the additional ranks created by the addition of contestants. Entry transformations correspond to situations where positive real gains can only be captured by a fixed number of contestants, regardless of the total number of contestants, e.g., adding participants to a WTA prototype design contest. Theorem 4 and Proposition 6 show that adding entrants always increases performance dispersion and also increases skewness when the original contest featured only two distinct prize values. Thus, we predict that opening up WTA product design contests to more entrants will lead to more risk taking (e.g., adopting more innovative but untested engineering designs).

Lastly, consider the effect of increasing contest size through scaling up the contest, i.e., increasing the number of contestants or equivalently the number of performance ranks, without changing the inequality of real gains. For example, scale a 2-player/1-winner contest by tenfold, resulting in a 20-player/10-winner contest. In both contests, half of the contestants win, capturing the same positive real gain, while the other half lose, capturing zero real gain. Theorem 3 shows that the dispersion of performance distributions increases with scaling. Scaling does not affect the range of contestant performance. Instead, performance becomes more "clustered," i.e., more probability weight is concentrated around a few performance levels. Contest scaling is relevant when prizes are based on performance terciles, quartiles, quintiles, etc. For example, the USN law school ranking discussed above divides law schools outside the top 50 into terciles: second tier (best), third tier, and fourth tier (worst) based on their performance, i.e., scores on the USN metric. The question addressed by our scaling result in this context is how increasing the number of non-elite law schools affects their risk-taking behavior (e.g., fudging data reports to USN).

We tested various predictions from our model through a lab experiment using a Distribution Builder (DB). A DB is a graphical interface that permits subjects to explicitly construct performance distributions. The range of distributions they can construct is limited only by the model's constraint on mean performance and the discreteness resulting from the finite number of performance levels represented in the DB. We recover performance distributions directly from the distributions produced using the DB rather than trying to recover performance distributions from realized draws of the participants' performance distributions, which is problematic given the high dimensionality of the set of feasible distributions and the limited number of performance realizations produced in the experiment.

The experimental results, presented in Section 6, are largely consistent with the theoretical predictions. As predicted, participants chose positively skewed performance distributions when the prize schedule was highly convex and negatively skewed distributions when the prize schedule was concave. Increasing the convexity of the prize schedule increased both dispersion and skewness. The effects of entry into contests with only two distinct prize levels were also consistent with the model's predictions: adding new contestants without increasing the number of winners' prizes increased dispersion and skewness. Although we did not find evidence for scaling increasing performance dispersion, we did observe the predicted increase in performance clustering.

Related literature Our paper contributes, both theoretically and experimentally, to the literature on risk-taking strategies in contests where contestants compete for rank-based "prizes." The contestants' risk-taking strategies can naturally be represented as the choices of probability distributions over realized performance. When choosing their risk-taking strategy, they face a *capacity constraint* that limits their expected performance but can choose from a menu of *performance distributions*. This approach has been adopted to model fund manager competition (Taylor, 2003; Seel and Strack, 2013; Chen et al., 2015; Jin and Noe, 2022), R&D races (Dasgupta and Stiglitz, 1980; Klette and de Meza, 1986; Cabral, 2003), status contests (Robson, 1992; Becker et al., 2005; Ray and Robson, 2012), electoral campaigns (Myerson, 1993), promotion contests (Hvide, 2002; Hvide and Kristiansen, 2003; Goel and Thakor, 2008; Gilpatric, 2009; Fang and Noe, 2022), and sales contests (Gaba and Kalra, 1999; Gaba et al., 2004).

Earlier theoretical studies in this literature commonly restrict contestants' choices of performance distributions either to symmetric distributions (Klette and de Meza, 1986; Hvide, 2002; Gaba et al., 2004; Goel and Thakor, 2008; Kräkel, 2008; Gilpatric, 2009) or to mixtures of two exogenously specified distributions (Hvide and Kristiansen, 2003; Taylor, 2003; Kräkel and Sliwka, 2004; Nieken and Sliwka, 2010). In these settings, risk choice essentially boils down to variance choice. The variance-choice models typically predict that each contestant plays a "bang-bang" strategy by choosing either zero or maximum risk in equilibrium. For example, Gaba et al. (2004) consider a multi-winner contest where winners receive identical prizes. They show that, when contestants can choose any symmetric distribution with the same mean, contestants will take maximum risk if the winner proportion is less than one half and will take no risk if the winner proportion is greater than one half.[2] Thus, a change of the number of prizes and/or contestants has no effect on risk taking behavior unless it causes the winner proportion to cross the one-half threshold.

In sharp contrast, in our setting where contestants are not restricted to symmetric distributions, (1) equilibrium solutions are always interior, (2) increased inequality of real gains and increased contest size or scale, no matter how small, always increase risk taking, and (3) in the case of identical winner prizes, the one-half winner proportion is no longer the boundary between regions of no risk and maximum risk; rather, it is the boundary between regions of negatively skewed and positively skewed performance.

More recent theoretical studies on risk-taking strategies in contests allow contestants to choose skewed distributions, making the analysis of performance skewness possible. However, unlike our paper, these studies do not examine the effect of contest structure on skewness but focus on the inefficiencies produced by rank-motivated risk taking (i.e., performance dispersion). Seel and Strack (2013) and Strack (2016) show that competition between fund managers leads to excessive managerial risk taking in financial markets. Fang and Noe (2022) show that soft rank-based promotion policies, through reducing the effects of risk taking, can increase the overall correlation between selection and ability and thus improve selection efficiency. Jin and Noe (2022) find that, while rank-based rewards and high-powered absolute performance rewards both engender risk taking by financial agents, combining these two types of rewards mollifies the risk-taking incentives produced by each single type. Importantly, whether risk taking reduces welfare is context dependent and, in some contexts, welfare can be affected by performance skewness. We do not restrict our analysis to performance dispersion or to any specific context of contest. Instead, we provide comprehensive comparative statics that relate contest structure to the shape of the equilibrium performance distribution. These characterization results can be applied to explore the implications of rank competitions on risk-taking strategies and economic outcomes.

Our results on skewness, modality, and tail behavior of performance can be applied to the financial-market frameworks developed by Seel and Strack (2013) and Strack (2016) to draw predictions for return distributions of funds with tournament incentives. Strack (2016) argues that the increased unsystematic dispersion generated by tournament incentives of mutual funds lowers investor welfare because of investor risk aversion. Our results point to an important qualification for this argument. Our characterizations of the effect of prize schedule convexity on performance skewness show

[2] To define maximum risk, Gaba et al. (2004) impose exogenous upper and lower bounds on the support of admissible performance distributions.

that the empirically documented convex reward schedules in mutual fund tournaments (Chevalier and Ellison, 1997; Sirri and Tufano, 1998), induce highly dispersed performance, as well as highly skewed performance. Moreover, as Mitton and Vorkink (2007) demonstrate, investors with high skewness preference are precisely the investors who are likely to be under-diversified and thus exposed to the idiosyncratic risk engendered by mutual fund tournaments. Thus, the welfare costs of mutual fund tournaments are likely to be primarily borne by investors who also receive a skewness benefit from such tournaments.

By employing a DB consistent with our theoretical framework, our experimental results not only generally support our theoretical predictions but also contribute new insights to the empirical and experimental literature on risk-taking in contests. This body of literature documents various factors that shape risk preferences, such as, among others, the number of contestants (Eriksen and Kvaløy, 2017), the proportion of winners (Gaba and Kalra, 1999; Fang et al., 2017), the distribution of prizes (Andersson et al., 2020), relative interim performance (Dijk et al., 2014), the frequency of feedback on peer performance (Eriksen and Kvaløy, 2014), financial professionals' intrinsic preferences for high rankings (Kirchler et al., 2018), and last-place aversion (Kuziemko et al., 2014).

Importantly, most existing experimental studies on risk-taking in contests adopt designs in which (i) performance is drawn from a symmetric distribution, with subjects choosing the variance, or (ii) performance is a weighted average of a safe and a risky component, with subjects choosing the exposure (i.e., the weight) to the risky component. These designs essentially reduce risk choice to variance choice. Two notable exceptions are Dijk et al. (2014) and Embrey et al. (2024). Dijk et al. (2014) allow subjects, acting as competing fund managers, to choose between several risky assets which vary in return skewness. However, because their subjects are given only a small set of different risky assets to choose from, the control over skewness is very limited. Embrey et al. (2024) allow competing subjects to decide when to stop a stochastic process, with the stopped value representing performance. Although the direct control is over stopping time, different stopping strategies can induce different (possibly skewed) performance distributions. When the process is a Brownian motion absorbed at zero with no drift, the set of performance distributions that can be produced by dynamic stopping decisions is the same as the set of admissible performance distributions used in our static setting (Seel and Strack, 2013). However, Dertwinkel-Kalt and Frey (2024) find that when stopping a stochastic process, subjects exhibit time-inconsistent behavior in line with the dynamic implications of salience theory (Bordalo et al., 2012). Embrey et al. (2024) also observe deviations between predicted and actual behavior, partly due to subjects' enjoyment from engaging with the process when stakes are involved. In contrast, our use of a DB enables subjects to explicitly construct their preferred performance distributions, without exposing them to the potential behavioral biases associated with dynamic risk decisions. Additionally, using a DB allows us to observe the performance distribution each subject chooses each round, which could not be inferred merely from a subject's choice of stopping time in a task involving stopping a stochastic process.

2 Risk-taking contests

2.1 The contest structure

Consider a contest with $n \geq 2$ contestants and n prizes. Each prize is identically valued by the contestants. Let v_i be the common value of the ith prize. We order the prizes so that $v_1 \leq \cdots \leq v_n$. To make it a contest, the n prizes must not be identical. Thus, we assume that $v_1 < v_n$. We allow a prize to have negative or zero value.

Prize allocation is based on performance in the contest. Contestants have a fixed expected performance $\mu > 0$, representing their *capacity*. Although their mean performance is fixed by their capacity, they can choose the "randomness" of their performance by undertaking risky actions. We assume that each contestant can choose any distribution of non-negative performance subject to the *capacity constraint* on mean performance.[3]

Contestants simultaneously choose performance distributions. Each contestant's realized performance is independently drawn from the performance distribution she chooses. The contestant with the highest realized performance wins the nth prize, v_n, the one with the next highest realized performance wins the $(n-1)$th prize, v_{n-1}, and so on. Ties are resolved arbitrarily. A contestant's payoff equals the prize she receives.

2.2 The equilibrium

We focus our analysis on symmetric equilibria, where all contestants choose the same performance distribution, F. In any symmetric equilibrium, all contestants face the

[3]We require performance to be non-negative. Non-negativity *per se* is not important for our analysis. What is important is that we require performance to be bounded below. This is because, in our model, high performance has only a shadow cost through the capacity constraint, and the lower bound on performance is required to make the shadow cost nonzero. We normalize the lower bound to zero. The non-negativity constraint is consistent with many real world contests. For example, in fund manager competitions, due to limited liability, the worst possible performance of a fund is zero gross return. In student examinations, the worst possible performance is getting zero points.

same *contest payoff function*, P, which maps a contestant's realized performance, x, to her expected payoff, $P(x)$, and is determined endogenously through the $n-1$ rivals' performance distributions, F. Our first lemma shows that, in any symmetric equilibrium, P must be continuous and increase over and only over a bounded interval support, with zero as the lower bound. This result follows from Claim 1 in Fang and Noe (2022), whose proof can be found in the Online Appendix A of Fang and Noe (2022).[4]

Lemma 1. *In any symmetric equilibrium, the contest payoff function, P, is (a) continuous and (b) increases over and only over the interval,* $[0, \bar{x}]$*, where* $\bar{x} < \infty$ *is endogenous.*

Part (a) is due to the requirement that, in any symmetric equilibrium, contestants must not place any point mass. Placing point mass on the same performance level, say x^o, would create a positive probability that all contestants tie at x^o. In this case, no matter what tie-breaking rule is used, there would exist at least one contestant who could be 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 ε to zero while, for all positive ε, no matter how small, the transfer would generate a gain that is bounded below by a strictly positive number. Hence, in any symmetric equilibrium, contestants must place no point mass, ensuring the continuity of P.

To understand part (b), note that, in any symmetric equilibrium, P cannot stay flat over an interval, say (x', x''), and then start to increase at x''. This is because, for x'' to be a point of increase for P, x'' must be in the support of F. However, given that P is continuous, if P stayed flat over (x', x''), we would have $P(x'') = P(x')$, in which case a contestant could obtain the same expected payoff simply by having the lower performance level, x', rather than the higher, x''. Given that the capacity constraint imposes a shadow price on performance, it would thus not be optimal for a contestant to have x'' in the support of her performance distribution, F, a contradiction.[5] This contradiction implies that P must increase over and only over an interval with a lower bound of 0. The upper bound of this interval is finite because performance has a positive shadow price, while the largest prize is bounded. Therefore, it is not optimal for a contestant to place probability weight on very high performance levels.

Lemma 1 in Fang and Noe (2022) implies that, in any symmetric equilibrium, all

---

[4]Although, in Fang and Noe (2022), each contestant has high capacity with probability θ and low capacity with probability $1-\theta$, their proof applies to the case where $\theta = 0$, in which all contestants have the same capacity. The prize schedule considered in Fang and Noe (2022) is more restrictive; they consider only contests with two distinct prize values. However, the proof of their Claim 1 does not rely on this restriction.

[5]Although performance imposes no direct cost in risk-taking contests, performance uses up capacity and the capacity constraint thus indirectly imposes a shadow price on performance.

performance/contest-payoff pairs, $(x, P(x))$, with x in the support of F, are collinear.[6] This result, combined with our Lemma 1, implies that P can be expressed as

$$P(x) = \alpha + \beta x, \quad x \in [0, \bar{x}]. \tag{1}$$

To pin down the endogenous parameters, α, β, and $\bar{x}$, in equation (1), note that, in a symmetric equilibrium, because no contestant places any point mass, a contestant will receive the lowest rank and thus v_1 if her realized performance is 0, and will receive the highest rank and thus v_n if her realized performance is $\bar{x}$. Hence, $P(0) = v_1$ and $P(\bar{x}) = v_n$. Also note that P increases only over points that are in the support of F. Thus, if X represents the equilibrium random performance of a contestant, given that P is linear over the support of X, this contestant's expected payoff satisfies $\mathbb{E}[P(X)] = P(\mathbb{E}[X]) = P(\mu)$. In a symmetric equilibrium, this is equal to the average prize: $P(\mu) = \left(\sum_{i=1}^{n} v_i\right)/n$. We, hence, obtain three equations that can be used to solve for α, β, and $\bar{x}$ and to pin down P.

Proposition 1 (Equilibrium contest payoff function)**.** *In any symmetric equilibrium,*

a. each contestant chooses a performance distribution supported by $[0, \bar{x}]$*, where*

$$\bar{x} = \frac{\mu\, n\,(v_n - v_1)}{\sum_{i=1}^{n}(v_i - v_1)}; \tag{2}$$

b. the contest payoff function faced by each contestant, P, is given by

$$P(x) = v_1 + (v_n - v_1)\,\frac{x}{\bar{x}}, \quad x \in [0, \bar{x}]. \tag{3}$$

Proposition 1 enables us to derive the equilibrium performance distribution, F. Under the rank-based prize allocation, a contestant will obtain v_{i+1} if she outperforms exactly i out of her $n-1$ competitors. The probability of this event equals $\binom{n-1}{i} F^i (1-F)^{n-1-i}$ when each competitor plays the same continuous distribution, F. Thus, the given contestant's contest payoff function, P, must satisfy that

$$P(x) = \sum_{i=0}^{n-1} v_{i+1} \binom{n-1}{i} F(x)^i (1 - F(x))^{n-1-i}. \tag{4}$$

Combining equations (3) and (4) gives the equilibrium performance distribution.[7]

Theorem 1 (Equilibrium)**.** *There exists a unique symmetric equilibrium. If* F_v *represents the equilibrium performance distribution associated with prize schedule* v*, then*

[6] More specifically, Fang and Noe (2022) consider two possible types of contestants, distinguished by capacity levels and with contestant type being private information. Their Lemma 1 shows that, in any symmetric equilibrium, for any given type, all performance/contest-payoff pairs, $(x, P(x))$, with x in the support of this given type's performance distribution, are collinear.

[7] The derived distribution, F, is clearly a best response to P, because the linearity of P over $[0, \bar{x}]$ ensures that any performance distribution supported by $[0, \bar{x}]$ is a best response.

$\operatorname{supp}(F_v) = [0, \mu n(v_n - v_1)/V]$, *where*

$$V = \sum_{i=1}^{n} (v_i - v_1), \tag{5}$$

and, over $\operatorname{supp}(F_v)$, F_v *is uniquely determined by*

$$\sum_{i=0}^{n-1} (v_{i+1} - v_1) \binom{n-1}{i} F_v(x)^i (1 - F_v(x))^{n-1-i} = \frac{V}{\mu n} x. \tag{6}$$

Equation (6) implies a "closed-form" representation of the quantile function associated with the equilibrium performance distribution. Because we have no closed-form representation, in general, for the equilibrium performance distribution, the quantile representation will be critical for our analysis.

Corollary 1 (Quantile representation)**.** *Define* V *by equation* (5). *If* Q_v *represents the quantile function associated with* F_v *expressed in equation* (6), *then* Q_v *satisfies*

$$Q_v(p) = \frac{\mu n}{V} \sum_{i=0}^{n-1} (v_{i+1} - v_1) \binom{n-1}{i} p^i (1-p)^{n-1-i}. \tag{7}$$

The quantile representation reveals a number of properties of the equilibrium performance distribution. First, note that, by equation (7), the effect of the prize schedule on performance is entirely determined by terms of the form $(v_i - v_1)/V$. The numerator, $v_i - v_1$, is the value of the ith prize in excess of the smallest prize. Since each contestant receives at least the value of v_1, what really matters for their incentives is the prize values in excess of v_1. Thus, we call $v_i - v_1$ the *real gain* of the ith prize. By equation (5), V thus represents the *total real gains* offered by the contest and, hence, $(v_i - v_1)/V$ represents the fraction of the total real gains captured by the ith prize. Such a fraction is invariant to increasing affine transformations of the prize schedule. Next, note that the equilibrium random performance is equal in distribution to $Q_v(\tilde{U})$, where $\tilde{U}$ is uniformly distributed on $[0,1]$. This implies, upon inspecting equation (7), that a change in capacity, μ, acts simply as a scale transformation of the equilibrium performance distribution. Finally, the minimum and the maximum performance levels are given by $Q_v(0) = 0$ and $Q_v(1) = \mu n (v_n - v_1)/V$, respectively. These observations imply the following result.

Lemma 2 (Basic properties)**.** *The equilibrium random performance is*

a. invariant under increasing affine transformations of the prize schedule,

b. proportional to capacity, μ,

c. and distributed continuously over its support $[0, \mu n (v_n - v_1)/V]$.

Part (c) of Lemma 2 has an important implication. Up to the scale factor, μ, the support of the performance distribution is determined entirely by the number of contestants, the total real gains, and the smallest and largest prizes. In fact, if the scale factor, μ, is fixed at 1, the upper bound of the support equals

$$\frac{(v_n - v_1)/V}{1/n}.$$

Thus, the upper bound is simply the fraction of total real gains allocated to the largest prize, $(v_n - v_1)/V$, relative to the "fair share" fraction, $1/n$, the fraction of real gains that each contestant would have received if real gains had been redistributed equally across all contestants. Because the lower bound of the support of the performance distribution is always 0, increasing the relative attractiveness of the largest prize stretches the support of the equilibrium performance distribution. Relative attractiveness can be increased either by increasing the fraction of real gains captured by the largest prize or by reducing each contestant's fair share of real gains by increasing the number of contestants. We will investigate both effects in the sequel.

3 The shape of the performance distribution

In this section, we investigate the relation between the prize schedule and the shape of the equilibrium performance distribution. Key to our analysis are the following representations of the quantile density and the derivative of the quantile density function for the equilibrium performance distribution. We derive these representations by differentiating the quantile function in equation (7).

Corollary 2 (Quantile density and its derivative)**.** *If q_v represents the quantile density function associated with Q_v expressed in equation* (7)*, then q_v satisfies*

$$q_v(p) = \frac{\mu\, n\,(n-1)}{V} \left(\sum_{i=0}^{n-2} \Delta v_{i+1} \binom{n-2}{i} p^i\,(1-p)^{n-2-i} \right), \tag{8}$$

where V is defined by equation (5) *and Δv_i by*

$$\Delta v_i = v_{i+1} - v_i. \tag{9}$$

If q_v' represents the derivative of q_v, then for $n \geq 3$ (otherwise, $q_v' = 0$), q_v' satisfies

$$q_v'(p) = \frac{\mu\, n\,(n-1)}{V} \left(\sum_{i=0}^{n-3} (\Delta v_{i+2} - \Delta v_{i+1}) \binom{n-2}{i} (n-2-i) p^i (1-p)^{n-3-i} \right). \tag{10}$$

Equation (8) shows that the quantile density function of the equilibrium performance distribution is, up to a scaling factor, a convex combination of the *prize differences*, $(\Delta v_1, \ldots, \Delta v_{n-1})$. For $n = 2$, the quantile density function is $q_v(p) = 2\mu$ for $p \in (0,1)$, in which case the quantile function is linear. For $n \geq 3$, equation (10) implies that the derivative of the quantile density and thus the convexity of the quantile function are determined by the *second differences*, $(\Delta v_2 - \Delta v_1, \ldots, \Delta v_{n-1} - \Delta v_{n-2})$.

3.1 Contests with monotone prize differences

Corollary 2 enables us to fully characterize the global behavior of the equilibrium probability density function (PDF) in contests with monotone prize differences. When $n = 2$, there is only one prize difference, and equation (8) implies a linear quantile function and thus a uniform performance distribution.

When $n \geq 3$, then if the prize-difference sequence $(\Delta v_1, \ldots, \Delta v_{n-1})$ is monotone, the second-difference sequence $(\Delta v_2 - \Delta v_1, \ldots, \Delta v_{n-1} - \Delta v_{n-2})$ remains non-sign-changing. Such contests fall into three categories:

(1) *Convex contests*: In convex contests, all second differences in the prize schedule are non-negative, and at least one second difference is positive. This category includes, as a special case, *WTA contests*, where only the top performer receives a positive real gain.

(2) *Concave contests*: In concave contests, all second differences are nonpositive, and at least one second difference is negative. An example of concave contests is *elimination contests*, where all contestants, except the lowest performer, receive the same prize.

(3) *Linear contests*: In linear contests, all second differences are zero.

Equation (10) implies that, for $p \in (0,1)$, the derivative of the quantile density function, q_v', is positive in convex contests, negative in concave contests, and zero in linear contests. Thus, the quantile density function, q_v, is increasing in convex contests, decreasing in concave contests, and constant in linear contests. Because, by the inverse function theorem, the direction of monotonicity of the PDF of the performance distribution is opposite to its quantile density's, the PDF is, over its support, decreasing in convex contests, increasing in concave contests, and constant in linear contests. We thus obtain the following result.

Proposition 2 (PDF in contests with monotone prize differences)**.** *In convex (concave) (linear) contests, the PDF of the equilibrium performance distribution is decreasing (increasing) (constant) over the support and thus the performance distribution is positively skewed (negatively skewed) (uniform).*

Proposition 2 implies that performance features positive skewness in "elite competitions," where the relation between rank and reward is commonly observed to be convex. A real-world example of such a competition is the mutual fund tournament.

Many studies identify a convex relation between a mutual fund's performance ranking and its future capital inflows (Chevalier and Ellison, 1997; Sirri and Tufano, 1998; Kaniel and Parham, 2017). Given that fees charged by mutual funds are tied to assets under management, and fund managers often receive asset-based compensation (Ma et al., 2019), increased capital inflows translate to higher profits for both the funds and the managers. Proposition 2 implies that this competition for fund flows induces fund managers to take positively skewed idiosyncratic risk, consistent with the findings in Wagner and Winter (2013). In contrast, Chevalier and Ellison (1999) find that, for young mutual fund managers, the probability of job termination is a convex function of relative performance. Since job termination can be considered as a negative reward, a convex termination-probability function implies a concave prize schedule. Proposition 2 thus predicts that, in contrast to senior managers motivated by the competition for external capital flows, young managers, motivated by the internal competition for employment, will opt for portfolios with negatively skewed idiosyncratic risk.

3.2 Contests with non-monotone prize differences

When the prize-difference sequence $(\Delta v_1, \ldots, \Delta v_{n-1})$ is non-monotone, the second-difference sequence $(\Delta v_2 - \Delta v_1, \ldots, \Delta v_{n-1} - \Delta v_{n-2})$ will have at least one sign change. To examine the global behavior of the equilibrium performance distribution in such a case, we can take out the common factor $(1-p)^{n-3}$ from the summand in equation (10). Then the derivative of the quantile density function can be expressed as

$$q_v'(p) = \frac{\mu\, n(n-1)}{V}(1-p)^{n-3}\left(\sum_{i=0}^{n-3} a_i \left(\frac{p}{1-p}\right)^i\right), \quad \text{where} \tag{11}$$

$$a_i = (\Delta v_{i+2} - \Delta v_{i+1})\binom{n-2}{i}(n-2-i). \tag{12}$$

It is clear from equation (11) that the sign of q_v' is given by the sign of the polynomial in $p/(1-p)$ with $(a_i)_{i=0}^{n-3}$ being the coefficient sequence. Since, for $p \in (0,1)$, the range of the function $p \hookrightarrow p/(1-p)$ is $(0,\infty)$, the number of times that q_v' changes its sign equals the number of times that the polynomial, $\sum_{i=0}^{n-3} a_i y^i$, changes its sign on the positive real line. Descartes' Rule of Signs shows that the number of positive roots of the polynomial, $\sum_{i=0}^{n-3} a_i y^i$, equals the number of sign changes in the sequence $(a_i)_{i=0}^{n-3}$ of coefficients or less than it by an even number (multiple roots of the same value are counted separately).[8] Thus, if the number of sign changes in the sequence $(a_i)_{i=0}^{n-3}$ of

[8]The number of sign changes for a sequence is determined as follows: consider a sequence, $(a_1, a_2 \ldots a_{n-3})$, of coefficients; replace every positive element by $+$ and every negative element by $-$, dropping all 0 coefficients. Call the resulting sequence the sign sequence. The sequence has k sign changes if there are k places in the sign sequence where the sign differs from the following sign.

coefficients equals 1, in which case, by equation (12), the number of sign changes in the sequence of second differences, $(\Delta v_{i+2} - \Delta v_{i+1})_{i=0}^{n-3}$, is 1, Descartes' Rule of Signs implies that the number of sign changes of $\sum_{i=0}^{n-3} a_i y^i$ is exactly 1. This implies, by equations (11) and (12), that the derivative of the quantile density function, q_v', changes sign exactly once. This property leads to the following characterization of the global behavior of the equilibrium PDF when the second differences of the prize schedule change sign exactly once, i.e., when the prize-difference sequence is non-monotone and either quasiconvex or quasiconcave.

Proposition 3 (PDF in contests with non-monotone and quasiconvex/quasiconcave prize differences)**.** *If the prize-difference sequence is non-monotone and quasiconvex (quasiconcave), then the equilibrium performance distribution is unimodal (has a U-shaped PDF).*

When the prize differences are non-monotone and quasiconvex, the prize schedule is "concave-convex." Proposition 3 shows that this type of prize schedule leads to a unimodal performance distribution. Conversely, when the prize differences are non-monotone and quasiconcave, i.e., when the prize schedule is "convex-concave," the equilibrium PDF becomes U-shaped. To understand Proposition 3, note that equation (10) implies that, for p sufficiently small, the sign of the derivative of the quantile density, $q_v'(p)$, must equal the sign of the first nonzero element in the sequence of second differences $(\Delta v_2 - \Delta v_1, \dots, \Delta v_{n-1} - \Delta v_{n-2})$. Thus, if the first nonzero second difference is negative, i.e., if the prize schedule exhibits concavity around the lower end, the quantile density will be decreasing in a neighborhood of 0. Because, as has been argued above, when the sequence $(\Delta v_2 - \Delta v_1, \dots, \Delta v_{n-1} - \Delta v_{n-2})$ has exactly one sign change, the derivative of the quantile density function, q_v', changes sign exactly once, a "concave-convex" prize schedule will lead to a quantile density function that changes its direction of monotonicity exactly once. Thus, the quantile density function must be U-shaped, implying an inverse U-shaped PDF and hence a unimodal distribution. Analogously, a "convex-concave" prize schedule will result in an inverse U-shaped quantile density function and hence a U-shaped PDF. An example of a "convex-concave" prize schedule is the one under which the top m performers receive the same "winner prize" and the remaining $n-m$ performers receive the same "loser prize" with $1 < m < n-1$. Proposition 3 implies that this type of prize schedule will induce a U-shaped PDF.[9]

When the second-difference sequence $(\Delta v_2 - \Delta v_1, \dots, \Delta v_{n-1} - \Delta v_{n-2})$ has more than one sign change, the PDF of the equilibrium performance distribution may change the direction of monotonicity multiple times. While determining the number of modes and

[9]By Proposition 2, over the support, the PDF is decreasing for $m = 1$ and increasing for $m = n-1$.

anti-modes of the PDF in this case can be challenging, we are able to sharply characterize the tail behavior of the PDF. We relegate this tail-behavior characterization to Appendix C, which shows that the PDF will be increasing (decreasing) in the lower end of the support if the prize schedule exhibits concavity (convexity) around the lower end and will be increasing (decreasing) in the upper end of the support if the prize schedule exhibits concavity (convexity) around the upper end. This result implies that, to endogenize a unimodal distribution, the prize schedule has to exhibit concavity around the lower end and convexity around the upper end, which requires $v_1 < v_2$, $v_{n-1} < v_n$, and some change in convexity, a condition that is violated by many commonly encountered prize schedules.[10]

When does the equilibrium performance distribution have a symmetric PDF? The answer is provided by the next result.

Proposition 4 (Symmetric distribution)**.** *The PDF of the equilibrium performance distribution is symmetric about its mean if and only if the prize differences are symmetric, i.e.,*

$$\Delta v_i = \Delta v_{n-i} \quad \forall i \in \{1, \dots, n-1\}, \tag{13}$$

where Δv_i *is defined in equation* (9)*. Condition* (13) *is equivalent to* $v_i + v_{n+1-i}$ *being a constant for all* $i \in \{1, \dots, n\}$*.*

Intuitively, the equilibrium performance distributions of the contestants must collectively produce, over their support, a linear contest payoff function for each contestant, implying a constant marginal gain of performance over the support of the distribution. The marginal gain from choosing a slightly higher performance level depends on both the increase in the marginal probability of attaining a higher rank, determined by the performance PDF, and the payoff increase associated with a higher rank, measured by the prize differences. Hence, to maintain constant marginal gains at all performance levels over the support of the distribution, the marginal increase in the probability of attaining a rank increase implied by the PDF must be inversely proportional to the associated prize increase implied by the prize differences. Thus, the performance PDF is symmetric if and only if the prize differences are symmetric, a condition that is equivalent to the ith highest rank and the ith lowest rank prizes summing to a constant for all i. Our findings imply that, despite the ubiquity of the assumption of symmetric and/or unimodal performance distributions, such distributions will arise endogenously only in very special cases.

[10]For example, convex, concave, and linear contests never induce unimodal performance distributions. Neither do contests with only two distinct prize values. Contests with three distinct prize values induce a unimodal distribution if and only if $v_1 < v_2 = \ldots = v_{n-1} < v_n$ with $n \geq 4$.

4 Dispersion and skewness of performance

In this section, we make a comprehensive analysis of how a change in the contest structure impacts the dispersion and skewness of the equilibrium performance distribution.

4.1 Dispersion

We investigate dispersion by employing the standard *convex order* relation. Dominance in convex order, which is sometimes referred to as a *mean-preserving increase in risk*, is defined below.

Definition 1. Let F and G be two CDFs. F is more dispersed than G in the sense of convex order if, for all convex functions, $w : \Re_+ \to \Re$

$$\int w(x)\, dF(x) \geq \int w(x)\, dG(x).$$

As the following lemma reveals, using convex order in the dispersion analysis aids in unveiling the welfare implications of altering the contest structure.

Lemma 3 (Convex order and welfare)**.** *Let $w = w(x_1, \ldots, x_n) : \Re^n_+ \to \Re$ represent a social welfare function, where x_i represents the realized performance of contestant i for all $i \in \{1, \ldots, n\}$. Suppose F and G are two performance distributions and suppose F is more dispersed than G in the sense of convex order. Then*

a. if w is convex in every argument, social welfare is higher if each of the n contestants' performance is independently drawn from F than from G;

b. if w is concave in every argument, social welfare is higher if each of the n contestants' performance is independently drawn from G than from F.

Many situations exist when welfare is concave in each contestant's performance. Examples include competition in delegated money management when investors are risk averse or contests when the social planner has egalitarian Rawlsian preferences, i.e., $w = \min[x_1, \ldots, x_n]$. In contrast, if social welfare is a convex function of the highest performance, such as in the R&D races *à la* Dasgupta and Stiglitz (1980), or equals the sum of a given number of highest performances, such as in a poetry competition in which only the k best submitted poems will be published in the anthology, the social welfare function will be convex in each contestant's performance.[11]

[11] In Dasgupta and Stiglitz (1980), contestants compete in innovation speed by each choosing a distribution of the discovery date of the innovation. Because society benefits from only the quickest innovation and because society discounts the future, social welfare is convex in the highest performance (the fastest innovation speed). More specifically, social welfare essentially equals $k\exp(r\max[x_1, \ldots, x_n])$, where $k > 0$ and $r > 0$ are constants and x_i represents innovation speed of contestant $i \in \{1, \ldots, n\}$. This welfare function is a convex transformation of the function, $\max[x_1, \ldots, x_n]$, a function convex in every argument. Thus, this welfare function is convex in every argument.

Note that when the compared performance distributions have the same mean, as is the case when the compared contests have equal contestant capacities, being more dispersed in the sense of convex order is equivalent to being second-order stochastically dominated. Henceforth, for expositional convenience, we will simply say that one distribution is *more dispersed* than another if it dominates the other in convex order.

4.1.1 The effect of prize schedule on dispersion

With contest size, n, fixed, what type of prize schedule change increases performance dispersion? As discussed in Section 2.2, equilibrium risk taking is determined by the real gains of prizes, where for any prize vector $v = (v_1, v_2, \ldots, v_n)$, the real gain of the ith prize is $v_i - v_1$ and $V = \sum_{i=1}^{n}(v_i - v_1)$ represents total real gains. In what follows, we show that performance dispersion increases when the real gains of prizes become more unequal. Our definition of real gain inequality follows from the Lorenz order, a simple, rich order frequently used in the economics literature to measure inequality.

Definition 2 (Real gain inequality)**.** Let v and $\hat{v}$ be two prize vectors, each containing n elements. Let V and $\hat{V}$ be the total real gains associated with v and $\hat{v}$, respectively. Then $\hat{v}$ has (strictly) higher real gain inequality than v if

$$\sum_{i=1}^{k}\left(\frac{\hat{v}_i - \hat{v}_1}{\hat{V}}\right) \leq \sum_{i=1}^{k}\left(\frac{v_i - v_1}{V}\right) \quad \forall k \in \{1, \ldots, n\} \text{ (with at least one strict inequality).}$$

In the above definition, $(v_i - v_1)/V$ represents the fraction of total real gains allocated to the ith rank. Thus, $\sum_{i=1}^{k}(v_i - v_1)/V$ represents the fraction of total real gains allocated to the k lowest ranks. If the condition in Definition 2 holds, then for any k, the k worst performers capture a lower share of total real gains under $\hat{v}$ than under v, in which case the real gain allocation is more tilted toward best performers under $\hat{v}$ than under v. Consistent with our definition of real gain inequality, increasing the highest-rank prize increases real gain inequality; increasing the lowest-rank prize also increases real gain inequality, provided that this prize remains to be the smallest after the increase. If we fix the lowest-rank prize, then a simple transformation that increases real gain inequality is to transfer value from a lower-rank prize to a high-rank prize while ensuring that the sequence of prizes is still nondecreasing in rank.

The following theorem shows that increased real gain inequality increases performance dispersion.

Theorem 2 (Real gain inequality effects)**.** *Consider two prize vectors, v and $\hat{v}$, each containing n elements. Let F and $\hat{F}$ be the equilibrium performance distributions associated with v and $\hat{v}$, respectively. If $\hat{v}$ has a strictly higher real gain inequality, then $\hat{F} \neq F$ and $\hat{F}$ is more dispersed than F.*

The dependence of performance dispersion on real gain inequality immediately implies that, fixing the number of contestants, performance dispersion is maximized in WTA contests and minimized in elimination contests.

Corollary 3 (Maximal and minimal dispersion)**.** *With the number of contestants, n, fixed, the equilibrium performance distribution is most dispersed in WTA contests, i.e., contests with* $v_1 = \cdots = v_{n-1} < v_n$*, and least dispersed in elimination contests, i.e., contests with* $v_1 < v_2 = \cdots = v_n$*.*

Corollary 3 implies that, for contests in which social welfare is convex in every contestant's performance (e.g., R&D contests), rewards in excess of the minimum reward should be concentrated on the best performer. As discussed in Section 2.2, the equilibrium performance distribution always has a lower bound of 0, while the upper bound is proportional to the fraction of total real gains allocated to the highest rank. Thus, by maximizing (minimizing) the fraction of total real gains captured by the highest rank, WTA (elimination) contests also maximize (minimize) the range of the equilibrium performance distribution.[12]

Theorem 2 offers an important caveat to the common belief that reducing the competitiveness of a prize schedule reduces risk taking. This caveat arises because real gain inequality and nominal prize inequality can move in opposite directions. For example, consider a contest involving four contestants with a prize vector $(1,2,3,4)$. Suppose we make the prize schedule less competitive by transferring a value of $3/4$ from the highest-rank prize to the lowest-rank prize, resulting in a new prize vector $(7/4,2,3,13/4)$. Clearly, such a transfer makes nominal prizes more equal. However, this transfer actually makes real gains more unequal, leading to an increase in risk taking.[13]

Theorem 2 has two interesting implications on the use of mixed incentive schemes, often referred to as "carrot-and-stick" schemes, where top performers are rewarded and bottom performers are punished. First, Theorem 2 implies that increasing the rewards for the best performer and/or reducing the punishments on the worst performer increases performance dispersion. This aligns with the intuition that individuals take higher risks when upside rewards become larger and/or when downside penalties become smaller. Second, perhaps surprisingly, adding more "sticks," i.e., punishing a larger group of lowest-ranked contestants without changing the total number of contestants, has a non-monotone effect on performance dispersion: dispersion decreases when the first stick

[12] Of course, if increased real gain inequality is not accompanied by an increase in the fraction of total real gains captured by the highest rank, the performance distribution will be more dispersed without any change in its range.

[13] The fractions of real gains change from $(0,1/6,1/3,1/2)$ before the transfer to $(0,1/12,5/12,1/2)$ after the transfer. This adjustment increases the fraction of real gains allocated to the second highest rank while reducing the fraction allocated to the third highest rank. Consequently, real gain inequality increases.

is introduced but gradually increases when more sticks are introduced. This occurs because when the first stick is introduced, the real gains from each non-minimal prize increase equally. Since real gains increase with rank, the proportional growth in real gains decreases with rank, thereby reducing real gain inequality and performance dispersion. In contrast, introducing more sticks after the first one primarily reduces the real gains of middling prizes without affecting higher-ranked prizes, thus increasing real gain inequality and consequently performance dispersion. A formal statement of the effect of adding sticks is as follows:

Corollary 4 (Adding sticks)**.** *Suppose $n \geq 3$. (a) With the rest of the prizes fixed, reducing v_1 decreases the dispersion of the equilibrium performance distribution if $v_2 < v_n$ and has no effect on the performance distribution if $v_2 = v_3 = \ldots = v_n$. (b) When $v_1 = \ldots = v_{k-1} < v_k$, reducing v_k by $\delta \in (0, v_k - v_{k-1}]$ increases the dispersion of the equilibrium performance distribution if $k \leq n-1$ and has no effect on the performance distribution if $k = n$.*

Kempf et al. (2009) find that, in the mutual fund industry, managerial risk-taking is constrained by the employment risk, i.e., the risk of losing jobs faced by low-ranked performers. Corollary 4 implies that, although penalizing the lowest-rank performer reduces managerial risk taking, penalizing other low-ranked performers can actually encourage risk taking and thus can reduce welfare when investors are risk averse.

4.1.2 The effect of contest size on dispersion

We now turn to the effect of a contest size change on performance dispersion. To disentangle the effect of a change in contest size from the effect of a change in real gain inequality, we begin by examining the effect of *contest scaling*, i.e., replacing each prize in the prize schedule by s copies of itself and multiplying the number of contestants by s, where s, the scaling factor, is a positive integer. Importantly, contest scaling does not alter the distribution of prizes or the Lorenz curve of real gains associated with the contest. It thus has no effect on real gain inequality. Contest scaling could result from identical contests being merged into a single larger contest.

Contest scaling scales up the number of contestants and the value of total real gains by the same multiple, without affecting the real gain of the largest prize. Thus, by Lemma 2, it does not change the support of the equilibrium performance distribution. However, it *does* affect equilibrium strategies. To develop an intuitive understanding of this effect, consider the simplest case. Normalize capacity, μ, to 1 and consider the effect of a 100-fold scaling of a contest with two contestants, one winner and one loser. In the 100-fold scaled contest, there are 200 contestants, 100 winners and 100 losers. In both these two cases, the threshold for winning is besting half of the total number of contestants, i.e., to win the competition, a contestant must have her realized performance exceed the "sample" median performance, where the sample is the vector of realized performance levels of all the contestants. Advancing from one position in the ranking to another has value only if the higher-ranked prize is larger than the lower-ranked prize. We call the quantiles of the prize schedule where prize value changes "prize-value-change quantiles." In both the original and the 100-fold scaled contest, the only prize-value-change quantile is the median.

In the original contest, the equilibrium performance distribution is the uniform distribution over $[0,2]$. Thus, the median of the contestant performance distribution equals one. Under the equilibrium performance distribution played in the original contest, a contestant's probability of winning increases uniformly with her performance level. If the same uniform distribution were played in the 100-fold scaled contest, the median of the performance distribution would also equal one. However, because the sample median converges to the median of the performance distribution as the number of contestants increases without bound, in the 100-fold scaled contest, the distribution of the sample median would be sharply increasing around one. Moreover, the median is the only prize-value-change quantile. Hence, to maximize the probability of winning, each contestant would have an incentive to deviate by playing a strategy that mixed zero performance with performance slightly above one, the median of the performance distribution. Thus, in order to maintain the necessary linearity of the contest payoff function, the equilibrium performance distribution of the scaled contest must hollow out around one. In fact, given that the 100-fold scaled contest has 100 identical winner prizes and 100 identical loser prizes, its prize-difference sequence is non-monotone and quasiconcave and, thus, by Proposition 3, the equilibrium performance PDF is U-shaped.

In this simple example, the contest offers prizes of two distinct values and contest scaling clusters mass at the two endpoints of the performance distribution. However, when a contest features prizes of more than two distinct values, and thus there are multiple prize-value-change quantiles, mass around the corresponding performance distribution quantiles that is swept away may approach another quantile of the distribution that corresponds to a prize-value-change quantile. Because too much mass also cannot accumulate around this quantile, probability mass is trapped in the interstices between the two quantiles, leading to "interior clustering" of mass in the interstices. This clustering increases dispersion but the increase in dispersion is limited because the support of the performance distribution is invariant to scaling. The results below formalize this intuition.

Theorem 3 (Scale effects)**.** *Let v and v^s be the prize vectors for contests with n and sn*

contestants, respectively, where $s > 1$ is an integer. Suppose v and v^s satisfy

$$v_i^s = v_{\lceil i/s \rceil} \quad \forall i \in \{1,\ldots,sn\}.$$

Let F and F^s be the equilibrium performance distributions associated with v and v^s, respectively. Then the following results hold:
a. the upper bound of $\mathrm{supp}(F^s)$ *equals the upper bound of* $\mathrm{supp}(F)$*;*
b. $F^s \neq F$ and F^s is more dispersed than F;
c. as $s \to \infty$, F^s converges weakly to a discrete limiting distribution F^∞ defined by

$$F^\infty(x) = \sup\left\{ p \in (0,1) : \frac{\mu n}{V} \times \sum_{\{k:\Delta v_k \neq 0\}} \Delta v_k \mathbb{1}_{k/n}(p) \leq x \right\},$$

where V is defined in (5)*, Δv_k is defined in* (9)*, and $\mathbb{1}_{k/n}$ is the indicator function for the interval $(k/n,\infty)$, i.e., $\mathbb{1}_{k/n}(p) = 0$ if $p \leq k/n$ and $\mathbb{1}_{k/n}(p) = 1$ if $p > k/n$.*

The characterization of the limiting distribution in Theorem 3.c is somewhat implicit. However, Remark A-1 in the proof of Theorem 3 presents a somewhat involved but closed-form algorithm for finding the limit points and their associated probabilities. Theorem 3.c demonstrates that, as scale increases, the PDF becomes more lumpy: valleys flatten and peeks rise. Dispersion increases with scaling simply because locally, around performance levels that correspond to prize-value-change quantiles, dispersion increases. This clustering effect, for a case with multiple distinct prize values and thus interior clustering, is illustrated by Figure 1.

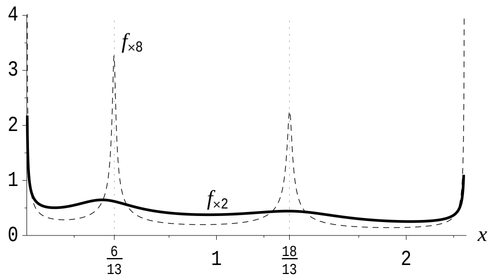


Figure 1: *The clustering effect of contest scaling on performance distributions.* The figure plots the equilibrium PDFs, for a two-fold, $f_{\times 2}$, and an eight-fold, $f_{\times 8}$, scaling of the prize schedule $v = (0, 1/2, 1/2, 3/2, 3/2, 5/2)$. Capacity is normalized to 1. The grid lines indicate the points in the interior of the distributions' supports, specified in Theorem 3.c, which are assigned mass by the limiting distribution, F^∞.

Theorem 3 implies that, if social welfare equals the sum of a fixed number of best performances (e.g., design competitions), grouping smaller local contests into a grand contest improves social welfare. Theorem 3 also predicts that, fixing the distribution of prizes, increasing contest size increases performance dispersion by making performance more clustered.

Now, consider another type of contest transformation that increases contest size: adding a number of contestants along with an equal number of "prizes," each having a value equal to the lowest-rank prize in the original contest. This type of contest size expansion mirrors real world situations in which only a certain number of contestants are rewarded, regardless of the total number of participants in the competition. Consequently, an increase in the number of contestants naturally results in more contestants receiving zero real gains. We refer to the effects of this contestant increase as *entrant effects*. The next result shows that contestant entry increases the range and the dispersion of equilibrium performance.

Theorem 4 (Entrant effects)**.** *Let v and v^e be the prize vectors for contests with n and $n+e$ contestants, respectively, where $e > 0$ is an integer. Suppose v and v^e satisfy*

$$v_i^e = \begin{cases} v_1 & \text{if } i \in \{1, \ldots, e+1\} \\ v_{i-e} & \text{if } i \in \{e+2, \ldots, e+n\} \end{cases}.$$

Let F and F^e be the equilibrium performance distributions associated with v and v^e, respectively. Then the following results hold:
a. the upper bound of $\text{supp}(F^e)$ *is larger than the upper bound of* $\text{supp}(F)$*;*
b. $F^e \neq F$ and F^e is more dispersed than F.

Theorem 4 implies that, if social welfare is convex and increasing in the best performance (e.g., R&D contests) or equals the sum of a fixed number of best performances, increasing contest participation through subsidizing entry with minimum rewards can improve social welfare via two channels. The first channel is purely statistical: increasing the number of contestants increases the number of random draws, which clearly (weakly) increases the values of the greatest draws. The second channel is strategic. Theorem 4 shows that increasing the number of contestants will induce contestants to take more risks, which further benefits society for any given number of contestants. Theorem 4 also implies that, in contests where society suffers from risk taking (e.g., banking and delegated money management), limiting the number of contestants through barriers to entry may be socially beneficial.

4.2 Skewness

Like dispersion, skewness can affect welfare through many channels, e.g., extreme levels of skewness in trading-strategy performance invalidate standard value-at-risk probability bounds for extreme losses and thus compromise systemic risk controls. In cases such as this, the welfare effects of skewness can be mitigated through better modeling or improved institutional arrangements. However, in some cases, the effect of skewness on welfare appears to be an inherent aspect of agent preferences. Many empirical studies document that individuals have rank-order social status concerns (e.g., Brown et al., 2008; Boyce et al., 2010), which have been modeled by Robson (1992) as a status contest in which contestant performance is represented by wealth and contestants take fair gambles in an attempt to reach a higher wealth rank. At the same time, empirical studies also find that overall happiness is highest for populations with negatively skewed wealth distributions (e.g., Smith et al., 1989; Brown et al., 2008).

To facilitate our analysis of performance skewness, we use a widely-adopted skewness order proposed in van Zwet (1964).

Definition 3. Let F and G be two CDFs, both of which are strictly increasing and twice continuously differentiable on their corresponding supports. The distribution G has more (less) skewness than F if and only if $G^{-1} \circ F$ is convex (concave) on the support of F.

Henceforth, when distribution G has more (less) skewness than F in the van Zwet sense, we will say that G is more positively (negatively) skewed than F, or simply that G is *more (less) skewed* than F. Note that this skewness order, like all skewness measures, is scale invariant, i.e., for any two random variables, X and Y, X is more skewed than Y if and only if aX is more skewed than bY, where a and b are arbitrary positive constants. In contrast, the convex order used to measure dispersion is not scale invariant. Nevertheless, when the compared contests have equal contestant capacities and thus equal mean performances, the skewness order implies the convex order.[14] Thus, holding capacity constant, if a change in the contest structure induces the performance distribution to be more skewed, it must also increase performance dispersion.

4.2.1 The effect of prize schedule on skewness

We first fix the number of contestants and study the prize schedule's effect on performance skewness. Our next result provides the condition which ensures that a prize schedule change increases performance skewness.

Theorem 5 (Increase performance skewness)**.** *Consider two contest prize vectors, v and $\hat{v}$, each containing n elements. Let*

$$\left(\frac{\Delta \hat{v}_i}{\Delta v_i}\right)_{i=1}^{n-1}, \qquad \text{where } \Delta v_i = v_{i+1} - v_i \text{ and } \Delta \hat{v}_i = \hat{v}_{i+1} - \hat{v}_i,$$

[14] It is known that the van Zwet skewness order implies the star order (see, e.g., Oja, 1981, Theorem 5.2). Arnold (1987, Theorem 6.6.2) shows that, with identical means between compared distributions, the star order implies the convex order.

be the sequence of the prize-difference ratio between the two prize schedules. If, after removing the possible $\frac{0}{0}$ terms in this sequence, the remaining sequence is nondecreasing, then the equilibrium performance distribution is more skewed under $\hat{v}$ than under v.[15]

Theorem 5 implies that modifying contest rewards in a way that increases the proportional marginal gain from climbing up the prize ladder always increases performance skewness. The increased marginal gain makes reaching for the top prizes more attractive. To maintain the required linearity in the equilibrium contest payoff function, the top end of the performance distribution must stretch out, thereby increasing skewness.

Note that the condition presented in Theorem 5 only requires the prize-difference ratio between the two prize schedules to be monotone in prize ranking over the ranks where prize differences are not vanishing under both schedules. This allows us to show that, fixing the number of contestants, performance skewness is maximized in WTA contests and minimized (i.e., most negative) in elimination contests.

Corollary 5 (Maximal and minimal skewness)**.** *With the number of contestants, n, fixed, the equilibrium performance distribution is most positively skewed in WTA contests, i.e., contests with $v_1 = \cdots = v_{n-1} < v_n$, and most negatively skewed in elimination contests, i.e., contests with $v_1 < v_2 = \cdots = v_n$.*

Kuziemko et al. (2014) find that individuals are "last-place averse." In our setting, last-place aversion can be captured by the first prize difference, Δv_1, and increasing last-place aversion is equivalent to reducing the value of the smallest prize, which increases Δv_1 without changing any other prize differences. Thus, Theorem 5 implies that increasing last-place aversion makes the equilibrium performance distribution less skewed. In fact, Corollary 5 implies that contestants' performance will be most negatively skewed if their incentives are purely driven by last-place aversion. Thus, in status contests where contestant performance is determined by wealth and contestants can choose wealth distributions by taking fair gambles, contestant wealth distributions will be most negatively skewed when status concerns are driven purely by last-place aversion. Since status contests typically involve large numbers of contestants, a negatively skewed contestant wealth distribution can lead to a negatively skewed population wealth distribution. The empirical evidence discussed at the beginning of Section 4.2 suggests that reducing skewness of the population wealth distribution increases overall happiness. Thus, contestant last-place aversion, by reducing skewness of the population wealth distribution, can increase social welfare.

[15] Here we adopt the standard convention that, when the denominator of a ratio is zero and its numerator is positive, the ratio equals infinity.

It is noteworthy that increasing the real gain inequality of prizes need not increase performance skewness, even though, as Theorem 2 shows, increased real gain inequality always increases performance dispersion. A simple example is given by $v = (0,1,2,3)$ and $\hat{v} = (0,0,3,3)$. Clearly, $\hat{v}$ has higher real gain inequality than v. However, by Proposition 4, both v and $\hat{v}$ result in a symmetric performance PDF and thus zero performance skewness.

Nevertheless, as the next result shows, if real gain inequality is increased through a *convexification of the prize schedule*, then both the dispersion and the skewness of performance increase.

Proposition 5 (Convexification effects)**.** *Consider two prize vectors, v and v^h, each containing n elements and $v_i^h = h(v_i)$ for all $i \in \{1,\dots,n\}$, where the function $h : \Re \to \Re$ is nondecreasing. If h is convex, then the equilibrium performance distribution is both more dispersed and more skewed under v^h than under v.*

Intuitively, convexifying the prize schedule reduces the marginal incentives in the lower end of the prize ladder and increase those in the higher end. This leads to a contraction of the lower part of the performance distribution and an extension of the upper part, resulting in an increase in both dispersion and skewness. Proposition 5 implies that enhancing the reward for the best performance increases both dispersion and skewness in performance. In contrast, if all prizes are non-negative, taxing the prizes progressively concavifies the prize schedule and, consequently, reduces both dispersion and skewness in performance.

4.2.2 The effect of contest size on skewness

In contrast to dispersion, the effect of contest size changes on skewness is generally indeterminate. Increasing contest scale by combining identical contests induces performance "clustering," as depicted in Figure 1. It is evident that clustering need not increase performance skewness. As we show in Appendix B, contestant entry can sometimes increase the spread in the lower part of the performance distribution so much that the increased spread in the upper part is insufficient to induce an increase in skewness.

Nevertheless, as the next result shows, adding contestants increases performance skewness if the contest has only two distinct prize values.

Proposition 6 (Entrant effect in contests with two distinct prize values)**.** *Consider an n-player contest with m winners and $n-m$ losers, where all the winners receive the same winner prize and all the losers receive the same loser prize. Add $e > 0$ contestants and e loser prizes into this contest. The equilibrium performance distribution is both more dispersed and more skewed in the new contest than in the baseline contest.*

5 Experimental design

5.1 Treatment design and predictions

Our experiment is designed to test some of the most intriguing predictions derived from our theoretical model. Specifically, we investigate whether increasing real gain inequality, increasing contest scale, or adding contestants increases performance dispersion (Theorems 2, 3, and 4 respectively) and whether convexifying a prize schedule or adding contestants to a contest with two distinct prize values increases performance skewness (Theorem 5 and Proposition 6 respectively). To this end, our experiment consists of seven treatments, with one 2-player treatment, one 6-player treatment, and five 4-player treatments. The prize schedule for each treatment is presented in Table 1.

Treatment	Prize schedule	Dispersion (SD)	Skewness
WTA4	$(10,10,10,370)$	10.205	1.058
Convex	$(10,30,90,270)$	7.492	0.637
Linear	$(10,70,130,190)$	5.196	0
Concave	$(10,100,140,150)$	4.083	-0.577
EC	$(10,130,130,130)$	3.402	-1.058
WTA2	$(10,370)$	5.196	0
3*WTA2	$(10,10,10,370,370,370)$	6.778	0

Table 1: *Prize schedules in different treatments and the associated standard deviation (SD) and skewness coefficient of the predicted equilibrium performance distributions.*

The five 4-player treatments have the same total prize and lowest-rank prize but differ in the degree of convexity in their prize schedules. Treatments WTA4 and Convex both have convex prize schedules, while WTA4 has the most convex prize schedule as it is a WTA contest. Treatment Linear has a linear prize schedule, while Concave and EC both have concave prize schedules. EC, which is an elimination contest, has the most concave prize schedule. These five treatments enable us to test the effects of real gain inequality and convexification. The associated predictions are as follows:

Prediction 1 (Real-gain inequality and convexification effects)**.** The dispersion and skewness of the predicted equilibrium performance distributions among the five 4-player treatments are ranked as follows:

i. WTA4 exhibits the highest performance dispersion, followed by Convex, Linear, Concave, and EC, which has the lowest dispersion.
ii. WTA4 exhibits the highest performance skewness, followed by Convex, Linear, Concave, and EC, which has the lowest skewness.

Part i of Prediction 1 follows from the dispersion effect of real gain inequality stated in Theorem 2, whereas Part ii follows from the convexification effect stated in Theorem 5.

To test the entrant effect, we compare treatment WTA4 to treatment WTA2, which is a two-player WTA contest offering the same prizes for the winner and the loser(s) as in treatment WTA4. By Theorem 4, adding contestants increases performance dispersion. Thus, we expect more dispersed performance distribution in WTA4 than in WTA2. Since WTA2 has two distinct prize values, by Proposition 6, adding contestants to it also increases performance skewness. Thus, we also expect more skewed performance in WTA4 than in WTA2.

Prediction 2 (Entrant effect)**.**

i. Performance dispersion is higher in WTA4 than in WTA2;
ii. Performance skewness is higher in WTA4 than in WTA2.

Treatment 3*WTA2 is a 3-fold scaling of WTA2, and we use 3*WTA2 and WTA2 to test the scale effect. By Theorem 3, scaling up a contest increases performance dispersion. We thus obtain the following prediction:

Prediction 3 (Scale effect)**.** Performance dispersion is higher in 3*WTA2 than in WTA2.[16]

Note that being larger in convex order implies a higher standard deviation (SD) and being larger in van Zwet skewness order implies a higher skewness coefficient (i.e., the third standardized moment). By applying the formula for moments of the equilibrium performance distribution provided in Appendix D to the prize schedule in each treatment, we obtain the theoretical values of SD and skewness presented in Table 1. Since empirical tests of changes in moments are more widely used and easier to handle than empirical tests of stochastic orders, we will test our predictions using SD as a measure of dispersion and the skewness coefficient as a measure of skewness.[17]

Table 1 also reflects the sign of skewness, consistent with the findings in Section 3. Specifically, convex prize schedules, like those in treatments WTA4 and Convex, generate positive performance skewness, whereas concave prize schedules, such as those in treatments Concave and EC, result in negative performance skewness. Symmetric prize differences, as in treatments Linear, WTA2, and 3*WTA2, lead to zero performance skewness. Due to the discrete approximation of the theoretical setup used in our

[16]As discussed in Section 4.2.2, in general, there is no clear prediction for the scale effect on performance skewness. However, in the case of a comparison between WTA2 and 3*WTA2, by Proposition 4, both the two associated prize schedules produce a symmetric performance distribution in equilibrium. We thus expect no difference in the skewness coefficient of performance between these two treatments.

[17]See Barrett and Donald (2003) for methods of testing for stochastic dominance, which require simulation or resampling for inference.

experiment, exact matches between theoretical and observed values are not expected. Nevertheless, we will still compare the observed values to the theoretical ones to identify any interesting patterns of deviations and provide explanations.

5.2 Experiment implementation

Our experiment involves the use of a distribution builder (DB) (Sharpe et al., 2000), a screenshot of which is shown in Figure 2. With the DB, subjects are able to build a probability distribution by distributing 100 blue markers (initially piled to the left of the y-axis) over the integers $0, 1, \ldots, 50$ on the x-axis. Each marker represents a probability of 1% and each integer represents a performance level.

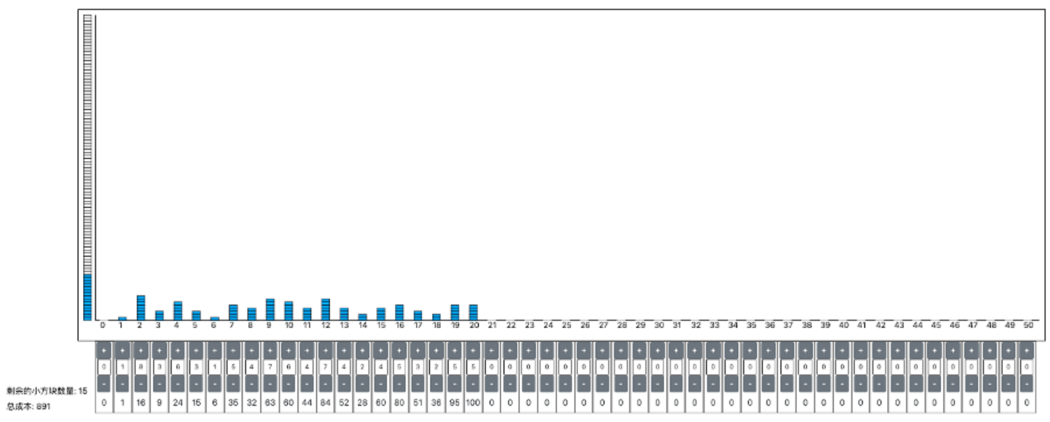


Figure 2: *Distribution builder.* Subjects allocate blue markers by placing them over the integers $0, 1, \ldots, 50$ on the x-axis. Under each integer, a budget box displays two numbers. The number between the +/- buttons shows the number of markers allocated to that integer and can be changed by clicking the +/- buttons or typing a new number. The number below the "-" button displays the cost of the markers allocated to that integer, which always equals the value of the integer times the number of markers allocated to the integer. The bottom-left corner of the DB presents a summary of (i) the number of unused markers (first row), and (ii) the total cost of allocated markers (second row).

To implement a mean constraint on performance, we endow subjects with a use-it-or-waste-it budget of 900 and assign a unit cost of k to allocate a marker to integer k. This way, a subject can adopt a "safe" strategy by placing all 100 markers on the integer 9, thereby consuming the entire budget and achieving a non-stochastic performance level of 9. Alternatively, subjects can opt for a risky strategy by distributing markers above and below 9, but the total cost of this strategy must not exceed 900. Such an experimental design implies a mean performance of 9 if subjects use up their entire budget, and such a mean performance in turn ensures that the theoretical predictions for all our treatments will not be affected by the exogenous upper bound of 50 we impose on admissible performance in the experiment.[18]

To assist subjects in making their decisions, we place a budget box at the bottom-left corner of the DB, which indicates the number of unused markers, and how much budget is consumed by the current distribution. To submit a distribution, subjects have to use all 100 markers, but they are not required to use up their entire budget. However, since the budget is use-it-or-waste-it, having unused budget is suboptimal for a subject.[19]

Our choice of a mean equal to 9 reflects several considerations. First, because performance is integer-valued in our experiment, choosing a too small mean would imply a too-coarse discretization. Since our theoretical predictions are derived from a continuous-performance setup, to better approximate this setup, we avoid choosing a mean that is too small. Second, if we were to choose a very large mean (and thus a large budget), we would also need to select a very large upper bound on performance to ensure it is non-binding in theory. However, it would be overly complex and time-consuming for subjects to allocate markers over a large number of integers with such a large budget. Therefore, we also avoid choosing a mean that is too large. Finally, we want the theoretical performance upper bound to be an integer for all of our treatments, which requires the mean to be a multiple of 3.[20]

We conducted our experiment at Wuhan University between February and March 2023. 336 students were recruited across the campus and participated in one of the seven treatments. Each session included only one treatment condition and consisted of 24 subjects, who were randomly assigned to groups of n for an n-player treatment, where $n = 2, 4, 6$.[21] To ensure an equal number of observation groups, we conducted two sessions for each of the five 4-player treatments, three sessions for the 6-player treatment, and one session for the 2-player treatment, resulting in a total of 14 sessions and 12 groups for each treatment. Sessions lasted approximately 90 minutes, and the average payment to each participant was 82 CNY.

Each session was divided into three parts. The first part served as an introduction and aimed to familiarize subjects with the session's structure, including the prize schedule for the contest part, and the DB. Detailed instructions on how to build an admissible performance distribution using the DB were provided, as were instructions on

[18]The largest theoretical performance upper bound in all our treatments is 36, significantly smaller than 50.

[19]If subjects have any unused budget, they can increase their performance (in the sense of first-order stochastic dominance) by reallocating a marker to a larger integer. In our experiment, out of a total of 6,720 constructed distributions, only 619 did not utilize the full budget. Among these, 444 distributions used at least 895 out of the total 900 budget, while 196 used 899 out of the total 900 budget.

[20]In our 4-player elimination contest treatment, all contestants except the worst performer receive the same prize. The theoretical performance upper bound equals $4/3$ times the mean performance, which is an integer if and only if the mean performance is a multiple of 3.

[21]We determined the order of treatments via random draw before conducting the experiment.

how prizes would be allocated.[22] Subjects also took a comprehension test to ensure they correctly understood the associated rules. The first part ended with a 5-minute practice round, where subjects were asked to build and submit admissible distributions using the DB and observe the realized number randomly drawn from the distribution they built. The practice round was not paid, and subjects were neither assigned to groups nor received any feedback on other participants' distributional choices or realized numbers.

The second part consisted of 20 rounds of one-shot contests under the same treatment conditions. At the beginning of the second part, subjects were randomly assigned to groups of n for n-player treatments, and the group assignments remained fixed throughout the 20 rounds of contests. Given the complexity of the game, the partner matching is to facilitate participants' learning. Since only one round will be randomly drawn for payment and only feedback on group members' drawn numbers (but not their chosen distributions) are revealed at the end of each round, we believe there is little room for reputation building across rounds. Moreover, the fixed-sum nature of the game does not allow strategic cooperation.

During each round, subjects used the DB to build their performance distribution. Two additional buttons were available to assist them: the "reload" button, available from round 2, allowed them to reload the distribution built in the previous round and subsequently make adjustments for the new round, while the "clear" button, available from round 1, quickly erased the existing distribution, giving them the option to start fresh at any time in the building process if desired. Once all group members built and submitted their distributions, each member's performance was randomly drawn from the distribution they built. The group members were then ranked according to their drawn performance, and prizes were allocated accordingly, with ties broken by fair randomization. At the end of each round, subjects received feedback information, including every group member's drawn number, ranking, and prize. We also provided a history box summarizing this information for all previous rounds above the DB.

In the last part of each session, subjects were asked to fill out a questionnaire that collected their demographic information and risk attitudes. Once part three was completed, participants were paid according to the number of experimental currency units (ECUs) they won from one randomly drawn round in part two, under the exchange rate of 1 ECU = 0.5 CNY, plus a show-up fee of 20 CNY.[23]

[22]The original instructions were in Chinese, which can be found in the online appendix. In Appendix G, we provide an English translation of the original instructions.

[23]The exchange rate at the time was approximately 1 CNY = 0.15 USD. The average payment per participant was slightly higher than the average hourly rate for a research assistant at the university.

6 Experimental results

6.1 Overview of the data

For an overview of the observed performance distributions, we first pool all the individual performance distributions in the same treatment and compare the resulting distribution with the simulated theoretical prediction generated from equation (6). Figure 3 displays the equilibrium performance distributions in the first row, while the actual distributions chosen by participants are presented in the second row. Overall, we observe a shift in the probability mass of the observed distributions across treatments in the direction predicted by our theory.[24] This is especially noticeable for the five 4-player treatments, illustrated in the first five columns in Figure 3: as the prize schedule becomes more concave, the probability mass shifts away from the lower end of the distribution while the upper end experiences a contraction, resulting in a more negatively skewed distribution.[25]

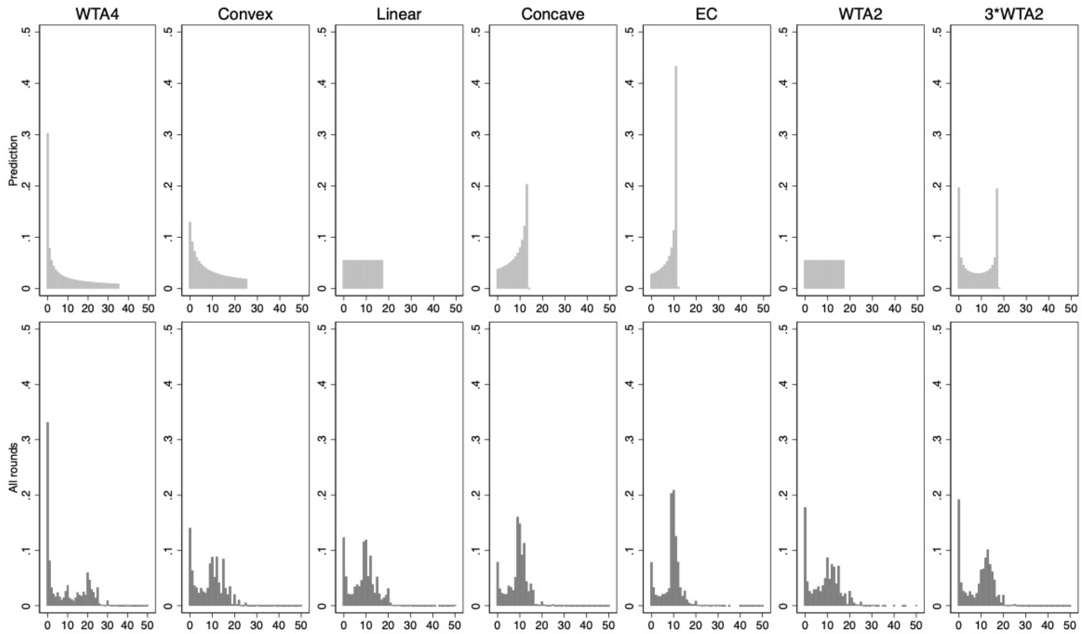


Figure 3: *Simulated equilibrium versus observed performance distributions in each treatment.* The first row displays the simulated distributions from the equilibrium quantile function in each treatment. The second row shows the aggregate performance distributions chosen by the subjects.

Recall that the theoretical upper bound of the support of the equilibrium perfor-

[24]Taking a subsample of all the observations, e.g., rounds 1-10, or rounds 11-20, or the first two rounds only, does not change the overall picture given in Figure 3.

[25]Kolmogorov–Smirnov tests which examine the overall similarity between the empirical distributions and the theoretical predictions are relegated to Appendix F to keep the main results concise.

mance distribution is governed by the fraction of total real gains allocated to the largest prize. In all treatments, enforced by the design of the DB, the observed distribution has a small amount of probability mass on the right tail which extends beyond the theoretical upper bound and reaches the physical upper bound of 50. While these extreme values affect both the dispersion and skewness of the observed distributions, their impact appears to be limited by the fact that only a small proportion of probability mass (5.5% on average) was allocated beyond the theoretical upper bound.[26]

To measure performance dispersion and skewness quantitatively, we use SD and the skewness coefficient. Both measurements are calculated for each individual in each round separately. Table 2 summarizes the observed average SD and skewness in each treatment and compares them with the predictions. Inspection of Table 2 reveals that the observed average SD and skewness are generally not significantly different from the theoretical predictions in treatments with less competitive prize schedules (Linear, Concave, and EC), whereas they significantly fall below the predicted values in treatments with more competitive structures (WTA4, Convex, 3*WTA2).

	SD			Skewness				N
	NE	Obs.	p-value	NE	Obs.	p-value	p-value (against 0)	
WTA4	10.205	9.162 (2.447)	$p < 0.01$	1.058	0.397 (0.600)	$p < 0.01$	$p < 0.01$	960
Convex	7.493	5.911 (2.523)	$p < 0.01$	0.637	-0.142 (1.072)	$p < 0.01$	$p = 0.145$	960
Linear	5.196	5.409 (2.608)	$p = 0.469$	0	-0.191 (1.098)	$p = 0.069$	$p = 0.069$	960
Concave	4.083	4.068 (2.237)	$p = 0.948$	-0.577	-0.741 (1.272)	$p = 0.161$	$p < 0.01$	960
EC	3.402	3.787 (2.421)	$p = 0.135$	-1.058	-0.972 (1.242)	$p = 0.523$	$p < 0.01$	960
WTA2	5.196	6.205 (2.407)	$p < 0.01$	0	-0.158 (0.943)	$p = 0.231$	$p = 0.231$	960
3*WTA2	6.778	5.986 (1.836)	$p < 0.01$	0	-0.415 (0.702)	$p < 0.01$	$p < 0.01$	960

Table 2: *Comparing the observed performance distributions to the simulated equilibrium predictions* The *p*-values for standard deviation (SD) and skewness are based on a one-sample *Wald*-test clustered on the individual level, comparing the observed and the predicted values. Column 8 '*p*-value (against 0)' tests how skewness compares to 0 using a one-sample *Wald*-test clustered on the individual level. Column "N" indicates the number of observations.

Lastly, we also compare the observed skewness measure to zero to test whether the distribution is positively or negatively skewed. As predicted by our theory, the

[26] The percentage of probability mass allocated within the theoretical upper bound in each treatment is as follows: 99.7% in WTA4, 99.5% in Convex, 93.4% in Linear, 90.8% in Concave, 88% in EC, 93% in WTA2, and 96.5% in 3*WTA2.

	Measure	WTA4 Diff	WTA4 p value	Convex Diff	Convex p value	Linear Diff	Linear p value	Concave Diff	Concave p value
Convex	SD	3.251	$p < 0.01$						
	Skew.	0.539	$p < 0.01$						
Linear	SD	3.753	$p < 0.01$	0.502	$p = 0.332$				
	Skew.	0.588	$p < 0.01$	0.049	$p = 0.756$				
Concave	SD	5.094	$p < 0.01$	1.843	$p < 0.01$	1.341	$p < 0.01$		
	Skew.	1.138	$p < 0.01$	0.599	$p < 0.01$	0.55	$p < 0.01$		
EC	SD	5.375	$p < 0.01$	2.124	$p < 0.01$	1.622	$p < 0.01$	0.281	$p = 0.479$
	Skew.	1.369	$p < 0.01$	0.83	$p < 0.01$	0.781	$p < 0.01$	0.231	$p = 0.354$

Table 3: *Real-gain inequality and convexification effects: summary statistics* Note: Column 'Diff' measures the pair-wise differences (column - row) between the observed distributions in different treatments. p-value is generated from two-sample t-tests with standard error clustered on the individual level.

distribution is positively skewed in WTA4, is negatively skewed in Concave and EC, and lacks skewness in Linear and WTA2. Overall, the results suggest that participants tend to choose positively skewed performance distributions under highly convex prize schedules and opt for negatively skewed ones under concave prize schedules.

6.2 Treatment effects

Following the layout of the predictions on treatment comparisons given in Section 5, we first examine the *real-gain inequality and convexification effects*. Table 3 displays the pair-wise comparison among the five 4-player treatments that are used for testing the real-gain inequality and convexification effects. The 'Diff' column presents the column - row treatment differences in SD and skewness. All the differences are positive, which suggests that, in line with our predictions, participants chose a more dispersed and more positively skewed distribution in treatments when the prize schedule is more convex. Testing these differences using two-sample t-tests against zero, we find that, except for the differences between Convex versus Linear and Concave versus EC, all the rest are highly significant.

Participants in our experiment are assigned into fixed groups, so variations among individuals and groups might affect the results. To further validate the observed real-gain inequality and convexification effects, we conduct multi-level mixed-effect regressions controlling for the individual and group-level random effects. Using treatment EC, which has the least convex prize schedule, as a baseline, we present in Table 4 the

	SD		Skewness	
	All rounds	Round 11-20	All rounds	Round 11-20
EC (Const.)	3.787***	3.379***	-0.972***	-1.217***
	(0.334)	(0.355)	(0.114)	(0.135)
Concave	0.281	0.467	0.231	0.236
	(0.472)	(0.502)	(0.161)	(0.191)
Linear	1.622***	1.777***	0.781***	0.792***
	(0.472)	(0.502)	(0.161)	(0.191)
Convex	2.124***	2.533***	0.830***	1.030***
	(0.472)	(0.502)	(0.161)	(0.191)
WTA	5.376***	5.902***	1.369***	1.622***
	(0.472)	(0.502)	(0.161)	(0.191)
Observations	4,800	2,400	4,800	2,400
Number of groups	60	60	60	60
Number of individuals	240	240	240	240

Table 4: *Real-gain inequality and convexification effects: mixed regression* Multi-level mixed regression controlled for individual and group level random effect. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The full regression table can be found in Appendix F.

regression results for all rounds and for the second half (round 11 - 20) of the experiment separately. Consistent with Table 3, all the 4-player treatments other than Concave have larger dispersion and skewness of performance compared with EC. The results do not differ from the overall picture shown in Table 3 concerning pair-wise comparisons either. We also run regressions with extra control variables for individual characteristics such as gender, major, etc., and the real-gain inequality and convexification effects stay robust.

To examine the *entrant effect*, we compare treatments WTA2 and WTA4. The upper panel in Figure 4 plots the average SD and skewness in these two treatments. The $p-$values are computed from a multi-level mixed-effect regression testing the treatment difference.[27] Consistent with our prediction 2, adding two more entrants and two lowest prizes into WTA2, performance dispersion increases (from 6.20 to 9.16, with $p < 0.01$) and performance skewness also increases (from -0.16 to 0.40, with $p < 0.01$). This result does not change qualitatively no matter whether we consider all rounds or only the last 10 rounds.

Finally, we investigate the *scale effect* by comparing WTA2 with its threefold counterpart, 3*WTA2. The comparison is presented in the lower panel in Figure 4. In contrast to the theoretical prediction, the observed dispersion is slightly smaller in treatment 3*WTA2 than in WTA2, but the difference is not significant ($p = 0.586$, see regression Table F-3 in Appendix F). We also observe a significant decrease in skewness in 3*WTA2 compared with WTA2, while our theory predicts zero skewness for both

[27]For the detailed regression results, see Table F-2 in Appendix F.

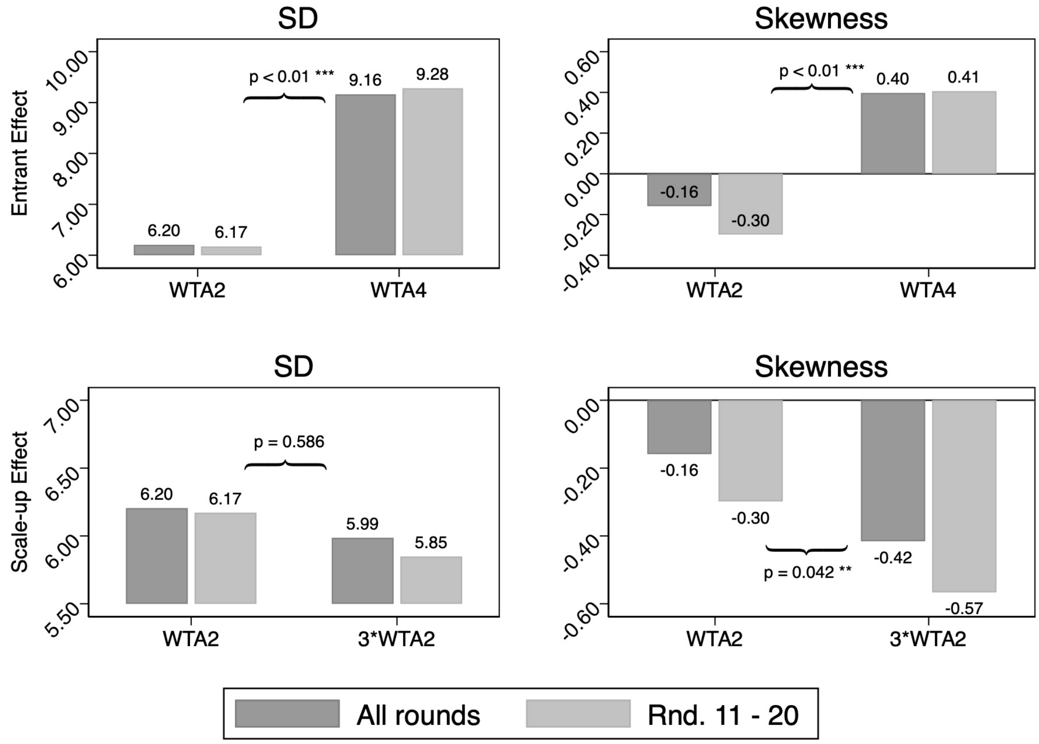


Figure 4: *Entrant effect and scale effect* The dark grey bar represents the average SD/skewness for all rounds, while the light grey bar represents the average for the last 10 rounds (second half) in each treatment. The p-values are computed from a multi-level mixed-effect regression testing the differences between treatments using data from all rounds. Detailed regression results are provided in Tables F-2 and F-3 in Appendix F.

treatments and thus no change of skewness between the two.

Our theory predicts that scaling up a contest increases performance dispersion not by expanding the upper bound of the support of the performance distribution, but rather by making the distribution more lumpy: valleys flatten and peeks rise. In the case of scaling up WTA2, our theory further predicts that the distribution will hollow out around the mean of 9. Inspection of the aggregate distributions in Figure 3 reveals that the observed performance distribution in treatment 3*WTA2 is indeed slightly more lumpy than in WTA2, with more mass on the two modes of the distribution and less mass around the mean of 9.[28] However, the percentage of probability mass allocated to performance levels beyond the theoretical upper bound, which equals 18 in both treat-

[28]The observed probability mass on performance levels 8, 9 (mean), and 10 is smaller in 3*WTA2 compared to WTA2 (2.25% vs. 3.69% on 8, 3.95% vs. 4.98% on 9, and 6.51% vs. 8.73% on 10). In both treatments, the performance level that received the highest probability in our experiment is 0. This (left) mode was chosen with a slightly larger probability in 3*WTA2 than in WTA2 (19.2% vs. 17.8%). The performance level that received the second highest probability is 13 in 3*WTA2 while being 10 in WTA2. The associated probability is again slightly larger in 3*WTA2 than in WTA2 (10.15% vs. 8.73%).

ments, is only 3.5% in 3*WTA2, whereas it is 7% in WTA2. The increased frequency of extreme values used in WTA2 tends to increase both dispersion and skewness, which might be a reason why the observed scale effect does not follow our prediction. However, given the complexity of participants' decisions in constructing their performance distributions and the potential sensitivity of our dispersion and skewness measures to noisy behavior, we are not surprised by the small discrepancies between our theoretical predictions and empirical observations.

Overall, our experimental findings well support the theoretical predictions that convexifying the prize schedule or adding contestants to a contest with a fixed number of identical winner prizes increases both the dispersion and skewness of performance. These results are encouraging for the empirical validity of the model given that the participants in our experiment only learned about the contest game and its mechanism in a short amount of time. We expect the underlying forces identified in this paper to also play an important role in shaping contestants' risk-taking behavior in field settings.

7 Conclusion

In this paper, we examined contestants' risk-taking strategies in contests in which contestants are free to choose performance distributions subject only to a capacity constraint on mean performance. In contrast to the results derived from variance-choice models, contestants do not choose risk-maximizing or risk-minimizing strategies in equilibrium. Instead, equilibrium performance distributions always have an interval support and, under many commonly observed prize schedules, feature nonzero skewness. We showed, both theoretically and experimentally, that contestants choose positively skewed performance distributions when the prize schedule is highly convex (such as in WTA contests) and opt for negatively skewed ones when the prize schedule is concave. Contestants take greater risks in larger contests, as well as in contests with higher real-gain inequality. While the increased risk taking is not universally accompanied by increased skewness, we identified two types of contest structure changes—convexifying the prize schedule and increasing the number of contestants competing for a fixed number of identical winner prizes—that increase both dispersion and skewness of performance.

Our theoretical and experimental results provide insights that can be applied to important real-world contexts. Given that the risk-taking strategies of agents influenced by rank-based incentives affect the stability of the financial system, the pace of invention and discovery, the wealth distributions, and the corporate strategies of CEOs, the insights offered by this paper provide guidance for policymakers attempting to affect the behavior of such agents.

It is interesting to note that the symmetric equilibrium derived in this paper coincides with the symmetric equilibrium derived from Myerson's electoral contest model in which each political party makes campaign promises to a continuum of voters who assign weighted votes to parties based on their ranking of these party promises (Myerson, 1993). Each party's campaign promises must satisfy a budget constraint, imposed by the quantity of public resources that the party is able to allocate in event of winning the election.[29] Myerson shows that each party makes an independent random promise to each voter in equilibrium and, while each party's strategy is expressed explicitly as a resource allocation strategy, it is essentially a choice of a distribution for the random promises subject to the constraint on its mean implied by the fixed budget. Using the maximal campaign promise (i.e., the upper bound of the support of the distribution of campaign promises) as an inequality measure, Myerson argues that *plurality voting*, in which each voter votes for one candidate and the winner is the candidate with the most votes, induces the most unequal distribution of campaign promises, whereas *negative-plurality voting*, in which each voter votes against one candidate and the winner is the candidate with the fewest such votes, induces the most equal distribution of campaign promises. If each voter must vote for exactly m out of n candidates and the winner is the candidate with the most votes, then the distribution becomes more unequal if either m decreases or n increases.

Applying our results to Myerson's framework, noting that plurality voting corresponds to a WTA prize schedule in our model, negative-plurality voting corresponds to an elimination-contest prize schedule, and the case in which each voter votes for exactly m of n candidates corresponds to a prize schedule with m identical prizes for n contestants, we obtain a much stronger and richer implication on how different rank-based voting rules affect the allocation of public resources. Our results on the convexification effect and the entrant effect imply that under *any* commonly used measure of inequality (e.g., Lorenz order, Gini mean difference, standard deviation), the distribution of campaign promises is most unequal and most positively skewed under plurality voting and most equal and most negatively skewed under negative-plurality voting; increasing the number of candidates, n, or reducing the number of votes each voter casts, m, will increase both the inequality and the skewness of the distribution of campaign promises.[30]

[29] If there are only two parties and a finite number of voters, this electoral contest resembles a Colonel Blotto game in which two generals fight in a war by each strategically allocating a fixed number of troops to different battlefields. See Roberson (2006) and Kovenock and Roberson (2021) for a comprehensive analysis of Colonel Blotto games.

[30] Although Myerson (1993) gives some analysis of how the standard deviation of the equilibrium distribution changes with a change of the voting rule, the analysis is based on numerical comparisons between several specific cases.

While our model provides insights into rank-based competitions in various contexts, the analysis could be extended in several interesting directions. First, while our focus was on the dispersion and skewness of performance, the analysis could be extended to higher (standardized) moments of performance. In Appendix D, we provide a formula for calculating any moment of the equilibrium performance distribution. This formula facilitates a comparison of performance kurtosis between different contests. Examining kurtosis could deepen our understanding of how contest structures affect the tail risk in performance distributions.

Second, while our main analysis takes capacity (i.e., mean performance) as given, we could endogenize it through a two-stage model, with capacity building through costly effort in the first stage, followed by capacity-constrained distributional choice in the second stage. This extension could evaluate contest designs from a wider perspective, including both risk taking and effort incentive provision. In Appendix E, we show that, if contestants have the same convex effort cost function and do not observe each other's effort choice, then this two-stage model will produce a symmetric equilibrium in which each contestant chooses the same deterministic effort level in the first stage and this effort level will be determined entirely by per capita real gains.[31] Because the effort level determines mean performance, all of our comparative static results on performance dispersion are robust to this extension if a contest structure change leaves per capita real gains unaffected (e.g., scaling up a contest or transferring value from one prize to another without changing the value of the lowest rank prize). Moreover, all of our theoretical results on the scale-invariant properties of the performance distribution, such as symmetry, modality, tail behavior, and skewness, are completely robust to this extension because mean performance determines only the scale of the distribution.

Third, we could extend our analysis by imposing an exogenous upper bound on performance. Such an extension could be applied to the Bayesian persuasion framework of Kamenica and Gentzkow (2011) to draw predictions about the effect of competition on information revelation. In a Bayesian persuasion model, a sender chooses a signal that transforms a commonly known prior into a distribution of posteriors, subject to the constraint that mean posterior equals the prior. Viewing the posterior as the sender's "performance," the choice becomes analogous to selecting a mean-preserving non-negative performance distribution from our model, with the caveat that the posterior must not ex-

[31] Thus, in this two-stage model, when effort costs are convex, then with per capita real gains fixed, each individual's equilibrium effort is independent of the prize structure (though their equilibrium performance distribution depends on the prize structure). In contrast, when effort costs are not convex, contestants may, in equilibrium, play a mixed strategy by choosing a random effort level in the first stage, and the prize structure can have an effect on effort. See Kim et al. (2023), which considers such a two-stage model with a general effort cost function and shows that the WTA prize schedule maximizes the expected contestant effort regardless of the shape of the effort cost function.

ceed one. Should our comparative static results on performance dispersion hold under this upper-bound constraint—a conjecture we find plausible—then it implies that when competing senders receive rewards based on the ranking of realized posteriors, increasing the real gain inequality of rewards, adding entrants, or scaling up the competition would induce them to choose signals that are (Blackwell) more informative.[32]

Finally, the analysis could be extended to consider contests where the effect of performance on contestant payoffs was not entirely captured by its effect on the prize allocation. One could, for example, consider contests where contestants receive a reward based on absolute performance in addition to the rank-based reward. With fixed mean performance, this extension would not change the equilibrium risk-taking strategy if contestants are risk neutral and the absolute-performance-based reward is linear in performance. In other cases, the equilibrium strategy can be affected.[33] Such extensions require a general "baseline" analysis of risk-taking incentives in pure rank-based competitions, the sort of baseline model provided by this paper.

References

Agarwal, Vikas, Lei Jiang, and Quan Wen, "Why do mutual funds hold lottery stocks?," *Journal of Financial and Quantitative Analysis*, 2022, *57* (3), 825–856.

Andersson, Ola, Håkan J Holm, and Erik Wengström, "Grind or gamble? an experiment on effort and spread seeking in contests," *Economic Inquiry*, 2020, *58* (1), 169–183.

Arnold, Barry C., *Majorization and the Lorenz Order: A Brief Introduction* Lecture Notes in Statistics, Springer, 1987.

Barrett, Garry F and Stephen G Donald, "Consistent tests for stochastic dominance," *Econometrica*, 2003, *71* (1), 71–104.

Becker, Gary S., Kevin M. Murphy, and Iván Werning, "The Equilibrium Distribution of Income and the Market for Status," *Journal of Political Economy*, 2005, *113* (2), 282–310.

[32]See Gentzkow and Kamenica (2016) for an analysis of the effect of competition on information revelation in a Bayesian persuasion model where not all mean-preserving distributions of posteriors are inducible via the available signals.

[33]See Jin and Noe (2022) for the effect on risk taking of adding an absolute-performance-based bonus scheme to a WTA contest.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, “Salience Theory of Choice Under Risk,” *Quarterly Journal of Economics*, 2012, *127* (3), 1243–1285.

Boyce, Christopher J., Gordon D.A. Brown, and Simon C. Moore, “Money and Happiness: Rank of Income, Not Income, Affects Life Satisfaction,” *Psychological Science*, 2010, *21* (4), 471–475.

Brown, Gordon D.A., Jonathan Gardner, Andrew J. Oswald, and Jing Qian, “Does Wage Rank Affect Employees’ Well-being?,” *Industrial Relations: A Journal of Economy and Society*, 2008, *47* (3), 355–389.

Cabral, Luis M. B., “R&D Competition When Firms Choose Variance,” *Journal of Economics and Management Strategy*, 2003, *12* (1), 139–150.

Chen, Joseph, Eric Hughson, and Neal Stoughton, “Strategic Mutual Fund Tournaments,” *Mimeo*, 2015.

Chevalier, Judith and Glenn Ellison, “Risk Taking by Mutual Funds as a Response to Incentives,” *Journal of Political Economy*, 1997, *105* (6), 1167–1200.

_ **and** _ , “Career Concerns of Mutual Fund Managers,” *Quarterly Journal of Economics*, 1999, *114* (2), 389–432.

Conrad, Jennifer, Nishad Kapadia, and Yuhang Xing, “Death and jackpot: Why do individual investors hold overpriced stocks?,” *Journal of Financial Economics*, 2014, *113* (3), 455–475.

Dasgupta, Partha and Joseph Stiglitz, “Uncertainty, Industrial Structure, and the Speed of R&D,” *Bell Journal of Economics*, 1980, *11* (1), 1–28.

Dertwinkel-Kalt, Markus and Jonas Frey, “Optimal stopping in a dynamic salience model,” *International Economic Review*, 2024, *65* (2), 885–913.

Diamond, Peter A. and Joseph E. Stiglitz, “Increases in Risk and in Risk Aversion,” *Journal of Economic Theory*, 1974, *8* (3), 337–360.

Dijk, Oege, Martin Holmén, and Michael Kirchler, “Rank Matters — The Impact of Social Competition on Portfolio Choice,” *European Economic Review*, 2014, *66*, 97–110.

Embrey, Matthew, Christian Seel, and J Philipp Reiss, “Gambling in risk-taking contests: Experimental evidence,” *Journal of Economic Behavior & Organization*, 2024, *221*, 570–585.

Eriksen, Kristoffer W and Ola Kvaløy, "Myopic risk-taking in tournaments," *Journal of Economic Behavior & Organization*, 2014, *97*, 37–46.

_ **and** _ , "No guts, no glory: An experiment on excessive risk-taking," *Review of Finance*, 2017, *21* (3), 1327–1351.

Fang, Dawei and Thomas Noe, "Less competition, more meritocracy?," *Journal of Labor Economics*, 2022, *40* (3), 669–701.

_ **, Martin Holmén, Daniel Kleinlercher, and Michael Kirchler**, "How tournament incentives affect asset markets: A comparison between winner-take-all tournaments and elimination contests," *Journal of Economic Dynamics and Control*, 2017, *75*, 1–27.

Gaba, Anil and Ajay Kalra, "Risk Behavior in Response to Quotas and Contests," *Marketing Science*, 1999, *18* (3), 417–434.

_ **, Ilia Tsetlin, and Robert L. Winkler**, "Modifying Variability and Correlations in Winner-Take-All Contests," *Operations Research*, 2004, *52* (3), 384–395.

Gentzkow, Matthew and Emir Kamenica, "Competition in persuasion," *The Review of Economic Studies*, 2016, *84* (1), 300–322.

Gilpatric, Scott M., "Risk Taking in Contests and the Role of Carrots and Sticks," *Economic Inquiry*, 2009, *47* (2), 266–277.

Goel, Anand M. and Anjan V. Thakor, "Overconfidence, CEO Selection, and Corporate Governance," *Journal of Finance*, 2008, *63* (6), 2737–2784.

Guercio, Diane Del and Paula A Tkac, "Star power: The effect of monrningstar ratings on mutual fund flow," *Journal of Financial and Quantitative Analysis*, 2008, *43* (4), 907–936.

Guo, Liang, "Product design and financial performance," *Design Management Journal*, 2010, *5* (1), 5–19.

Hardy, Godfrey Harold, John Edensor Littlewood, and George Pólya, *Inequalities*, Cambridge university press, 1952.

Hvide, Hans K., "Tournament Rewards and Risk Taking," *Journal of Labor Economics*, 2002, *20* (4), 877–898.

— **and Eirik G. Kristiansen**, "Risk Taking in Selection Contests," *Games and Economic Behavior*, 2003, *42* (1), 172–179.

Jin, Dunhong and Thomas Noe, "The Golden Mean: The Risk-Mitigating Effect of Combining Tournament Rewards with High-Powered Incentives," *The Journal of Finance*, 2022, *77* (5), 2907–2947.

Jones, M. C., "The Complementary Beta Distribution," *Journal of Statistical Planning and Inference*, 2002, *104* (2), 329–337.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615.

Kaniel, Ron and Robert Parham, "WSJ category kings–The impact of media attention on consumer and mutual fund investment decisions," *Journal of Financial Economics*, 2017, *123* (2), 337–356.

Kempf, Alexander, Stefan Ruenzi, and Tanja Thiele, "Employment Risk, Compensation Incentives, and Managerial Risk Taking: Evidence from the Mutual Fund Industry," *Journal of Financial Economics*, 2009, *92* (1), 92–108.

Kim, Kyungmin, Vijay Krishna, and Dmitry Ryvkin, "Choosing your own luck: Strategic risk taking and effort in contests," *Mimeo*, 2023.

Kim, Min S, "Changes in mutual fund flows and managerial incentives," *Available at SSRN 1573051*, 2019.

Kirchler, Michael, Florian Lindner, and Utz Weitzel, "Rankings and risk-taking in the finance industry," *The Journal of Finance*, 2018, *73* (5), 2271–2302.

Klette, Tor and David de Meza, "Is the Market Biased against Risky R&D?," *RAND Journal of Economics*, 1986, *17* (1), 133–139.

Kovenock, Dan and Brian Roberson, "Generalizations of the general lotto and colonel blotto games," *Economic Theory*, 2021, *71*, 997–1032.

Kräkel, Matthias, "Optimal Risk Taking in an Uneven Tournament Game with Risk Averse Players," *Journal of Mathematical Economics*, 2008, *44* (11), 1219–1231.

— **and Dirk Sliwka**, "Risk Taking in Asymmetric Tournaments," *German Economic Review*, 2004, *5*, 103–116.

Kuziemko, Ilyana, Ryan W. Buell, Taly Reich, and Michael I. Norton, "“Last-Place Aversion”: Evidence and Redistributive Implications," *Quarterly Journal of Economics*, 2014, *129* (1), 105–149.

Ma, Linlin, Yuehua Tang, and Juan pedro Gómez, "Portfolio manager compensation in the US mutual fund industry," *Journal of Finance*, 2019, *74* (2), 587–638.

Mitton, Todd and Keith Vorkink, "Equilibrium Underdiversification and the Preference for Skewness," *Review of Financial Studies*, 2007, *20* (4), 1255–1288.

Myerson, Roger B., "Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems," *American Political Science Review*, 1993, *87* (04), 856–869.

Nieken, Petra and Dirk Sliwka, "Risk-Taking Tournaments — Theory and Experimental Evidence," *Journal of Economic Psychology*, 2010, *31*, 254–268.

Oja, Hannu, "On Location, Scale, Skewness and Kurtosis of Univariate Distributions," *Scandinavian Journal of Statistics*, 1981, *8*, 154–168.

Pinelis, Iosif, "L'Hospital Type Rules for Monotonicity, with Applications," *Journal of Inequalities in Pure & Applied Mathematics*, 2002, *3*, Article 5.

Quah, John K-H and Bruno Strulovici, "Aggregating the Single Crossing Property," *Econometrica*, 2012, *80* (5), 2333–2348.

Ray, Debraj and Arthur Robson, "Status, Intertemporal Choice, and Risk-Taking," *Econometrica*, 2012, *80* (4), 1505–1531.

Reiss, R.-D., *Approximate Distributions of Order Statistics*, Springer, 1980.

Roberson, Brian, "The Colonel Blotto Game," *Economic Theory*, 2006, *29*, 1–24.

Robson, Arthur J., "Status, the Distribution of Wealth, Private and Social Attitudes to Risk," *Econometrica*, 1992, *60* (4), 837–857.

Sauder, Michael and Ryon Lancaster, "Do rankings matter? The effects of US News & World Report rankings on the admissions process of law schools," *Law & Society Review*, 2006, *40* (1), 105–134.

Seel, Christian and Philipp Strack, "Gambling in Contests," *Journal of Economic Theory*, 2013, *148* (5), 2033–2048.

Sharpe, William F, Daniel G Goldstein, and Phil W Blythe, "The distribution builder: A tool for inferring investor preferences," *preprint*, 2000.

Sirri, Erik R. and Peter Tufano, "Costly Search and Mutual Fund Flows," *Journal of Finance*, 1998, *53* (5), 1589–1622.

Smith, Richard H., Edward Diener, and Douglas H. Wedell, "Intrapersonal and Social Comparison Determinants of Happiness: A Range-Frequency Analysis," *Journal of Personality and Social Psychology*, 1989, *56* (3), 317–325.

Strack, Philipp, "Risk-Taking in Contests: The Impact of Fund-Manager Compensation on Investor Welfare," *Mimeo*, 2016.

Taylor, Jonathan, "Risk-Taking Behavior in Mutual Fund Tournaments," *Journal of Economic Behavior & Organization*, 2003, *50* (3), 373–383.

van Zwet, W. R., *Convex Transformations of Random Variables*, Math. Centre Tracts 7, Mathematisch Centrum, Amsterdam, 1964.

Wagner, Niklas and Elisabeth Winter, "A New Family of Equity Style Indices and Mutual Fund Performance: Do Liquidity and Idiosyncratic Risk Matter?," *Journal of Empirical Finance*, 2013, *21* (C), 69–85.

Appendix A: Proofs of results

Proof of Proposition 1. The discussion in the main text shows that P increases only over $[0,\bar{x}]$, where $P(x) = \alpha + \beta x$, and P satisfies that $P(0) = v_1$, $P(\bar{x}) = v_n$, and $P(\mu) = \left(\sum_{i=1}^{n} v_i\right)/n$. The proposition follows from these facts. □

Proof of Theorem 1. Follows from Proposition 1, equation (4), and the Binomial Theorem that $\sum_{i=0}^{n-1} \binom{n-1}{i} F^i (1-F)^{n-1-i} = 1$. □

Proof of Corollary 1. Obvious. □

Proof of Lemma 2. See the argument right before Lemma 2 in the main text. □

Proof of Corollary 2. Applying Abel summation to the sum in equation (7) yields

$$Q_v(p) = \frac{\mu n}{V} \sum_{i=0}^{n-2} \left((v_{i+2} - v_{i+1}) \sum_{k=i+1}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \right). \tag{A-1}$$

Re-index using $j = i+1$ and note that $v_{j+1} - v_j = \Delta v_j$. Then equation (A-1) becomes

$$Q_v(p) = \frac{\mu n}{V} \sum_{j=1}^{n-1} \left(\Delta v_j \sum_{k=j}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \right). \tag{A-2}$$

By the relation between the CDF of a Binomial random variable and the CDF of a Beta random variable, we have

$$\sum_{k=j}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = I_p(j, n-j), \tag{A-3}$$

where $I_p(a,b)$ represents the CDF of the Beta distribution with parameters a and b. Inserting (A-3) into equation (A-2) yields

$$Q_v(p) = \frac{\mu n}{V} \sum_{i=1}^{n-1} \Delta v_i I_p(i, n-i). \tag{A-4}$$

To find out the expression for q_v, note that the PDF of the Beta distribution $I_p(i, n-i)$ equals $p^{i-1}(1-p)^{n-i-1}/B(i, n-i)$, where $B(i, n-i)$ is the Beta function equal to $1/\left((n-1)\binom{n-2}{i-1}\right)$. Differentiate Q_v with respect to p, using equation (A-4), apply the expression for the PDF of the Beta distribution, and re-index using $j = i-1$. This yields the expression for q_v in equation (8).

To find out the expression for q_v', differentiate the quantile density function given by equation (8), noting that $n \geq 3$. This yields

$$q_v'(p) = \frac{\mu\, n(n-1)}{V} \left[\left(\sum_{i=1}^{n-2} \Delta v_{i+1} \binom{n-2}{i}\, i\, p^{i-1} (1-p)^{n-2-i} \right) \right.$$
$$\left. - \left(\sum_{i=0}^{n-3} \Delta v_{i+1} \binom{n-2}{i} (n-2-i) p^{i} (1-p)^{n-3-i} \right) \right]. \quad \text{(A-5)}$$

Re-index the first summation in (A-5), using $k = i-1$, and note that $\binom{n-2}{k+1}(k+1) = \binom{n-2}{k}(n-2-k)$. Then the first summation in equation (A-5) is equivalent to

$$\sum_{k=0}^{n-3} \Delta v_{k+2} \binom{n-2}{k} (n-2-k) p^{k} (1-p)^{n-3-k}. \quad \text{(A-6)}$$

Substitute the expression in (A-6) for the first summation in (A-5) and combine common factors. This yields the expression for q_v' given by equation (10). □

Proof of Proposition 2. See the discussion in Section 3.1. □

Proof of Proposition 3. Let $K = \{k \in \{1, \ldots n-2\} : \Delta v_{k+1} - \Delta v_k \neq 0\}$. When the prize-difference sequence is non-monotone, it must be that $K \neq \emptyset$. Thus, $\underline{k} = \min(K)$ exists. Note that q_v' given by (10) can be expanded as a polynomial of p and the lowest power of p is $\underline{k}-1$, with the associated coefficient being $\binom{n-2}{\underline{k}-1}(n-1-\underline{k})\left(\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}}\right)$, i.e.,

$$q_v'(p) = \binom{n-2}{\underline{k}-1}(n-1-\underline{k})\left(\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}}\right) p^{\underline{k}-1} + o(p^{\underline{k}-1}). \quad \text{(A-7)}$$

Suppose the prize-difference sequence is non-monotone and quasiconvex. In this case, it must be that $\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}} < 0$. Then equation (C-2) implies that q_v' is negative in a neighborhood of 0. The argument right before Proposition 3 in the main text shows that q_v' changes sign exactly once if the prize-difference sequence is non-monotone and quasiconvex. Thus, over $p \in (0,1)$, q_v' is first negative and then positive, implying a U-shaped q_v and hence an inverse U-shaped PDF, i.e., a unimodal distribution.

The result for the case of a non-monotone and quasiconcave prize-difference sequence can be established analogously. We thus omit its proof. □

Proof of Proposition 4. We first show sufficiency. Suppose condition (13) holds. The quantile density function given by equation (8) implies that

$$q_v(1-p) = \frac{\mu\, n(n-1)}{V} \left(\sum_{j=0}^{n-2} \Delta v_{n-1-j} \binom{n-2}{j} p^{j} (1-p)^{n-2-j} \right). \quad \text{(A-8)}$$

Equations (8), (13), and (A-8) imply that $q_v(1-p) = q_v(p)$. The symmetry of q_v implies the symmetry of the associated PDF.

Next, we show necessity. Suppose the PDF is symmetric. The symmetry of the PDF implies the symmetry of q_v and, hence, $q_v(1-p) - q_v(p) = 0$ for all $p \in (0,1)$. Given that $\mu n(n-1)/V$ is a positive constant, by (8) and (A-8), this further implies that

$$\sum_{j=0}^{n-2} (\Delta v_{n-1-j} - \Delta v_{j+1})\, b_{j,n-2}(p) = 0, \quad \forall p \in (0,1)\,, \text{ where} \tag{A-9}$$

$$b_{j,n-2}(p) = \binom{n-2}{j} p^j (1-p)^{n-2-j}.$$

The function $p \hookrightarrow \sum_{j=0}^{n-2} (\Delta v_{n-1-j} - \Delta v_{j+1})\, b_{j,n-2}(p)$ is a polynomial and thus analytic. Hence, if it vanishes on any interval, it must identically equal 0 everywhere. Thus, if (A-9) holds, then

$$\sum_{j=0}^{n-2} (\Delta v_{n-1-j} - \Delta v_{j+1})\, b_{j,n-2}(p) = 0, \quad p \in \Re. \tag{A-10}$$

However, the set of Bernstein polynomials, $\{b_{j,n-2}\}_{j=0}^{n-2}$, is a basis for the $n-2$ degree polynomials. Thus, a linear combination of the basis elements produces the zero polynomial if and only if all the coefficients of the linear combination equal 0. Hence, (A-10) implies equation (13). □

Proof of Lemma 3. Suppose w is convex in every argument and suppose F dominates G in the convex order. For each contestant $i \in \{1,\ldots,n\}$, with the rest of the contestants' performance fixed, w is convex in x_i. Thus, by the definition of convex order and the independence of performance, fixing the rest of the contestants' performance distributions, the expected value of w is higher if contestant i's performance is drawn from F than from G. Applying this argument to every contestant establishes part (a). Part (b) can be established analogously. We thus omit its proof. □

Proof of Theorem 2. An *inequality increasing transfer* is a transfer of value from a lower-rank prize to a higher-rank prize, without any change in total prize value or the ranking of each prize. To prove the theorem, we first show that

Lemma A-1. *Fixing v_1, an inequality increasing transfer makes the equilibrium performance distribution more dispersed.*

Proof. Consider an inequality increasing transfer of $\delta > 0$ from the ith to the kth prize, where $2 \le i < k \le n$. Let v and v' be the prize schedules before and after the transfer, respectively. It is clear that

$$v'_j = \begin{cases} v_j - \delta & \text{if } j = i \\ v_j + \delta & \text{if } j = k \\ v_j & \text{otherwise} \end{cases}. \tag{A-11}$$

Note that, such a transfer does not change the total real gains, V. Thus, by Corollary 1, the difference between the quantile functions, $Q_{v'}$ and Q_v, equals

$$Q_{v'}(p) - Q_v(p) = \frac{\mu n \delta}{V} p^{k-1}(1-p)^{n-k} H(p), \quad p \in (0,1), \quad \text{where}$$
$$H(p) = -\binom{n-1}{i-1}\left(\frac{1-p}{p}\right)^{k-i} + \binom{n-1}{k-1}.$$

Clearly, for each $p \in (0,1)$, the sign of $Q_{v'}(p) - Q_v(p)$ is determined by the sign of $H(p)$. Given $k > i$, it is clear that $\lim_{p \downarrow 0} H(p) < 0$, $\lim_{p \uparrow 1} H(p) > 0$, and $p \hookrightarrow H(p)$ is continuous and strictly increasing. Thus, $Q_{v'}$ and Q_v must cross exactly once, in the way that $Q_{v'}$ crosses Q_v from below. Hence, their associated CDFs, $F_{v'}$ and F_v, being their inverse functions, must also cross exactly once, in the way that $F_{v'}$ crosses F_v from above. Since $F_{v'}$ and F_v have the same mean, this single-crossing property implies that $F_{v'}$ differs from F_v by a simple mean-preserving spread (Diamond and Stiglitz, 1974). Because, by Diamond and Stiglitz (1974), a simple mean-preserving spread implies an increase in dispersion in the sense of convex order, such an inequality increasing transfer makes the equilibrium performance distribution more dispersed. □

Let v^r be the real gain vector associated with v, i.e., $v_i^r = (v_i - v_1)/V$ for all $i \in \{1,\ldots,n\}$. Let $\hat{v}^r$ be the real gain vector associated with $\hat{v}$. Hardy, Littlewood, and Polya results imply that $\hat{v}$ is more unequal than v if and only if there exits a (perhaps countably infinite) collection of inequality increasing transfers that transform v^r into $\hat{v}^r$ (Hardy et al., 1952, Proof of Theorem 42). Because $\hat{v}_1^r = v_1^r = 0$, none of these transfers affects the smallest real gain. Thus, by Lemma A-1, each of these transfers makes the equilibrium performance distribution more dispersed in the convex order. The convex order is transitive. Thus, the composition of any number of such transformations preserves this order relation. Thus, the performance distribution under $\hat{v}^r$ is more dispersed than the one under v^r. By part (a) of Lemma 2, the former distribution is the same as the one under $\hat{v}$ and the latter is the same as the one under v. The result thus follows. □

Proof of Corollary 3. Follows immediately from Theorem 2. □

Proof of Corollary 4. Let v and $\hat{v}$ be two prize vectors, each with $n \geq 3$ elements, such that $\hat{v}_i = v_i - \delta$, $\delta > 0$, if $i = k$, $\hat{v}_i = v_i$ if $i \neq k$, and $v_1 = \ldots = v_{k-1}$ if $k > 1$. Let V and $\hat{V}$ be the total real gains associated with v and $\hat{v}$, respectively.

(a): Suppose $k = 1$. If $v_2 = \ldots = v_n$, reducing v_1 leads to an affine transformation of the prize schedule, which, by Lemma 2, does not affect the performance distribution. Now consider the case where $v_2 < v_n$. Since the first real gain is always 0 and the sum of real gains normalized by total real gains is always 1, by Theorem 2, to prove that the dispersion under $\hat{v}$ is less than that under v when $k = 1$ and $v_2 < v_n$, it suffices to show the following:

$$\frac{\sum_{i=1}^{j}(\hat{v}_i - \hat{v}_1)}{\hat{V}} \geq \frac{\sum_{i=1}^{j}(v_i - v_1)}{V}, \quad \forall j \in \{2, \ldots, n-1\}, \tag{A-12}$$

with strict inequality holding for some j. We show this as follows.

$$\text{For } j \geq 2, \quad \frac{\sum_{i=1}^{j}(\hat{v}_i - \hat{v}_1)}{\hat{V}} = \frac{\sum_{i=2}^{j}(v_i - v_1 + \delta)}{\sum_{i=2}^{n}(v_i - v_1 + \delta)} = \frac{(j-1)\delta + \sum_{i=2}^{j}(v_i - v_1)}{(n-1)\delta + V}. \tag{A-13}$$

Thus, by equation (A-13),

$$\frac{\sum_{i=1}^{j}(\hat{v}_i - \hat{v}_1)}{\hat{V}} - \frac{\sum_{i=1}^{j}(v_i - v_1)}{V} = c\left((j-1)V - (n-1)\sum_{i=2}^{j}(v_i - v_1)\right), \tag{A-14}$$

where $c = \delta/(V(V + (n-1)\delta)) > 0$. Thus, to show (A-12), it suffices to show that

$$(j-1)V - (n-1)\sum_{i=2}^{j}(v_i - v_1) \geq 0, \quad \forall j \in \{2, \ldots, n-1\}, \tag{A-15}$$

with strict inequality for some j. Note that

$$(j-1)V - (n-1)\sum_{i=2}^{j}(v_i - v_1) = (j-1)\left(\sum_{i=j+1}^{n}(v_i - v_1)\right) - (n-j)\left(\sum_{i=2}^{j}(v_i - v_1)\right)$$

$$\geq (j-1)(n-j)(v_{j+1} - v_1) - (n-j)(j-1)(v_j - v_1) = (j-1)(n-j)(v_{j+1} - v_j) \geq 0.$$

When $v_2 < v_n$, the last inequality must be strict for some $j \in \{2, \ldots, n-1\}$. This implies, through (A-14) and (A-15), that (A-12) holds with strict inequality for some j. Thus, the result is immediate from Theorem 2.

(b): If $k = n$, given that $v_1 = \ldots = v_{k-1}$, reducing v_k leads to an affine transformation of the prize schedule, which, by Lemma 2, does not affect the performance distribution. Now consider the case where $1 < k \leq n-1$. By Theorem 2, to prove that the dispersion under $\hat{v}$ is greater than that under v when $1 < k \leq n-1$, it suffices to show that

$$\frac{\sum_{i=1}^{j}(\hat{v}_i - \hat{v}_1)}{\hat{V}} \leq \frac{\sum_{i=1}^{j}(v_i - v_1)}{V}, \quad \forall j \in \{2, \ldots, n-1\},$$

with strict inequality for some j. Note that $v_1 = \ldots = v_{k-1}$ and $v_n - v_1 > 0$, and by the hypothesis in the corollary, $\delta \leq v_k - v_{k-1}$, which implies $V > \delta$. Thus,

$$\frac{\sum_{i=1}^{j}(\hat{v}_i - \hat{v}_1)}{\hat{V}} = \begin{cases} 0 = \frac{\sum_{i=1}^{j}(v_i - v_1)}{V} & \text{if } j < k \\ \frac{\left(\sum_{i=1}^{j}(v_i - v_1)\right) - \delta}{V - \delta} < \frac{\sum_{i=1}^{j}(v_i - v_1)}{V} & \text{if } k \leq j \leq n-1 \end{cases}. \tag{A-16}$$

The result thus follows from equation (A-16) and Theorem 2. □

Proof of Theorem 3. (a): Note that scaling does not change the real gain of the largest prize. Nor does it change the ratio of the total real gains to the number of contestants. Thus, by Lemma 2, there is no change in the upper bound of the support of the equilibrium distribution.

(b): First note that the duplication of prizes moves a prize from position k in the baseline prize schedule to position sk in the new prize schedule. The duplicate does not generate any nonzero prize differences, so the values and number of nonzero prize differences are the same as they were in the baseline schedule; the only effect of scaling is to advance prize difference Δv_k from position k to position sk. Hence, expressing the equilibrium quantile function for the scaled contest using equation (A-4) and noting that the ratio n/V is the same for the scaled contest and the baseline contest, we see that the quantile distribution associated with the scaled contest is given by

$$\frac{\mu n}{V} \times \sum_{\{k:\Delta v_k \neq 0\}} \Delta v_k I_p(sk, s(n-k)) \tag{A-17}$$

and the quantile function associated with the baseline contest is

$$\frac{\mu n}{V} \times \sum_{\{k:\Delta v_k \neq 0\}} \Delta v_k I_p(k, n-k),$$

where $I_p(a,b)$ represents the CDF of the Beta distribution with parameters a and b. Without loss of generality, assume that the scaler, $\mu n/V$, satisfies $\mu n/V = 1/(v_n - v_1)$, in which case the sum of the nonzero differences multiplied by $\mu n/V$ will equal 1. Let Q and Q^s be the quantile functions for the baseline and the scaled contest, respectively. Under the assumption we made on $\mu n/V$, Q and Q^s are also distribution functions. The first step in the proof is to show that

Result A-1. $I_p(sk, s(n-k))$ is dominated by $I_p(k, n-k)$ in the convex order.

Proof. To prove this result, it suffices to show that $I_p(k, n-k)$ differs from $I_p(sk, s(n-k))$ by a simple mean-preserving spread. Given that $I_p(k, n-k)$ and $I_p(sk, s(n-k))$ have the same mean, a sufficient condition for the simple mean-preserving spread relation is that their densities, $q_{n,k}$ and $q^s_{n,k}$, satisfy the condition that $q_{n,k} - q^s_{n,k}$ changes sign twice with the pattern $+-+$. This will be true if and only if $\log \circ q_{n,k} - \log \circ q^s_{n,k}$ changes sign twice with the pattern $+-+$. Note that

$$\log \circ q_{n,k}(p) - \log \circ q^s_{n,k}(p) = \\ \log\left(\frac{B(sk, s(n-k))}{B(k, n-k)}\right) - (s-1)\Big((n-k)\log(1-p) + k\log(p)\Big).$$

The function $p \hookrightarrow \log \circ q_{n,k}(p) - \log \circ q^s_{n,k}(p)$ tends to positive infinity as p tends to 0 or 1. Since both $q_{n,k}$ and $q^s_{n,k}$ are densities and are not identically equal, it is not possible for the ratio $q_{n,k}/q^s_{n,k}$ to be no less than 1 for all $p \in (0,1)$. Thus, there must be some point where $\log \circ q_{n,k}(p) - \log \circ q^s_{n,k}(p) < 0$. Thus, $p \hookrightarrow \log \circ q_{n,k}(p) - \log \circ q^s_{n,k}(p)$ crosses the x-axis at least twice. Since $p \hookrightarrow \log \circ q_{n,k}(p) - \log \circ q^s_{n,k}(p)$ is convex, it crosses the

x-axis at most twice. Thus, it crosses the x-axis twice, first from above and next from below. Thus, the mean-preserving spread relation holds and the result follows. □

Since convex order is closed under mixtures, Result A-1 permits us to conclude that

Result A-2. Q^s is dominated by Q in the convex order.

Now let F and F^s be the CDFs of the performance distribution associated with the baseline and the scaled contest, respectively. Define

$$L(p) = \frac{1}{\mu} \int_0^p Q(t)\, dt \quad \text{and} \quad L^s(p) = \frac{1}{\mu} \int_0^p Q^s(t)\, dt.$$

By definition, L and L^s are the Lorenz curves associated with F and F^s, respectively. Result A-2 implies that

$$L^s(p) \le L(p), \quad p \in (0,1). \tag{A-18}$$

Equation (A-18), combined with Arnold (1987, Corollary 3.2.1), implies that F^s dominates F in the convex order.

(c): By the argument given in the proof of part (b) of the theorem, the quantile function associated with the scaled contest, Q^s, is given by (A-17). As $s \to \infty$, $I_p(sk, s(n-k)) \xrightarrow{dist.} \delta_{k/n}$, where $\delta_{k/n}$ represents the point mass distribution function for a point mass on k/n. Thus

$$Q^s \to Q^\infty(p) = \frac{\mu\, n}{V} \times \sum_{\{k:\Delta v_k \neq 0\}} \Delta v_k \delta_{k/n}(p-). \tag{A-19}$$

The left-continuous version of $\delta_{k/n}(p)$, given by $\delta_{k/n}(p-)$, is used in the summand in (A-19) because, by convention, quantile functions, in contrast to distribution functions, are represented by left-continuous nondecreasing functions. The generalized inverse of the quantile function can be used to yield the distribution function using a standard formula (e.g.,Theorem 1.2.8 Reiss, 1980). This formula and the observation that $\delta_{k/n}(p-) = \mathbb{1}_{k/n}(p)$, where $\mathbb{1}_{k/n}(p)$ is the indicator function for $(k/n, \infty)$, yields the expression for F^∞ in the theorem. Finally, note that the performance distribution associated with the scaled contest, F^s, converges to F^∞ in distribution if and only if $Q^s \to Q^\infty$ at every continuity point $p \in (0,1)$ of Q^∞ (Theorem 1.2.9 Reiss, 1980).

The following algorithm can be used to compute the jump points and the mass on the jump points for the limiting distribution, F^∞, defined in the theorem.

Remark A-1 (Algorithm for explicit computation of F^∞)**.** F^∞ stated implicitly in Theorem 3.c can be explicitly computed as follows: Let $K = \{k \in \{1, \ldots, n-1\} : \Delta v_k \neq 0\}$, $\underline{k} = \min(K)$, $\bar{k} = \max(K)$, and $\ell = \#(K)$. Define F^∞ by the following conditions:

a. F^∞ places weight only on $\ell + 1$ points, including 0 and $\{\mu\, n\,(v_{k+1} - v_1)/V : k \in K\}$.

b. The weight on 0 equals $\underline{k}/n$,

c. The weight on $\mu\, n\,(v_{k+1}-v_1)/V$ for all $k \in K\backslash\{\bar{k}\}$ equals $\frac{\min\{j\in K: j>k\}-k}{n}$.

d. The weight on $\mu\, n\,(v_{\bar{k}+1}-v_1)/V$ equals $1-\frac{\bar{k}}{n}$.

□

Proof of Theorem 4. (a): Note that adding contestants does not change the real gain of the largest prize or the total real gains. Thus, its effect on the upper bound of the support purely comes from the increase in n. The result then follows from Lemma 2.

(b): To show the convex order, it suffices to show that F second-order stochastically dominates F^e, which is to show that $\int_0^x (F(t)-F^e(t))dt \leq 0$ for all $x \geq 0$, with strict inequality at some x. This condition, translated into the quantile functions (Arnold, 1987, Corollary 3.2.1), is

$$\int_0^p (Q(t)-Q^e(t))dt \geq 0 \quad \forall p \in [0,1], \quad \text{with strict inequality at some } p, \tag{A-20}$$

where Q and Q^e represent the quantile functions associated with F and F^e, respectively. Below, we prove that the condition in (A-20) holds. Let

$$G_k(p) = I_p(k, n-k), \qquad G_k^e(p) = I_p(k+e, n-k), \tag{A-21}$$

$$Q_k(p) = n\mu\, G_k(p), \qquad Q_k^e(p) = (n+e)\mu\, G_k^e(p). \tag{A-22}$$

Note that adding contestants does not change the total real gains, V. Also, the prize differences satisfy that $\Delta v_{k+e}^e = \Delta v_k$ for all $k \in \{1,\ldots,n-1\}$ and $\Delta v_k^e = 0$ for all $k \leq e$. Thus, by equations (A-4), (A-21), and (A-22),

$$\int_0^p (Q(t)-Q^e(t))dt = \frac{1}{V}\sum_{k=1}^{n-1} \Delta v_k \int_0^p (Q_k(t)-Q_k^e(t))dt. \tag{A-23}$$

Now consider the terms, $Q_k(p)-Q_k^e(p)$. We first argue that the ratio $Q_k^e(p)/Q_k(p)$ is increasing in p. Let $q_k(p)$ and $q_k^e(p)$ be the quantile densities associated with $Q_k(p)$ and $Q_k^e(p)$, respectively. Note that the density of $I_p(k,n-k)$, the CDF of a Beta distribution, is the PDF of the Beta distribution, $p^{k-1}(1-p)^{n-k-1}/B(k,n-k)$, where $B(k,n-k)$ is the Beta function. Thus, by equations (A-21) and (A-22), for $p \in (0,1)$,

$$\frac{q_k^e(p)}{q_k(p)} = \left(\frac{n+e}{n}\right) \frac{B(k,n-k)}{B(k+e,n-k)}\, p^e. \tag{A-24}$$

Because $e > 0$, it is clear from (A-24) that

$$p \hookrightarrow q_k^e(p)/q_k(p) \text{ is increasing.} \tag{A-25}$$

Since $Q_k(p)$ and $Q_k^e(p)$ are continuous and differentiable on $[0,1]$ and $Q_k(0) = Q_k^e(0) = 0$, by the monotone form of l'Hôspital's rule (Pinelis, 2002, Proposition 1.1), (A-25) implies that

$$p \hookrightarrow Q_k^e(p)/Q_k(p) \text{ is increasing.} \tag{A-26}$$

Next, note that

$$Q_k^e(1) = \mu(n+e)G_k^e(1) = \mu(n+e) > \mu n = \mu n G_k(1) = Q_k(1).$$

Thus, given that Q_k^e and Q_k are continuous,

$$Q_k^e(p) > Q_k(p) \quad \text{on some open neighborhood of 1.} \tag{A-27}$$

Finally, note that

$$\int_0^1 Q_k(t)dt = n\mu \int_0^1 G_k(t)dt. \tag{A-28}$$

Since G_k is the CDF of the Beta distribution with parameters k and $n-k$, by Jones (2002), G_k is also the quantile function for the *complementary Beta distribution* with parameters k and $n-k$. The mean of the complementary Beta distribution with parameters a and b is $b/(a+b)$ (Jones, 2002, §2.5, the first sentence). Thus, $\int_0^1 G_k(t)dt$, which represents the mean of the complementary Beta with parameters k and $n-k$, equals $(n-k)/n$. Hence, by (A-28),

$$\int_0^1 Q_k(t)dt = n\mu \int_0^1 G_k(t)dt = n\mu \text{ Mean[CompBeta}[k, n-k]] = \mu(n-k).$$

Similarly,

$$\int_0^1 Q_k^e(t)dt = (n+e)\mu \int_0^1 G_k^e(t)dt = (n+e)\mu \text{ Mean[CompBeta}[k+e, n-k]] = \mu(n-k).$$

Thus,

$$\int_0^1 Q_k^e(t)dt = \int_0^1 Q_k(t)dt. \tag{A-29}$$

Equations (A-27) and (A-29) imply that the functions $p \hookrightarrow Q_k^e(p)$ and $p \hookrightarrow Q_k(p)$ must cross somewhere on $(0,1)$. Then (A-26) implies that they can only cross once with Q_k initially above Q_k^e and ultimately below. This, combined with (A-29), implies that

$$\int_0^p (Q_k(t) - Q_k^e(t))dt \geq 0 \quad \forall p \in (0,1), \quad \text{with strict inequality at some } p. \tag{A-30}$$

Equations (A-23) and (A-30) imply (A-20), and the result follows.[34] □

Proof of Theorem 5. Let F and $\hat{F}$ be the equilibrium performance distributions associated with v and $\hat{v}$, respectively. Let Q and $\hat{Q}$ be their associated quantile functions and q and $\hat{q}$ their associated quantile density functions. By the definition of quantile functions, $\hat{F}^{-1} \circ F$ being convex is equivalent to $\hat{Q} \circ Q^{-1}$ being convex; that is, $\hat{Q}$ is convex with respect to Q. Given that equilibrium distributions are differentiable,

$$p \hookrightarrow \frac{\hat{q}(p)}{q(p)} \text{ is nondecreasing } \Leftrightarrow \hat{Q} \text{ is convex with respect to } Q.$$[35]

[34] See Strack (2016) for a proof of the result using the established result in Quah and Strulovici (2012) for aggregating the single-crossing property.

[35] Cargo, G. T. (1965). Comparable means and generalized convexity, *Journal of Mathematical Anal-*

Thus,

$$\hat{F} \text{ is more skewed than } F \Leftrightarrow p \hookrightarrow \frac{\hat{q}(p)}{q(p)} \text{ is nondecreasing} . \tag{A-31}$$

By equation (8),

$$\frac{\hat{q}(p)}{q(p)} = k \frac{\sum_{i=1}^{n-1} \Delta\hat{v}_i b[n-2,p](i-1)}{\sum_{i=1}^{n-1} \Delta v_i b[n-2,p](i-1)}, \tag{A-32}$$

where $k = V/\hat{V}$ is a positive constant and $b[n-2,p](i-1) = \binom{n-2}{i-1} p^{i-1}(1-p)^{n-1-i}$ represents the probability mass function of the Binomial distribution with parameters $n-2$ and p evaluated at point $i-1$. Let $D_n = \{1,\ldots,n-1\}$ represent the set of prize-difference indices. Define

$$\begin{aligned} J &= \{i \in D_n : \Delta v_i = \Delta\hat{v}_i = 0\} \\ J' &= \{i \in D_n : \Delta v_i = 0, \Delta\hat{v}_i \neq 0\} . \end{aligned} \tag{A-33}$$

We prove the theorem in two steps. The first step is Result A-3.

Result A-3. If $J' = \emptyset$ and, for all $i \in D_n \backslash J$, $\Delta\hat{v}_i/\Delta v_i$ is nondecreasing, then $\hat{F}$ is more skewed than F.

Proof. Define the set

$$\mathscr{I} = \left\{ j \in D_n : \Delta v_j \neq 0 \right\} .$$

Note that, by the definitions of J and J' in (A-33),

$$\mathscr{I} = D_n \backslash (J \bigcup J').$$

Given $J' = \emptyset$, if there exists $i \in D_n$ such that $\Delta v_i = 0$, it must be that $\Delta\hat{v}_i = 0$. Then equation (A-32) implies that

$$\frac{\hat{q}(p)}{q(p)} = k \frac{\sum_{i \in \mathscr{I}} \frac{\Delta\hat{v}_i}{\Delta v_i} \Delta v_i b[n-2,p](i-1)}{\sum_{i \in \mathscr{I}} \Delta v_i b[n-2,p](i-1)}. \tag{A-34}$$

For each $p \in (0,1)$, define the probability measure $\Pi[p]$ over $\mathscr{I}$ as follows: First let

$$S(p) = \sum_{i \in \mathscr{I}} \Delta v_i b[n-2,p](i-1);$$

then define

$$\Pi[p](i) = \frac{\Delta v_i b[n-2,p](i-1)}{S(p)}.$$

Note that, if $p_2 > p_1$, $\Pi[p_2]$ dominates $\Pi[p_1]$ in the monotone likelihood ratio order and thus, *a fortiori*, $\Pi[p_2]$ first-order stochastically dominates $\Pi[p_1]$. Also note that

$$\frac{\hat{q}(p)}{q(p)} = k \sum_{i \in \mathscr{I}} \frac{\Delta\hat{v}_i}{\Delta v_i} \Pi[p](i). \tag{A-35}$$

---

ysis and Applications, 12(3), 387-392.

Thus, if $\Delta\hat{v}_i/\Delta v_i$ is nondecreasing, the stochastic dominance of $\Pi[p_2]$ over $\Pi[p_1]$ implies that $\frac{\hat{q}(p_2)}{q(p_2)} \geq \frac{\hat{q}(p_1)}{q(p_1)}$. Thus, by (A-31), $\hat{F}$ is more skewed than F. □

The next step is Result A-4, which shows that the theorem also holds if J' is not empty and if all the prize-difference indices contained in J' are larger than those contained in $D_n\backslash(J\bigcup J')$.

Result A-4. If $J' \neq \emptyset$, $\min[J'] > \max[D_n\backslash(J\bigcup J')]$, and, for all $i \in D_n\backslash(J\bigcup J')$, $\Delta\hat{v}_i/\Delta v_i$ is nondecreasing, then $\hat{F}$ is more skewed than F.

Proof. Define another prize vector $\tilde{v}$ such that $\Delta\tilde{v}_i = \Delta\hat{v}_i$ if $i \in D_n\backslash J'$ and $\Delta\tilde{v}_i = 0$ if $i \in J'$. Consider the effect on skewness of a prize schedule change from v to $\tilde{v}$. Note that, by our construction of $\tilde{v}$, there does *not* exist any i such that $\Delta v_i = 0$ while $\Delta\tilde{v}_i \neq 0$. Also, given that, for all $i \in D_n\backslash(J\bigcup J')$, $\Delta\hat{v}_i/\Delta v_i$ is nondecreasing, by construction of $\tilde{v}$, $\Delta\tilde{v}_i/\Delta v_i$ is nondecreasing for all $i \in \left\{j \in \{1,\ldots,n-1\} : \Delta v_j \neq 0\right\}$. Thus, by Result A-3, which has been proved above, the distribution under $\tilde{v}$, $\tilde{F}$, must be more skewed than that under v, F. Now consider the effect on skewness of a prize schedule change from $\hat{v}$ to $\tilde{v}$. By construction of $\tilde{v}$, there does *not* exist any i such that $\Delta\hat{v}_i = 0$ while $\Delta\tilde{v}_i \neq 0$. Also, by construction and the condition that $\min[J'] > \max[D_n\backslash(J\bigcup J')]$, $\Delta\tilde{v}_i/\Delta\hat{v}_i$ is nonincreasing for all $i \in \left\{j \in \{1,\ldots,n-1\} : \Delta\hat{v}_j \neq 0\right\}$. Thus, by Result A-3, $\tilde{F}$ must be less skewed than the distribution under $\hat{v}$, $\hat{F}$. Thus, given that $\tilde{F}$ is more skewed than F, by the transitivity of van Zwet skewness order, $\hat{F}$ must be more skewed than F. □

Now we use Results A-3 and A-4 to establish the theorem. First, consider the case in which $J' = \emptyset$. In this case, if, after removing the possible $\frac{0}{0}$ terms in the sequence, $(\Delta\hat{v}_i/\Delta v_i)_{i=1}^{n-1}$, the remaining sequence is nondecreasing, it must be that, for all $i \in D_n\backslash J$, $\Delta\hat{v}_i/\Delta v_i$ is nondecreasing. Thus, by Result A-3, $\hat{F}$ is more skewed than F. Next, consider the case in which $J' \neq \emptyset$. By our convention that ratios for which the denominator is zero and the numerator is positive equal infinity, if, after removing the possible $\frac{0}{0}$ terms in the sequence, $(\Delta\hat{v}_i/\Delta v_i)_{i=1}^{n-1}$, the remaining sequence is nondecreasing, it must be that $\min[J'] > \max[D_n\backslash(J\bigcup J')]$ and that, for all $i \in D_n\backslash(J\bigcup J')$, $\Delta\hat{v}_i/\Delta v_i$ is nondecreasing. Thus, by Result A-4, $\hat{F}$ is more skewed than F. □

Proof of Corollary 5. In WTA contests, all the prize differences are vanishing except the last whereas, in elimination contests, all the prize differences are vanishing except the first. The result then follows from a straightforward application of Theorem 5. □

Proof of Proposition 5. When h is convex, the sequence, $(\Delta v_i^h/\Delta v_i)_{i=1}^{n-1}$, must be nondecreasing. Thus, by Theorem 5, the equilibrium performance distribution under v^h is more skewed in the sense of van Zwet than that under v. Because convexification does not change mean performance and also because, when the compared distributions have equal mean, the van Zwet skewness order implies the convex order, the equilibrium performance distribution under v^h must also be more dispersed in the sense of convex order than that under v. □

Proof of Proposition 6. Let v be the prize vector with m identical winner prizes and $n-m$ identical loser prizes. Let v^e be the prize vector with m identical winner prizes and $n-m+e$ identical loser prizes. Note that v has only one non-vanishing prize difference, which is $\Delta v_{n-m} = v_{n-m+1} - v_{n-m} > 0$, and v^e also has only one non-vanishing prize difference, which is $\Delta v^e_{n+e-m} = \Delta v_{n-m}$. Thus, by equation (8),

$$\frac{q^e(p)}{q(p)} = c\,p^e, \tag{A-36}$$

where c is a positive scaler and q and q^e represent the quantile density functions for the contest before and after adding contestants, respectively. By (A-36), for $e > 0$, $p \hookrightarrow \frac{q^e(p)}{q(p)}$ is increasing. The result then follows immediately from (A-31). □

Appendix B: An example in which adding contestants does not increase performance skewness

Consider a contest with 8 contestants and a prize schedule, v, given by

$$v = \left(0, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{2}{9}, \frac{2}{9}\right). \tag{B-1}$$

After adding one contestant into this contest, the new prize schedule, v^e, satisfies

$$v^e = \left(0, 0, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{2}{9}, \frac{2}{9}\right). \tag{B-2}$$

Normalize capacity, μ, to 1. The PDFs associated with the original prize schedule, f, and the new prize schedule, f^e, are plotted in Figure B-1.

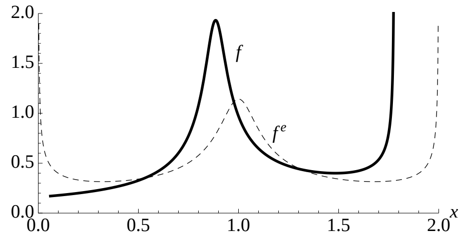


Figure B-1: The PDFs associated with the original and the new prize schedule

Intuitively, if there is an increase in the skewness of a distribution in the sense of van Zwet, there will be an extension of the upper part of the distribution together with a contraction of the lower part. However, in Figure B-1, f^e, the performance PDF under v^e, extends the upper part of f, the performance PDF under v, with a spread in the lower part. This suggests that f^e is not more skewed in the sense of van Zwet than f. Below we verify this result by referring to the definition of van Zwet order.

Note that F^e, the performance CDF under v^e, is more skewed than F, the performance CDF under v, in the sense of van Zwet if and only if $(F^e)^{-1} \circ F$ is convex. Expressed in terms of the quantile functions associated with these distribution functions, this condition becomes that $Q^e \circ Q^{-1}$ is convex. Next, note that the graph of $x \hookrightarrow Q^e \circ Q^{-1}(x)$ is given by $\{(x,y) : y = Q^e \circ Q^{-1}(x)\}$. Letting $x = Q(p)$, we see that the graph of $x \hookrightarrow Q^e \circ Q^{-1}(x)$ is also given by $\{(Q(p), Q^e(p)) : p \in (0,1)\}$, the image of $p \hookrightarrow (Q(p), Q^e(p))$, $p \in (0,1)$. Thus, for F^e to be more skewed than F, it must be that the graph of $\{(Q(p), Q^e(p)) : p \in (0,1)\}$ represents a convex function. In Figure B-2, we plot this graph for the prize schedules in the example with v and v^e given by (B-1) and (B-2), respectively. The graphed function in Figure B-2 is clearly not convex. Thus, F^e is not more skewed than F in the sense of van Zwet in the example.

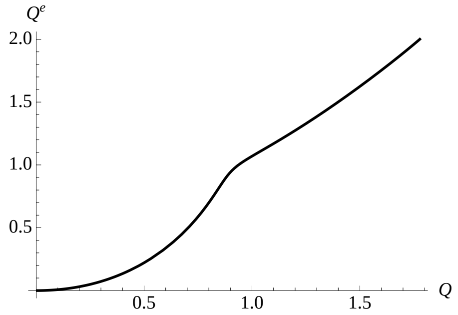


Figure B-2: Parametric quantile plot of performance distributions associated with the original and the new prize schedule

Appendix C: Tail behavior of the performance distribution

In this appendix, we investigate the tail behavior of the equilibrium performance distribution. To shorten our discussion, we employ the following definition.

Definition 4. For a PDF, f, of a distribution, F, with support given by the interval $[a,b]$, we will say that f is *initially increasing (decreasing)* if there exists a neighborhood of a on which f is increasing (decreasing). We will say that f is *ultimately increasing (decreasing)* if there exists a neighborhood of b on which f is increasing (decreasing).

The following result characterizes the tail behavior of the equilibrium PDF.

Proposition C-1 (Tail behavior of PDF)**.** Suppose $n \geq 3$ (otherwise, the equilibrium distribution is uniform). Define Δv_i by equation (9). Let $K = \{k \in \{1, \ldots n-2\} : \Delta v_{k+1} - \Delta v_k \neq 0\}$. If $K \neq \emptyset$, let $\underline{k} = \min(K)$ and $\bar{k} = \max(K)$. The following results hold:

a. If $K = \emptyset$, the equilibrium performance distribution is uniform.
b. If $\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}} > (<) 0$, the equilibrium PDF is initially decreasing (increasing).
c. If $\Delta v_{\bar{k}+1} - \Delta v_{\bar{k}} > (<) 0$, the equilibrium PDF is ultimately decreasing (increasing).

Proof. (a): Immediate from equation (10).

(b) and (c): The proof is based on identifying the sign of q_v' given by equation (10). Since in (10), $\mu\, n(n-1)/V$ is a positive scaler, which does not affect the sign of q_v', without loss of generality, we assume that $\mu\, n(n-1)/V = 1$.

When $K \neq \emptyset$, by the definitions of $\underline{k}$ and $\bar{k}$, $\Delta v_{k+1} - \Delta v_k = 0$ if $k < \underline{k}$ or $k > \bar{k}$. Thus, by equation (10) and the assumption we made that $\mu\, n(n-1)/V = 1$,

$$q_v'(p) = \sum_{k=\underline{k}-1}^{\bar{k}-1} (\Delta v_{k+2} - \Delta v_{k+1}) \binom{n-2}{k} (n-2-k) p^k (1-p)^{n-3-k}. \tag{C-1}$$

Note that q_v' given by (C-1) can be expanded as a polynomial of p and the lowest power of p is $\underline{k}-1$, with the associated coefficient being $\binom{n-2}{\underline{k}-1}(n-1-\underline{k})\left(\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}}\right)$, i.e.,

$$q_v'(p) = \binom{n-2}{\underline{k}-1}(n-1-\underline{k})\left(\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}}\right) p^{\underline{k}-1} + o(p^{\underline{k}-1}). \tag{C-2}$$

By writing the part, p^k, as $(1-(1-p))^k$ in equation (C-1), we see that q_v' given by (C-1) can also be expanded as a polynomial of $1-p$ and the lowest power of $1-p$ is $n-2-\bar{k}$, with the associated coefficient being $\binom{n-2}{\bar{k}-1}(n-1-\bar{k})\left(\Delta v_{\bar{k}+1} - \Delta v_{\bar{k}}\right)$, i.e.,

$$q_v'(p) = \binom{n-2}{\bar{k}-1}(n-1-\bar{k})\left(\Delta v_{\bar{k}+1} - \Delta v_{\bar{k}}\right) (1-p)^{n-2-\bar{k}} + o((1-p)^{n-2-\bar{k}}). \tag{C-3}$$

Equation (C-2) implies that, if $\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}} > 0$, then q_v' is initially positive and thus q_v is initially increasing and, if $\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}} < 0$, then q_v' is initially negative and thus q_v is initially decreasing. Similarly, equation (C-3) implies that, if $\Delta v_{\bar{k}+1} - \Delta v_{\bar{k}} > 0$, then q_v is ultimately increasing and, if $\Delta v_{\bar{k}+1} - \Delta v_{\bar{k}} < 0$, then q_v is ultimately decreasing. By the inverse function theorem, $f_v = 1/(q_v \circ F_v)$. Thus, q_v being initially (ultimately) increasing (decreasing) implies that f_v is initially (ultimately) decreasing (increasing). This completes the proof of parts (b) and (c). □

Proposition C-1 implies that the tail behavior of the equilibrium performance PDF is determined by the first, $\underline{k}$, and the last, $\bar{k}$, non-vanishing second difference of the prize schedule. These second differences determine the local convexity or concavity of the quantile function and thus the equilibrium CDF around the two endpoints of the support. For example, if $\Delta v_{\underline{k}+1} - \Delta v_{\underline{k}} > 0$, then the quantile function exhibits convexity around 0, and thus, its inverse, the CDF of the performance distribution, exhibits concavity around its lower endpoint, i.e., the PDF is initially decreasing.

Note that, because prizes are nonconstant and nondecreasing in rank, if $v_1 = v_2$, the first non-vanishing second difference must exist and be positive, and thus the PDF must be initially decreasing. Similarly, if $v_{n-1} = v_n$, the PDF must be ultimately increasing. Thus, if the prize schedule has a flat end, the equilibrium performance PDF will never be inverse U-shaped, i.e., the performance distribution will never be unimodal.

Appendix D: Moments of the equilibrium performance distribution

Let Q_v be the quantile function associated with the equilibrium performance distribution under the prize vector v. The expression for Q_v is given by Corollary 1. Note that the equilibrium random performance is equal in distribution to $Q_v(\tilde{U})$, where $\tilde{U}$ is uniformly distributed on $[0,1]$. Thus, if X represents equilibrium random performance, then for any continuous function $g : \mathbb{R} \mapsto \mathbb{R}$, we have

$$\mathbb{E}[g(X)] = \mathbb{E}[g \circ Q_v(\tilde{U})] = \int_0^1 g \circ Q_v(p) dp. \tag{D-1}$$

We hence obtain the following result.

Lemma D-1. Let Q_v be defined in Corollary 1. The nth raw moment of the equilibrium performance distribution satisfies

$$\mathbb{E}[X^n] = \int_0^1 Q_v(p)^n dp. \tag{D-2}$$

The nth central moment of the equilibrium performance distribution satisfies

$$\mathbb{E}[(X-\mu)^n] = \int_0^1 (Q_v(p) - \mu)^n dp. \tag{D-3}$$

The nth standardized moment of the equilibrium performance distribution satisfies

$$\frac{\mathbb{E}[(X-\mu)^n]}{(\mathbb{E}[(X-\mu)^2])^{\frac{n}{2}}} = \frac{\int_0^1 (Q_v(p) - \mu)^n dp}{\left(\int_0^1 (Q_v(p) - \mu)^2 dp\right)^{n/2}}. \tag{D-4}$$

Appendix E: Endogenizing mean performance

In our risk-taking contest model, contestants can choose any distribution of non-negative random performance subject only to a capacity constraint on mean performance. In this appendix, we consider an extension of this model which enables us to endogenize capacity and, thus, mean performance.

Suppose, before contestants simultaneously choose their capacity-constrained performance distributions, they have to first build their capacity through costly investment/effort. Assume that contestants have the same cost function, c, which is increasing, weakly convex, and continuously differentiable in capacity. Assume that contestants do not observe each other's capacity choice when they make their risk-taking decisions.

The following lemma shows that this extended model has a symmetric equilibrium in which all the contestants choose the same capacity level in the capacity-building stage, followed by the same distributional choice in the risk-taking stage, and the equilibrium capacity is determined entirely by per capita real gains.

Lemma E-1. The extended model where capacity is endogenous has a unique symmetric equilibrium. In this equilibrium, in the capacity-building stage, all the contestants choose the same capacity level, $\mu > 0$, which is implicitly and uniquely determined by

$$\mu c'(\mu) = \frac{V}{n}, \tag{E-1}$$

where $V = \sum_{i=1}^{n}(v_i - v_1)$ represents total real gains and thus, V/n represents per capita real gains. In the risk-taking stage, all the contestants choose the same performance distribution, F_v, given by Theorem 1, with μ endogenized through equation (E-1).

Proof. We show that the prescribed strategy profile indeed sustains an equilibrium. For any given contestant, if all of her $n-1$ rivals play the prescribed strategy, then, by Proposition 1, this contestant will face a contest payoff function given by

$$P(x) = \min\left[v_1 + (v_n - v_1)\frac{x}{\bar{x}}, v_n\right], \tag{E-2}$$

where $\bar{x} = \frac{\mu n(v_n - v_1)}{\sum_{i=1}^{n}(v_i - v_1)}$. Note that this contest payoff function is weakly concave, implying that, conditioned on any capacity choice made by this contestant, it is always a best response for her to play safe in the risk-taking stage. Thus, if she chooses capacity μ' in the capacity-building stage, then by playing a best response to the prescribed P in the risk-taking stage, she will earn a prize equal to $P(\mu')$ in expectation. Hence, in the capacity-building stage, her problem is to choose μ' to maximize

$$P(\mu') - c(\mu').$$

In a symmetric equilibrium, μ' must equal μ and must satisfy the first-order condition, $P'(\mu) - c'(\mu) = 0$. Thus, by equation (E-2), we have $\frac{V}{\mu n} = c'(\mu)$, and equation (E-1) hence follows.

To see that μ is uniquely determined by equation (E-1), note that, because c is increasing, weakly convex, and continuously differentiable, c' must be positive, nondecreasing, and continuous. Thus, the left hand side of equation (E-1), $\mu \hookrightarrow \mu c'(\mu)$, must be increasing and continuous, go to 0 as $\mu \to 0$, and go to infinity as $\mu \to \infty$. Hence, given that the right hand side of equation (E-1) is a positive constant, there must exist a unique μ that satisfies equation (E-1). □

Appendix F: Other experimental results

Bootstrap Kolmogorov-Smirnov test

To compare the overall distribution among treatments and against predictions, we employ a non-parametric *Kolmogorov-Smirnov* (KS) test. The KS test statistic, denoted as D, measures the maximum discrepancy between two cumulative distributions at each point x along the distribution, indicating how different the two distributions are.[36]

The standard KS test assumes distributions are generated by independent and identically distributed observations. However, the distributions we observe do not meet this assumption: each marker is part of a set of 100 markers allocated by a participant, and a participant's choice to allocate one marker depends on their choices for their other markers due to the budget constraint. Therefore, we employ the bootstrap test, generating a distribution of KS statistics through 10,000 iterations of the following process.

Consider two treatments, each with n observations. Let F_1 and F_2 be the aggregate distributions from Treatment 1 and Treatment 2, respectively, and let $\overline{F}$ be the distribution obtained by pooling these two distributions. The null hypothesis H_0 is defined as $H_0 : F_1 = F_2 = \overline{F}$. This null hypothesis suggests that if there is no difference between the distributions from these two treatments, they can be regarded as independent samples drawn from the same pooled distribution.

In each iteration, we randomly reassign the treatment labels to the observations in $\overline{F}$ to create new treatment groups F_1^p and F_2^p. Subsequently, we calculate the KS statistic D between F_1^p and F_2^p. This process is repeated 10,000 times to generate a distribution of D_j ($j \in \{1, \cdots, 10{,}000\}$) drawn from $\overline{F}$.

We then calculate the KS statistics D^* between F_1 and F_2. The null hypothesis H_0 is rejected if D^* exceeds the 95_{th} percentile of the distribution of D_j. This indicates that only if D^* significantly surpasses D_j (derived from the pooled data of the two treatments) can we reject the null hypothesis that F_1 and F_2 are drawn from the same distribution. The results of this bootstrap KS test are presented in Figures F-1 and F-2.

This permutation method can be used not only to test the treatment effect but also to compare the observed distribution against the predicted distribution. To do this, we replace F_2 with the simulated distribution generated by equation (6). The results are displayed in Figures F-3 and F-4.

[36] $D = \sup_x |F_{1,n}(x) - F_{2,m}(x)|$, where $F_{1,n}$ and $F_{2,m}$ are the empirical distribution functions of sample 1 and sample 2, respectively.

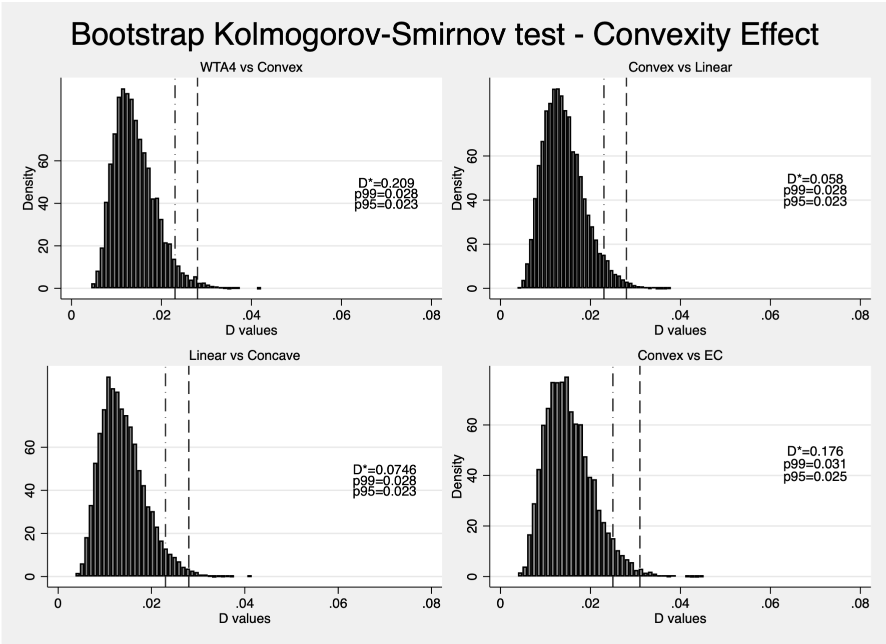


Figure F-1: *Bootstrap Kolmogorov-Smirnov test for the convexification effect* D^* represents the KS statistic between F_1 and F_2. The dash-dot line and the dashed line indicate the 95th percentile and the 99th percentile of the bootstrapped KS statistics, respectively. It is evident that in all pairwise comparisons, the null hypothesis H_0 is rejected.

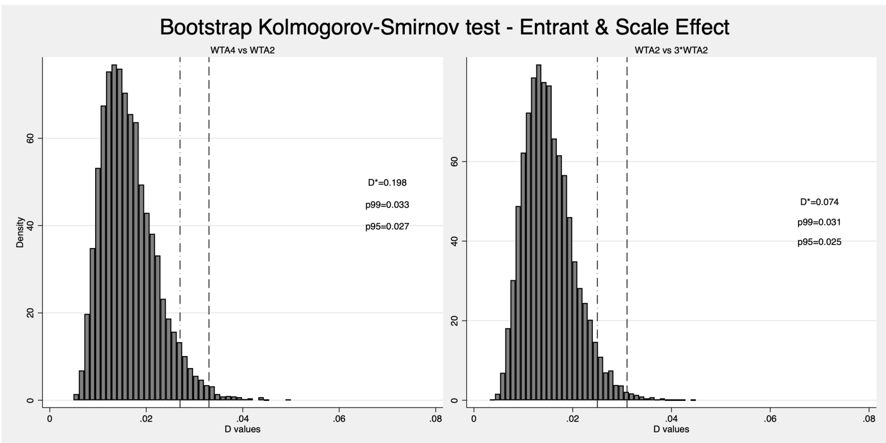


Figure F-2: *Bootstrap Kolmogorov-Smirnov test for the entrant and the scale effects* D^* represents the KS statistic between F_1 and F_2. The dash-dot line and the dashed line indicate the 95th percentile and the 99th percentile of the bootstrapped KS statistics, respectively. It is evident that in all pairwise comparisons, the null hypothesis H_0 is rejected.

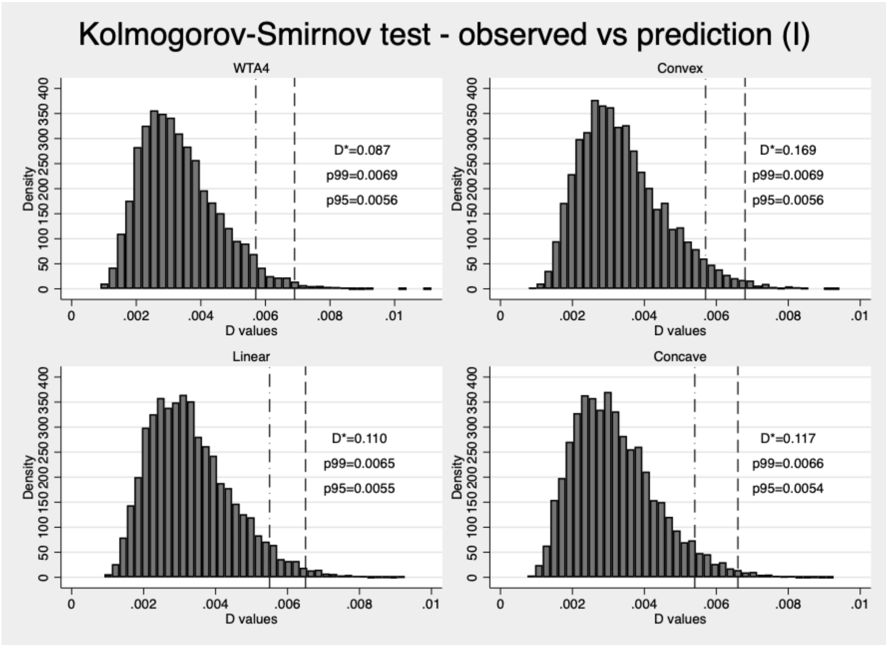


Figure F-3: *Bootstrap Kolmogorov-Smirnov test comparing the observed distribution with the predicted distribution for treatments WTA4, Convex, Linear, and Concave* D^* represents the KS statistic between F_1 and F_2. The dash-dot line and the dashed line indicate the 95th percentile and the 99th percentile of the bootstrapped KS statistics, respectively. It is evident that in all pairwise comparisons, the null hypothesis H_0 is rejected.

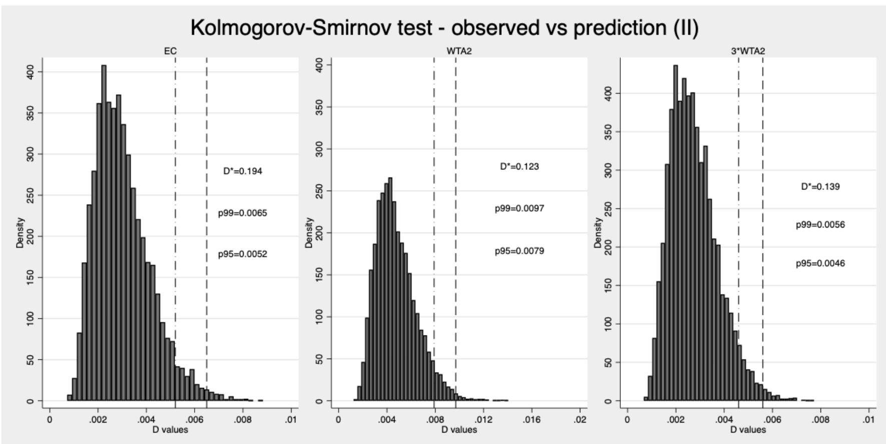


Figure F-4: *Bootstrap Kolmogorov-Smirnov test comparing the observed distribution with the predicted distribution for treatments EC, WTA2, and 3*WTA2* D^* represents the KS statistic between F_1 and F_2. The dash-dot line and the dashed line indicate the 95th percentile and the 99th percentile of the bootstrapped KS statistics, respectively. It is evident that in all pairwise comparisons, the null hypothesis H_0 is rejected.

	SD				Skewness			
	Model 1		Model 2		Model 1		Model 2	
	All rounds	Round 11-20	All rounds	Round 11-20	All rounds	Round 11-20	All rounds	Round 11-20
EC (Const.)	3.787***	3.379***	2.469***	2.659***	-0.972***	-1.217***	-1.165***	-1.524***
	(0.334)	(0.355)	(0.608)	(0.606)	(0.114)	(0.135)	(0.256)	(0.277)
Concave	0.281	0.467	0.243	0.442	0.231	0.236	0.226	0.244
	(0.472)	(0.502)	(0.476)	(0.505)	(0.161)	(0.191)	(0.160)	(0.192)
Linear	1.622***	1.777***	1.531***	1.694***	0.781***	0.792***	0.762***	0.793***
	(0.472)	(0.502)	(0.477)	(0.506)	(0.161)	(0.191)	(0.161)	(0.192)
Convex	2.124***	2.533***	2.140***	2.536***	0.830***	1.030***	0.816***	1.021***
	(0.472)	(0.502)	(0.476)	(0.505)	(0.161)	(0.191)	(0.160)	(0.192)
WTA4	5.376***	5.902***	5.360***	5.875***	1.369***	1.622***	1.341***	1.617***
	(0.472)	(0.502)	(0.479)	(0.507)	(0.161)	(0.191)	(0.162)	(0.194)
Female			0.292	0.126			0.072	-0.026
			(0.221)	(0.213)			(0.097)	(0.103)
Post_grad			0.198	0.104			-0.024	0.047
			(0.234)	(0.226)			(0.103)	(0.109)
Major.Business			0.020	-0.077			-0.010	0.103
			(0.275)	(0.267)			(0.119)	(0.127)
Major.Art			-0.088	-0.262			0.096	0.168
			(0.300)	(0.291)			(0.130)	(0.139)
Major.Others			-0.068	-0.033			-0.005	0.253
			(0.482)	(0.465)			(0.212)	(0.225)
Risk Seeking			0.594***	0.460**			0.069	0.065
			(0.213)	(0.205)			(0.096)	(0.100)
σ^2_{group}	0.744	0.985	0.798	1.018	0.033	0.089	0.035	0.091
	(0.252)	(0.282)	(0.256)	(0.284)	(0.031)	(0.042)	(0.032)	(0.043)
$\sigma^2_{individual}$	2.212	2.007	2.078	1.931	0.451	0.486	0.445	0.478
	(0.249)	(0.222)	(0.236)	(0.214)	(0.051)	(0.054)	(0.051)	(0.054)
Observations	4,800	2,400	4,800	2,400	4,800	2,400	4,800	2,400
Number of groups	60	60	60	60	60	60	60	60

Table F-1: *Regression results for the convexification effect* Multi-level mixed regression controlled for the individual and group level random effect. The control variables include Gender (Baseline = Male), Post_grad (Baseline = Undergraduate), Major (Baseline = Science and Engineer), Risk Seeking (Baseline = risk-averse). Risk is measured on a scale from 0 to 10, with a risk tolerance of 0 representing the lowest and a tolerance of 10 representing the highest. Individuals with a risk score no greater than 5 are categorized as risk-averse, while those with a score greater than 5 are categorized as risk-seeking. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	SD				Skewness			
	Model 1		Model 2		Model 1		Model 2	
	All rounds	Round 11-20	All rounds	Round 11-20	All rounds	Round 11-20	All rounds	Round 11-20
WTA2 (Const.)	6.204***	6.171***	5.955***	5.836***	-0.159	-0.298***	-0.334**	-0.497***
	(0.409)	(0.423)	(0.557)	(0.596)	(0.102)	(0.100)	(0.158)	(0.145)
WTA4	2.958***	3.109***	3.116***	3.254***	0.555***	0.704***	0.638***	0.778***
	(0.553)	(0.569)	(0.564)	(0.573)	(0.131)	(0.129)	(0.136)	(0.123)
Female			-0.400	-0.085			0.160	0.105
			(0.330)	(0.375)			(0.110)	(0.103)
Post_grad			0.040	0.006			0.088	0.057
			(0.384)	(0.433)			(0.125)	(0.116)
Major.Bus			0.299	0.239			-0.086	0.009
			(0.441)	(0.490)			(0.136)	(0.126)
Major.Art			0.382	0.156			-0.238*	-0.191
			(0.447)	(0.499)			(0.141)	(0.131)
Major.Others			-0.183	-0.272			-0.135	0.050
			(0.759)	(0.856)			(0.247)	(0.228)
Risk Seeking			0.481	0.441			0.242**	0.275***
			(0.316)	(0.358)			(0.105)	(0.098)
σ^2_{group}	1.319	1.314	1.356	1.249	0.039	0.069	0.001	0.012
	(0.535)	(0.549)	(0.551)	(0.547)	(0.030)	(0.038)	(0.021)	(0.019)
$\sigma^2_{individual}$	1.203	1.495	1.078	1.465	0.210	0.200	0.202	0.171
	(0.281)	(0.337)	(0.259)	(0.336)	(0.039)	(0.038)	(0.038)	(0.031)
Observations	1,440	720	1,440	720	1,440	720	1,440	720
Number of groups	24	24	24	24	24	24	24	24

Table F-2: *Regression results for the entrant effect* Multi-level mixed regression controlled for the individual and group level random effect. The control variables include Gender (Baseline = Male), Post_grad (Baseline = Undergraduate), Major (Baseline = Science and Engineer), Risk Seeking (Baseline = risk-averse). Risk is measured on a scale from 0 to 10, with a risk tolerance of 0 representing the lowest and a tolerance of 10 representing the highest. Individuals with a risk score no greater than 5 are categorized as risk-averse, while those with a score greater than 5 are categorized as risk-seeking. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	SD				Skewness			
	Model 1		Model 2		Model 1		Model 2	
	All rounds	Round 11-20	All rounds	Round 11-20	All rounds	Round 11-20	All rounds	Round 11-20
WTA2 (Const.)	6.204***	6.171***	5.584***	5.325***	-0.159	-0.298***	-0.381	-0.469*
	(0.309)	(0.294)	(0.635)	(0.616)	(0.105)	(0.101)	(0.270)	(0.254)
3*WTA2	-0.218	-0.323	-0.253	-0.316	-0.257**	-0.268**	-0.259**	-0.271**
	(0.401)	(0.379)	(0.395)	(0.367)	(0.126)	(0.123)	(0.118)	(0.119)
Female			-0.079	-0.089			0.229**	0.163*
			(0.230)	(0.223)			(0.102)	(0.095)
Post_grad			-0.348	-0.245			-0.115	-0.128
			(0.264)	(0.255)			(0.116)	(0.108)
Major.Bus			0.073	0.215			-0.092	0.017
			(0.310)	(0.298)			(0.131)	(0.123)
Major.Art			0.331	0.214			0.079	0.101
			(0.321)	(0.309)			(0.136)	(0.129)
Major.Others			-0.511	-0.754			-0.179	-0.190
			(0.603)	(0.584)			(0.263)	(0.247)
Risk seeking			0.528**	0.707***			0.089	0.131
			(0.218)	(0.212)			(0.098)	(0.091)
σ^2_{group}	0.607	0.510	0.564	0.409	0.020	0.027	0.001	0.012
	(0.324)	(0.271)	(0.310)	0.244	(0.027)	(0.023)	(0.021)	(0.019)
$\sigma^2_{individual}$	0.944	0.919	0.846	0.795	0.208	0.176	0.202	0.171
	(0.188)	(0.181)	(0.173)	(0.163)	(0.038)	(0.031)	(0.038)	(0.031)
Observations	1,920	960	1,920	960	1,920	960	1,920	960
Number of groups	24	24	24	24	24	24	24	24

Table F-3: *Regression results for the scale effect* Multi-level mixed regression controlled for the individual and group level random effect. The control variables include Gender (Baseline = Male), Post_grad (Baseline = Undergraduate), Major (Baseline = Science and Engineer), Risk Seeking (Baseline = risk-averse). Risk is measured on a scale from 0 to 10, with a risk tolerance of 0 representing the lowest and a tolerance of 10 representing the highest. Individuals with a risk score no greater than 5 are categorized as risk-averse, while those with a score greater than 5 are categorized as risk-seeking. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix G: Experimental instructions

The experiment was conducted in Chinese. The following instruction is an English translation of the instruction used in the WTA4 treatment of the experiment. Apart from the description of the prize schedule (and the number of participants in a group if this number is not 4), the instructions for all other treatments were the same. The original Chinese instruction can be found in the online appendix.

Experimental instructions used in the WTA4 treatment

Thank you very much for participating in today's experiment! Since you arrived on time, you will receive a 20 RMB participation fee. Please carefully read the experimental instructions in your hand. They will help you make better decisions during the experiment to achieve higher earnings. The experiment is expected to last 1.5 hours. We will use experimental currency (EC) as the unit of measurement for earnings. At the end of the experiment, we will convert your earnings into RMB and pay you.

Your total compensation for this experiment will be the sum of:

(1) Your participation fee of 20 RMB;
(2) Your earnings from the experimental decision-making part.

To make sure you understand this experiment, our lab assistant will read the experimental instructions aloud before the experiment begins and will be available to provide assistance throughout the experiment. Please remain quiet during the experiment and do not discuss with other participants. If you cannot adhere to this rule, we will ask you to leave the experiment, and you will not receive compensation for your earnings from the experimental decision-making part. If you have any questions during the experiment, please raise your hand. Our lab assistant will assist you.

The Experiment

This experiment consists of 20 rounds of decision-making. Before the experiment begins, the computer will randomly assign you to a group of four participants and randomly number each member. In each round, you will compete for a set of prizes with the other members of your group. The composition of your group, each member's random number within the group, and the prize settings used in this experiment will remain the same throughout all 20 rounds.

The prize settings for this experiment are shown in the figure below:

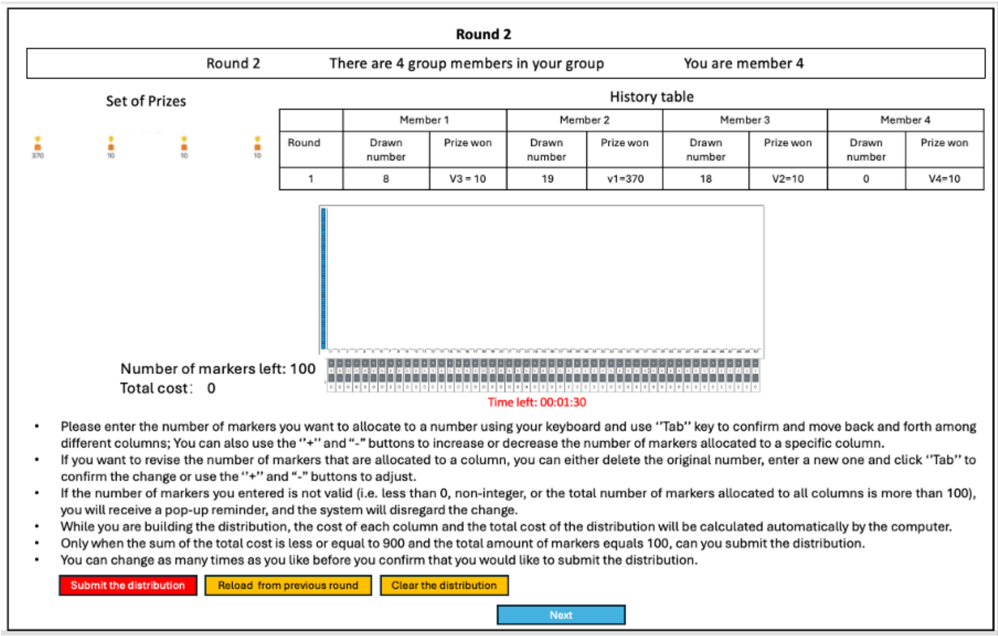
1	2	3	4
370	10	10	10

In the first row, the numbers on the trophies indicate the rankings within your group (i.e., from 1st place to 4th place). The second row shows the prizes corresponding to each rank. The value of the prizes is measured in EC. As illustrated, if you rank first in your group, you will receive 370 EC units. If you rank second, third, or fourth, you will receive 10 EC units.

Your Task

In each round of the competition, you need to allocate 100 markers to integers from 0 to 50 (a total of 51 numbers) to construct a probability distribution. Each marker represents a 1% probability of being drawn. The distribution you build will determine the probability of each number being drawn. Therefore, the more markers you place on a particular number, the higher the probability that this number will be drawn. The computer will randomly draw one number out of the 51 integers according to the distribution you build. Additionally, the computer will also draw a number for each of the other group members based on their respective distributions. Group members will be ranked from highest to lowest based on the numbers drawn, determining the prize allocation for that round. For example, if your drawn number is 14 and the other members' drawn numbers are 10, 20, and 0, respectively, your rank in the group would be second, and you would receive the 2nd place prize (i.e., 10 EC units) for that round.

In the experiment, you will use the computer interface (the distribution builder) shown below to construct your probability distribution:



(This is an English translation of the original screen.)

When constructing the distribution, you will face the following budget constraint: each marker placed on a different number will incur a different cost. Although increasing the number of markers on larger numbers can increase your probability of winning higher prizes, it will also correspondingly consume more of your budget. The total cost of the 100 markers must be less than or equal to 900. The specific cost calculation method is as follows:

The cost of each marker placed on a number = the value of that number

The total cost of all markers placed on each number = the total number of markers placed on that number * the cost of each marker placed on that number

For example, if you decide to place 10 markers on number 10, which will cost you 100 in total, there will be a 10% probability that 10 is drawn by the computer. Please note that **the cost of all markers placed on the number 0 is zero. If the total cost of the distribution you construct is less than 900, the remaining budget will be forfeited and will not be added to your earnings.**

As shown in the figure above, you can enter the number of markers you want to allocate to a number using your keyboard, then press " Tab " key to confirm each entry and move between columns. You can also use the " + " and " - " buttons to increase or decrease the number of markers allocated to each column. If you want to change the number of markers allocated to a particular column, first delete the current number, then re-enter the desired number, and press the " Tab " key to confirm the change in that column. Alternatively, you can directly use the " + " and " - " buttons to adjust the number of markers. If the number of markers you enter in any column exceeds the valid range (i.e., less than 0, non-integer, or the total number of markers in the allocation table exceeds 100), the system will automatically prompt you and reset the number of markers in that column to its previous value. During the distribution construction process, the computer will automatically calculate the cost spent on each number and the total cost. **You can submit the distribution only when the total number of markers in the probability distribution table equals 100 and the total cost does not exceed 900.** You can make unlimited modifications before confirming the submission.

To facilitate constructing and adjusting your distribution, we provide an option to "Load Previous Round's Distribution." You can choose (but are not required) to use this button to adjust and build a new probability distribution based on your submission from the previous round (except for the first round). You can also use the "Clear Distribution" button at any time to clear the existing distribution and start fresh. Above the distribution builder, we provide a history table containing the results of all previous rounds (except for the first round) to help you make better decisions.

After the distributions are submitted, the computer will automatically draw one number for each group member according to their submitted distribution and rank them within the group. In the case of a tie, the computer will break the tie randomly. You will observe the computer using a lottery wheel to draw a number for you. You will also receive information about your own as well as your group members' drawn numbers, rankings, and the prize allocation at the end of each round.

Your Earnings

At the end of the 20 rounds, the computer will randomly select one round as the payment round. The prize you receive in that round will be converted to RMB at an exchange rate of **1 EC = 0.45 RMB** and will be paid to you along with your participation fee at the end of the experiment. All information about the payment round (which round was selected, the numbers drawn for you and the other group members in that round, your ranking, and the prize awarded) will be displayed on your computer screen.

Pre-experiment Questionnaire and Practice Round

In order to ensure that you fully understand the experiment, you need to answer a few test questions related to the experimental instructions before the experiment begins.

After all of you answer the test questions correctly, you will enter a practice interface. This interface will help you further understand how to use the distribution builder and how the computer will draw numbers using a lottery wheel. This practice round will NOT be included in the formal experiment and will NOT affect your earnings today. If you have any questions about the experimental instructions or the test questions, please seek assistance from our lab assistant promptly.

Thank you again for your participation and patience! The experiment will start soon...

Post-experiment Questionnaire

Finally, after all experimental decisions have been completed, you will be asked to fill out a brief information collection questionnaire, including questions about your gender, education level, major, and decisions made during the experiment. Please rest assured that all information you provide in this experiment is anonymous and strictly confidential, and the data will only be used for academic research analysis.