

# Turning Up the Heat: The Discouraging Effect of Competition in Contests

---

Dawei Fang

*University of Gothenburg*

Thomas Noe

*University of Oxford*

Philipp Strack

*Yale University*

We study contests in which contestants are homogeneous and have convex effort costs. Increasing contest competitiveness, by making prizes more unequal, scaling up the competition, or adding new contestants, always discourages effort. These results have significant implications: although often criticized as evidence of laxity or cronyism, muting competition (e.g., adopting softer grading curves or less high-powered promotion systems) can both reduce inequality and increase output. Holding promotion contests at the division level rather than the firm level can boost employees' effort. Our results are also consistent with personnel policies that feature egalitarian pay systems and dismissal of worst-performing employees.

We thank Michael Baye, Paul Heidhues, Kai Konrad, Dan Kovenock, Jingfeng Lu, Benny Moldovanu, John Quah, Aner Sela, Xianwen Shi, Ron Siegel, Sebastian Sieglösch, Jun Xiao, and seminar audiences at the Bank of Finland, Bergen, Berkeley, Berlin, Bonn, Caltech, Chapman, Durham University, Gothenburg, Harvard/Massachusetts Institute of Technology, Munich, Regensburg, and various conferences for helpful comments and suggestions. We would like to also thank the editor, Emir Kamenica, and reviewers of this journal for extremely insightful comments, which have greatly contributed to the presentation of our results. Fang gratefully acknowledges financial support from Browaldh Stiftelsen and Vinnova (grant 2010-02449).

Electronically published April 15, 2020

[*Journal of Political Economy*, 2020, vol. 128, no. 5]

© 2020 by The University of Chicago. All rights reserved. 0022-3808/2020/12805-0009\$10.00

## I. Introduction

Contests, by allocating prizes based on relative performance, create competition between contestants. Naturally, increasing competition increases contestants' incentives to exert high effort and reach for the top, even when doing so comes at great cost. However, there is also an indirect discouragement effect: if some contestants reach for the top, it is harder for their rivals to beat them, which can reduce their rivals' gains from exerting effort. Thus, sometimes, contestants may give up and exert little or no effort. We show that, perhaps surprisingly, in settings where contestants are homogeneous and have convex effort costs, this indirect effect always dominates the direct effect—increased competition always reduces expected effort of individual contestants.

In our setup, each contestant simultaneously chooses a costly effort level. The contestant with the highest effort wins the largest prize, and so forth until all the prizes are allocated, with the worst performer always receiving no prize.<sup>1</sup> The competitiveness of the competition depends on the number of contestants and on how contest rewards are tilted toward better performers.

We find that increased competition discourages effort under three disparate definitions of increased competition: increased prize inequality, increased contest scale, and contestant entry. Increased prize inequality results from transferring value from lower- to higher-ranked prizes, for example, offering the entire bonus budget to the salesperson with the highest sales revenue instead of splitting it between a group of best performers. Increased contest scale results from proportionally scaling up the number of contestants and the number of prizes, for example, fixing the promotion rate while holding promotion contests at the organizational level instead of the division level. Entry results from adding contestants without adding prizes, for example, adding outsiders to the competition for a firm's senior position.

In equilibrium, each contestant chooses a random effort. Increasing competition, through increasing prize inequality or contest scale, increases the likelihood that contestants will choose either very high or very low effort levels. We show that, when effort costs are convex, the reduction in effort caused by the increased likelihood of low effort levels exceeds the increase in effort caused by the increased likelihood of high effort levels, and consequently, more "spread-out" effort distributions yield lower expected effort. Entry by new contestants has a more profound adverse effect on

<sup>1</sup> Throughout this paper, "no prize" denotes a prize of zero. Because the worst performer receives no prize, there always exists some competition between the contestants for the positive prizes offered in the contest. Contests in which the value of the worst performer's prize is nonzero can be translated into our setup simply by redefining the value of each prize to be its value in excess of the worst performer's prize.

individual effort because entry lowers the average reward received by the contestants and thus the cost of effort contestants are willing to incur.

Our results appear to be consistent with the anomalously uncompetitive characteristics of many actual contests, for example, coarse grading schemes in universities, weak compensation-performance relationships, and sales contests that reward the majority of salespeople.<sup>2</sup> Our results for contest scale predict that “forced-ranking systems,” which mandate the percentage of employees assigned a given grade, will be more effective in motivating effort when the comparison groups are smaller.<sup>3</sup>

*Related literature.*—We contribute to the literature on deterministic (noiseless) contests, that is, so-called all-pay contests, by showing that increasing competition, whether defined by increased prize inequality, increased contest scale, or entry, always reduces each contestant’s expected effort. The class of environments we consider is restrictive along one dimension—we assume that contestants are homogeneous—but quite general along other dimensions. It encompasses contests with an arbitrary number of contestants, general prize structures, and arbitrary convex costs of effort.

Our analysis is built upon the all-pay-contest literature, which aims to characterize how contestant heterogeneity and effort-cost convexity affect the relationship between prize structure and contestant effort. The root of this literature is the seminal work by Barut and Kovenock (1998). Barut and Kovenock establish an irrelevance result: in all-pay contests with complete information (regarding abilities), homogeneous contestants, and linear effort costs and with the sum of prizes fixed, all prize structures that offer the worst performer a zero prize produce the same expected total effort. Thus, absent contestant heterogeneity and cost convexity, the expected effort of individual contestants is not affected by prize inequality or contest scale. Because expected total effort is fixed by the sum of prizes, adding contestants through contest entry does have a negative effect on individual contestant effort.

Subsequent research branches off by modifying one or more of the assumptions in Barut and Kovenock (1998) and focuses extensively on identifying optimal (effort-maximizing) distributions of prizes. Introducing contestant heterogeneity while retaining the complete-information assumption, even with linear effort costs, produces a fairly intractable setting. Xiao (2018) shows that in the case of complete information, linear

<sup>2</sup> Table 1 in Baker, Jensen, and Murphy (1988) shows that the salary difference between managers receiving the highest evaluation rating and those receiving the lowest rating is less than 8%. Churchill, Ford, and Walker (1997) report that some firms employ sales contests that reward more than 80% of salespeople.

<sup>3</sup> Forced-ranking systems implement a contest in which employees compete for performance grades (Gladwell 2002). Survey evidence suggests that at least 20% of large corporations use these systems (Bates 2003).

costs, and three contestants, offering two equal prizes is optimal if the two weakest contestants are relatively similar, whereas winner-take-all contests are optimal otherwise. However, extending the optimal contest design analysis in Xiao (2018) to the general  $n$ -contestant case is infeasible.<sup>4</sup>

More determinant results can be obtained by either assuming incomplete information or examining asymptotic contestant behavior. In an incomplete-information setting, where contestants are heterogeneous ex post but not ex ante, Moldovanu and Sela (2001) show that winner-take-all contests are optimal when effort costs are linear or concave and are never optimal when costs are sufficiently convex.<sup>5</sup> However, even in Moldovanu and Sela's setting, when costs are strictly convex, general characterizations of the relationship between prize structure and contestant effort are difficult to obtain.

Olszewski and Siegel (2018c) are able to characterize optimal contest designs when contestants are ex post heterogeneous (contestants can be ex ante symmetric or asymmetric, and information can be complete or incomplete) by considering the limit of contestant behavior as the number of contestants goes to infinity. They show that winner-take-all contests are optimal when costs are linear or concave, but when costs are strictly convex, offering many heterogeneous prizes is optimal.<sup>6</sup>

Our paper, like these papers, is rooted in Barut and Kovenock (1998) but considers the effect of convex effort costs without introducing contestant heterogeneity. Because we analyze contests with a finite number of contestants rather than the limit of contestant behavior, we are able to separate the effect of the number of contestants from the effect of the distribution of prizes. This separation is productive. In particular, we identify increased prize inequality and increased contest scale, two orthogonal prize-structure changes, one operating on the distribution of prizes with

<sup>4</sup> We provide very limited analysis of the optimal design of asymmetric complete-information all-pay contests with at least three contestants in sec. S.4 of the online appendix. We show that, in these contests, when there are at least two contestants with the greatest ability, the winner-take-all structure is optimal if effort costs are linear but can be suboptimal if effort costs are strictly convex.

<sup>5</sup> In incomplete-information settings, each contestant's type (ability) is a privately observed independent random draw from a type distribution. Thus, ex post (conditioned on the draw), contestants are heterogeneous. Contestants are ex ante homogeneous (heterogeneous) if the type distribution is identical (not identical) for all contestants.

<sup>6</sup> Other factors that can render highly competitive contest designs suboptimal have been identified in various settings and under various designer objectives: contestant risk aversion (Glazer and Hassin 1988; Kalra and Shi 2001; Olszewski and Siegel 2018c), social loss aversion (Lim 2010), sabotage (Lazear 1989), learning (Halac, Kartik, and Liu 2017), heavy-tailed fluctuations (Drugov and Ryvkin 2019), participation constraints (Megidish and Sela 2013), free-riding (Chowdhury and Kim 2017), risk taking (Hvide 2002), and wasteful effort (Olszewski and Siegel 2018b). For discussions of the optimality of winner-take-all contests outside the all-pay-contest setting, see Clark and Riis (1998), Krishna and Morgan (1998), Kalra and Shi (2001), Szymanski and Valletti (2005), Schweinzer and Segev (2012), and Drugov and Ryvkin (2019).

contest size fixed and the other operating on contest size with the distribution of prizes fixed, as two distinct notions of increased competition that uniformly push effort in the same direction. They both always reduce expected individual effort and always increase effort dispersion when contestants are homogeneous and have convex effort costs. Thus, when cost convexity is introduced into the framework of Barut and Kovenock (1998), the discouraging effect of competition not only encompasses increased competition caused by entry but also extends much more broadly, to increased competition caused by increased contest scale and prize inequality.

To our knowledge, this paper is the first to provide general characterizations of the effect of contest scale in the all-pay-contest framework.<sup>7</sup> Given the existing literature, which finds that convex costs push optimal prize structures away from the most unequal prize structure, winner take all, toward more equal prize structures, one implication from our analysis, that cost convexity can render sufficiently unequal prize structures suboptimal, is not surprising. The perhaps more surprising results of our analysis of prize inequality are that (1) there exists a natural, general, prize-inequality partial order under which increased prize inequality always reduces expected effort and (2) the same partial order also ranks effort dispersion. Thus, when contestants are homogeneous and effort costs are strictly convex, winner-take-all contests are not merely suboptimal. In fact, among all contests that offer the worst performer a zero prize, winner-take-all contests induce the lowest expected effort and the highest effort dispersion, whereas offering all contestants except the worst performer identical prizes induces the highest expected effort and the lowest effort dispersion.

## II. Setup

Consider an all-pay contest between  $n \geq 2$  homogeneous risk-neutral contestants. Each contestant  $i$  simultaneously chooses an effort  $x_i \geq 0$ . By choosing  $x_i$ , contestant  $i$  incurs effort cost (i.e., disutility of effort)  $c(x_i)$ . Effort cost  $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is differentiable, strictly increasing, and convex, with  $c(0) = 0$ .<sup>8</sup>

There is an ordered vector of prizes  $v \in \mathbb{R}_+^n$ :

$$v_1 \leq v_2 \leq \dots \leq v_n,$$

<sup>7</sup> Andreoni and Brownback (2017) develop an example, under specific parametric assumptions, in which scaling up an ex ante symmetric incomplete-information all-pay contest with linear costs increases expected individual effort. Outside the all-pay-contest setting, Fu and Lu (2009) identify a positive effect of scale on effort when contest winners are selected sequentially via a lottery process. In contrast, Chowdhury and Kim (2017) find a negative effect of scale on effort when contest winners are selected simultaneously via a lottery process or contest losers are eliminated sequentially via a lottery process.

<sup>8</sup> These assumptions imply that  $c$  maps  $\mathbb{R}_+$  one-to-one onto  $\mathbb{R}_+$ .

where  $0 = v_1 < v_n$ . A vector is called a *prize vector* if it satisfies the above condition. Let  $\mathcal{P}^n$  be the set of all  $n$ -dimensional prize vectors. Prize allocation is based on effort ranking: the contestant with the highest effort wins the  $n$ th prize,  $v_n$ , the one with the next-highest effort wins the  $(n - 1)$ th prize,  $v_{n-1}$ , and so on.<sup>9</sup> Ties are broken by fair randomization. A contestant's payoff equals the prize he receives less his effort cost.

### A. Equilibrium

We restrict attention to symmetric equilibria and, thus, suppress the index of a contestant's identity. The rationale for this restriction is discussed at the end of this section. The characterization of equilibria in our model is an immediate corollary of the beautiful result in Barut and Kovenock (1998), which provides a characterization of all equilibria in symmetric complete-information all-pay auctions, where the cost of a bid simply equals the amount bid.

Barut and Kovenock show that these all-pay auctions have a unique symmetric equilibrium (plus possibly a continuum of asymmetric equilibria) and that, in this symmetric equilibrium, all of the bidders continuously randomize their bids over  $[0, v_n - v_1]$  and receive an expected payoff equal to  $v_1$ . The intuition for this symmetric equilibrium characterization is as follows: the symmetric equilibrium must be in mixed strategies without any mass point, because otherwise, there would exist a positive probability that all of the bidders place the same bid, say bid  $s$ , and thus tie at  $s$ . A bidder would then be strictly better off moving mass from  $s$  to a bid that is higher but arbitrarily close to  $s$ . The lower bound of the support of each bidder's bid distribution must be zero, because if it were  $\underline{s} > 0$ , given no mass point, a bidder, by bidding  $\underline{s}$ , could win only the lowest-rank prize,  $v_1$ , a prize that he would win anyway simply by bidding zero. A similar argument implies that each bidder's bid distribution must have an interval support.<sup>10</sup> Because zero is the lower bound of this interval support and because there is no mass point, the expected payoff to each bidder by placing any bid from this interval support must equal  $v_1$ , the payoff from a zero bid. Thus, the upper bound of this interval support must be  $v_n - v_1$ , in which case a bidder, by choosing this upper-bound bid, wins the highest-rank prize,  $v_n$ , for sure while also obtaining a payoff equal to  $v_n - (v_n - v_1) = v_1$ .

<sup>9</sup> We assume that prize allocations are a deterministic function of rank. However, under contestant risk neutrality, this assumption is without loss of generality, since if contestants, instead, received different lotteries over prizes based on their rank, the value of a lottery would depend only on the expectation of the associated random prize.

<sup>10</sup> Otherwise, the support would have a gap, say  $(s_1, s_2)$ ,  $0 \leq s_1 < s_2$ . Then, by a similar argument, bidding  $s_2$  would be strictly dominated by bidding  $s_1$ , which would lead to a contradiction.

To apply the result in Barut and Kovenock (1998) to our model, simply note the one-to-one correspondence between effort levels and effort costs. This one-to-one correspondence implies the equivalence between effort ranking and effort-cost ranking. Thus, we can think of prize allocation as being based on effort-cost ranking. Hence, if we think of contestants as competing by choosing an “effort-cost” bid, such a competition is exactly the symmetric all-pay auction studied by Barut and Kovenock, given the fact that the cost of an effort-cost bid simply equals the amount bid.

Barut and Kovenock’s result implies, given our normalization of  $v_1 = 0$ , that our game, when viewed as an all-pay auction with bids being effort costs, has a unique symmetric equilibrium and that in this equilibrium, every contestant continuously randomizes his effort-cost bid over  $[0, v_n]$ , using the same mixed strategy, and receives zero expected payoff.

Our all-pay contest has a unique symmetric equilibrium, and in this equilibrium, every contestant continuously randomizes effort over  $[0, c^{-1}(v_n)]$ , using the same effort distribution, denoted by  $F_v$ , for random effort, denoted by  $X^v$ , and receives zero expected payoff. This result implies an indifference condition: in the symmetric equilibrium, every contestant is indifferent between all the effort levels over  $[0, c^{-1}(v_n)]$ , and any such effort level gives a contestant zero expected payoff.

We now use this indifference condition to solve for the equilibrium effort distribution,  $F_v : [0, c^{-1}(v_n)] \rightarrow [0, 1]$ . Note that the probability that a contestant obtains the  $i$ th-lowest-rank prize,  $v_i$ , is the probability that the contestant outperforms exactly  $i - 1$  of his  $n - 1$  rivals. Thus, if  $\pi_v(p)$  denotes the expected reward obtained by a contestant who outperforms each rival independently with probability  $p$ ,  $\pi_v$  is given by

$$\pi_v(p) = \sum_{i=1}^n v_i \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i}. \quad (1)$$

Because there is no mass point in the symmetric equilibrium, if a contestant’s effort equals  $x$ , his probability of outperforming any given rival in the symmetric equilibrium equals  $F_v(x)$ , and hence his expected reward equals  $\pi_v \circ F_v(x)$ . The indifference condition implies that

$$\pi_v \circ F_v(x) - c(x) = 0, \quad x \in [0, c^{-1}(v_n)]. \quad (2)$$

Because the function  $\pi_v : [0, 1] \rightarrow [0, v_n]$  is continuous and increasing,  $\pi_v$  has a continuous increasing inverse function,  $\pi_v^{-1} : [0, v_n] \rightarrow [0, 1]$ . The following characterization of equilibrium effort is thus immediate from equation (2).

**THEOREM 1** (Barut and Kovenock). There exists a unique symmetric equilibrium. In this equilibrium, all contestants receive zero expected

payoff, and each contestant's effort distribution  $F_v : [0, c^{-1}(v_n)] \rightarrow [0, 1]$  is given by

$$F_v(x) = (\pi_v^{-1} \circ c)(x),$$

where  $\pi_v^{-1}$  is the inverse function of  $\pi_v$  and  $\pi_v$  is given by equation (1).

An immediate corollary of theorem 1 shows that increasing the rewards being contested induces contestants to increase their effort in the sense of first-order stochastic dominance (FSD).

**COROLLARY 1.** If two prize vectors,  $w, v \in \mathcal{P}^n$ , satisfy that  $w \geq v$ , then  $X^w \succ_{\text{FSD}} X^v$ .<sup>11</sup>

Corollary 1 follows immediately from theorem 1, because, by equation (1),  $\pi_w \geq \pi_v$  if  $w \geq v$ . Given that  $\pi_w$  and  $\pi_v$  are both increasing,  $\pi_w \geq \pi_v$  implies that  $\pi_w^{-1} \leq \pi_v^{-1}$ , which, by theorem 1, further implies that  $F_w \leq F_v$ , that is,  $X^w \succ_{\text{FSD}} X^v$ .

### B. Symmetric Equilibria

The rationale for our restriction to symmetric equilibria is as follows. First, theorem 2 in Barut and Kovenock (1998) implies that asymmetric equilibria exist only when the two lowest-rank prizes are equal, that is, only when  $v_2 = v_1$ . Thus, for prize vectors with  $v_2 \neq v_1$ , the symmetric equilibrium uniquely characterizes contestant behavior. Second, for prize vectors with  $v_2 = v_1$ , a natural requirement for the selection from the equilibrium correspondence is continuity. By the continuity of  $F_v$  in  $v$  implied by theorem 1, picking the symmetric equilibrium is a continuous selection. In contrast, by Barut and Kovenock (1998, their theorem 2), in any asymmetric equilibrium, at least one contestant has a gap in the support of his effort distribution. Such a gap, no matter how large, will vanish if the prize vector is slightly modified to make the two lowest-rank prizes unequal.<sup>12</sup> Thus, none of the asymmetric equilibria can be part of a continuous selection. Third, theorem 2 in Barut and Kovenock (1998) shows that a continuum of asymmetric equilibria arises if  $v_2 = v_1$ . Thus, if we selected asymmetric equilibria to characterize contestant behavior for prize vectors with  $v_2 = v_1$ , it is not obvious which asymmetric equilibria we should select or

<sup>11</sup> We use  $\geq$  to denote the pointwise comparison of vectors.  $X^w \succ_{\text{FSD}} X^v$  if and only if, for any nondecreasing function  $u : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\mathbb{E}[u(X^w)] \geq \mathbb{E}[u(X^v)]$ , provided that the expectations exist.

<sup>12</sup> In particular, if  $v \in \mathcal{P}^n$  has  $v_2 = v_1 = 0$ , a modification that ensures the resulting prizes to be nondecreasing in rank with the two lowest-rank prizes being unequal is as follows: let  $i^* = \min\{i : v_{i+1} - v_i > 0\}$ , and modify  $v$  by adding an arbitrarily small  $\varepsilon > 0$  to each  $v_i$ ,  $i = 2, \dots, i^*$ , and subtracting  $(i^* - 1) \times \varepsilon$  from  $v_{i^*+1}$ . By Barut and Kovenock (1998, their theorem 2), the contest with the resulting prize vector has only the symmetric equilibrium, and, in the symmetric equilibrium, contestants' effort distributions have an interval support (and thus no gap in the support).



how we would define “the” effort distribution for contestants, given that different contestants would have different effort distributions. Finally, section S.3 of the online appendix shows that the symmetric equilibrium dominates all asymmetric equilibria (if any exist) in the sense of maximizing expected total effort.

### III. An Illustration: Motivating the Marketeers

This section illustrates the discouraging effect of competition in a simplified example. Consider the problem of Jill, a retail broker for a small investment bank. Jill’s compensation is based on her clients’ trading commissions. Thus, she would like her clients to trade as often as possible. Jill has three marketing assistants. Her assistants’ job is to make “cold calls” to clients of the bank that pitch trading opportunities. If a cold call results in a “lead,” a client interested in considering the bank’s suggested trade, Jill follows up on the lead with a personal visit or telephone conference.

The assistants must exert costly effort to generate leads. For all three assistants, the cost of effort is determined by the same quadratic function,  $c(x) = 0.001x^2$ , where  $x$  is the number of leads generated (or effort).

Jill has received a \$1,000 incentive fund from the bank earmarked for awards to high-performing assistants. To motivate her assistants, Jill has decided to use this fund to sponsor a rank-order contest. She conjectures that a one-winner contest, which allocates the entire prize to the best-performing assistant, is the optimal way to motivate exceptional effort. In this contest, the prize vector is  $v^1 = (0, 0, 1000)$ .

However, as a reality check, Jill decides to consider an alternative contest structure: a two-winner contest in which the \$1,000 prize is split equally between the two best-performing assistants. In the two-winner contest, the prize vector is  $v^2 = (0, 500, 500)$ . The two-winner contest seems to blunt the incentive for really exceptional performance.

In addition to the reality check provided by the two-winner contest, Jill ponders how she might be able to generate an even more competitive contest. She comes up with an idea: scale up the size of the contest by consolidating her one-winner contest with the identical one-winner contest that her fellow broker, Jack, is planning to run. If Jack’s and Jill’s contests are consolidated, Jack and Jill’s six assistants will compete in a single contest with a prize vector of  $v^{\text{con}} = (0, 0, 0, 0, 1000, 1000)$ , as opposed to competing in two separate one-winner contests.

Although, in both the consolidated and one-winner contests, one-third of the assistants will receive a prize of \$1,000 and two-thirds will receive a prize of 0, the consolidated contest might be more challenging because winning the \$1,000 reward in the consolidated contest requires an assistant to outperform four rivals as opposed to outperforming just two rivals in the one-winner contest.

Jill now turns to the problem of computing the results of the three contests—one-winner, two-winner, and consolidated. By theorem 1, in the one-winner contest, the maximum number of leads produced by an assistant,  $\bar{x}$ , satisfies,  $c(\bar{x}) = v_3 = 1,000$ , implying that  $\bar{x} = 1,000$ . Thus, the equilibrium lead distribution,  $F_v$ , is supported by  $[0, 1,000]$ . By theorem 1, over this support,  $F_v$  satisfies the indifference condition

$$\pi_v \circ F_v(x) = c(x) = 0.001x^2. \tag{3}$$

In equation (3),  $\pi_v(p)$  represents the expected reward to an assistant, given that the assistant outperforms an arbitrary rival with probability  $p$ . Because, in the one-winner contest, a prize can be won only if an assistant outperforms two rivals,  $\pi_v(p) = v_3p^2 = 1,000p^2$ . Thus, for the one-winner contest,  $\pi_v \circ F_v(x) = 1,000F_v(x)^2$ , and hence, equation (3) reduces to  $1,000F_v(x)^2 = 0.001x^2$ . Thus,  $F_v(x) = 0.001x$ ,  $x \in [0, 1,000]$ ; that is, the equilibrium lead distribution for all the assistants is uniform over  $[0, 1,000]$ , implying that the expected number of leads per assistant equals 500.

Performing similar calculations for the two other contests, Jill records in table 1 the results, rounding to the nearest integer.

Jill finds these results rather surprising: the least competitive contest, the two-winner contest, on average produces the most leads per assistant. Jill’s conjecture was that making the contest more competitive would increase the incentive to exert the extraordinary effort required to capture rewards tied to extraordinary performance.

In some sense, Jill’s conjecture is reasonable. Holding the effort distributions of other assistants fixed, it is indeed the case that switching to a more competitive contest increases the incentive to exert high effort. Figure 1 depicts the payoff to an individual assistant in the more competitive contest, given that other assistants play their equilibrium strategies in the less competitive contest. In particular, figure 1A depicts individual assistant payoffs in the one-winner contest if other assistants play their equilibrium two-winner strategies. Figure 1B depicts individual assistant payoffs in the consolidated contest if other assistants play their equilibrium

TABLE 1  
EQUILIBRIUM RESULTS IN THE THREE CONTESTS UNDER CONSIDERATION

CONTEST	PRIZE VECTOR	LEADS PER ASSISTANT			EXPECTED LEADS FOR JILL
		Mean	Minimum	Maximum	
One winner	(0, 0, 1000)	500	0	1,000	1,500
Two winners	(0, 500, 500)	555	0	707	1,665
Consolidated	(0, 0, 0, 0, 1000, 1000)	463	0	1,000	1,389

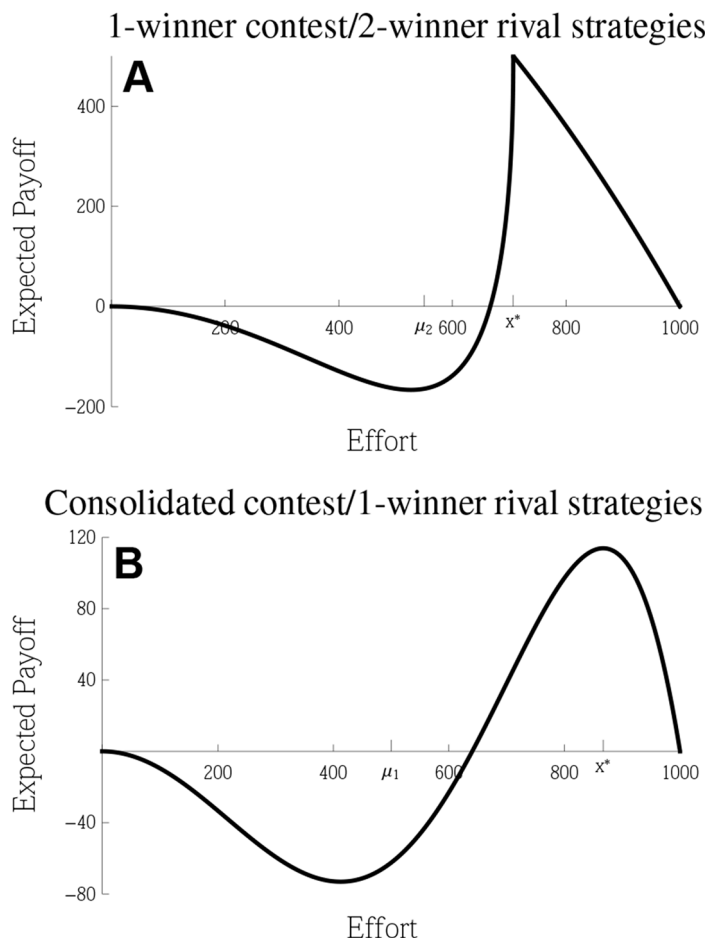


FIG. 1.—Payoff from marketeer effort in the more competitive contest when other assistants play equilibrium strategies in the less competitive contest.  $x^*$  represents the optimal level of effort for an individual assistant.  $\mu_i$  represents mean effort in the  $i$ -winner contest. In A,  $x^* = 707$ , and in B,  $x^* = 867$  (rounded to the nearest integer).

one-winner strategies. In both panels, intermediate levels of effort produce very low assistant payoffs, and each assistant's best reply entails effort far in excess of expected effort in the less competitive contest. Thus, holding the strategies of other assistants fixed, making the contest more competitive does indeed reward extraordinary performance and punish mediocre performance.

However, in both panels, the optimal level of effort is greater than 600, which implies effort costs in excess of 360. Given that only one-third of

the assistants can actually win a prize of 1,000 in the one-winner and consolidated contests, the average reward equals  $1,000/3 < 360$ . For this reason, all the assistants increasing effort to more than 600 cannot be sustained in equilibrium. Thus, to reach an equilibrium, strategies must adjust. The flaw in Jill's conjecture is that it ignores this adjustment.

To reach a new equilibrium associated with the more competitive contests, strategies must adjust so that the assistants do not find it strictly optimal to exert high effort or strictly suboptimal to exert intermediate effort. This requires an increase in the payoff from intermediate effort levels and a decrease in the payoff from high effort levels. The payoff from intermediate effort levels can increase only if the probability that these effort levels win a prize increases. This probability can increase only if the probability of low effort by other assistants increases. Similarly, the payoff from a high effort level can decrease only if other assistants increase their probability of choosing even higher effort levels. This argument leads to the conjecture that equilibrium effort will be more spread out in the more competitive contests.

This conjecture is confirmed in figure 2. Figure 2A compares effort distributions in the one-winner and two-winner contests. Figure 2B compares effort distributions in the consolidated and one-winner contests. In both panels, the equilibrium effort distribution in the more competitive contest crosses the equilibrium effort distribution in the less competitive contest once and from above. This implies that both very low and very high effort levels are more likely in the more competitive contest.

At this point, cost convexity comes into play. Because effort costs are convex, the expected effort is lower for any fixed expected effort cost when effort is more spread out. Because the average reward is the same in all three contests, and the equilibrium condition (theorem 1) requires expected effort cost to equal the average reward, expected assistant effort is lower in the more competitive contest.

Our example implies that the effect of tilting rewards toward better performers and the effect of scaling up the competition are completely analogous: both reduce expected individual effort and increase effort dispersion. The subsequent analysis will extend this insight to more general contests and changes of the prize structure.

#### IV. The Discouraging Effect of Competition

##### A. *Competitiveness Measured by Prize Inequality*

Contests, by assigning prizes based on contestants' performance rankings, create competition between contestants. Such competition is supposed to motivate contestants. One way to increase competitiveness is to tilt rewards toward better performers. However, as the example in section III suggests,

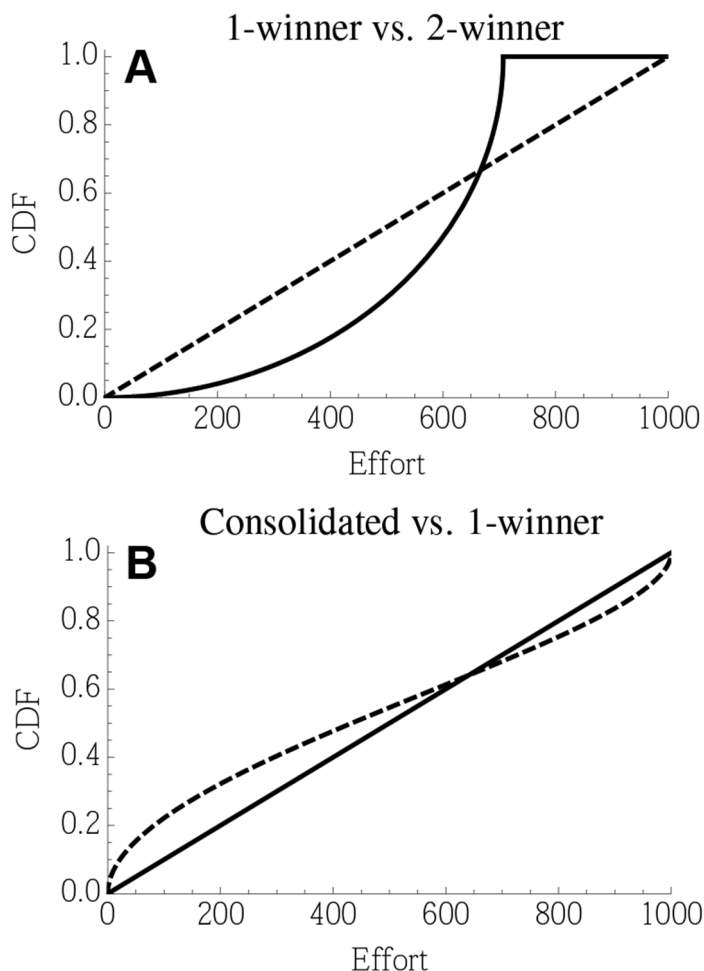


FIG. 2.—Equilibrium effort distributions in marketing contests. The dashed line represents the equilibrium effort distribution in the more competitive contest (one-winner contest in *A* and consolidated contest in *B*). The solid line represents the equilibrium effort distribution in the less competitive contest (two-winner contest in *A* and one-winner contest in *B*). CDF = cumulative distribution function.

increasing competitiveness by offering the entire prize purse to the best performer, as opposed to splitting the total reward equally between the two top performers, induces less expected effort. In fact, this discouragement effect extends to contests with an arbitrary number of homogeneous contestants, an arbitrary convex cost function, and a more general notion of tilting rewards toward better performers, defined under the Lorenz order for prize vectors with the same length and sum. Because the Lorenz

order is a simple, rich order frequently used in the economics literature to measure inequality, it is the natural choice for measuring prize inequality in contests.

**DEFINITION 1 (Prize inequality).** Let  $w \in \mathcal{P}^n$  and  $v \in \mathcal{P}^n$  be two prize vectors with the same length and sum of prizes,  $\sum_{i=1}^n w_i = \sum_{i=1}^n v_i$ . Vector  $w$  is more unequal than  $v$  if  $w$  is more unequal than  $v$  in the Lorenz order, that is,

$$\sum_{i=1}^k w_i \leq \sum_{i=1}^k v_i$$

for all  $k \in \{1, \dots, n\}$ .<sup>13</sup>

### 1. Transferring Value from Higher- to Lower-Rank Prizes

Our definition of prize inequality captures many simple prize-vector transformations. With the lowest-rank prize fixed, one transformation that reduces prize inequality is to transfer value from a higher-rank prize to a lower-rank prize while ensuring that the sequence of prizes is still nondecreasing in rank. For example, changing the prize vector  $(0, 1, 2, 6)$  to  $(0, 1, 3, 5)$  is inequality decreasing, as the highest-rank prize was decreased by 1 while the second-highest-rank prize was increased by 1. Such an inequality-decreasing transfer is called a Pigou-Dalton (P-D) transfer in the economics literature on social inequality. Its opposite operation, the reverse P-D transfer, is clearly inequality increasing.

### 2. Assigning Prizes Randomly and Coarsening Grading Schemes

In the context of contests, Olszewski and Siegel (2018b) develop the idea of “pooling,” that is, pooling an interval of performance ranks and assigning prizes randomly among contestants with the associated ranks. Because pooling makes contestant reward less dependent on contestant effort rank, Olszewski and Siegel interpret pooling as reducing competition. Under our assumption of contestant risk neutrality, pooling is equivalent to an averaging transformation that takes a connected subset of ranks not containing the lowest rank and replaces the prizes associated with these ranks by their average.<sup>14</sup> Like an inequality-decreasing transfer, averaging

<sup>13</sup> To separate the prize-inequality effect, we restrict our definition of prize inequality to prize vectors with the same length and sum, in which case the Lorenz order coincides with majorization developed by Hardy, Littlewood, and Pólya (1952). See sec. IV.B for a discussion of relaxing this restriction.

<sup>14</sup> Formally,  $I$  is a connected set of ranks if  $i, j \in I$ ,  $j > i$ , and  $i < k < j$  implies that  $k \in I$ . Then  $v$  is an averaging transformation of  $w$  if  $I$  is a connected subset,  $1 \notin I$ ,  $v_k = w_k$  for  $k \notin I$ , and  $v_k = 1/|I| \sum_{j \in I} w_j$  for  $k \in I$ . It is well known that such an averaging transformation reduces inequality (see Arnold 1990).

also reduces prize inequality. For example, consider a rank-based grading scheme, often called “grading on a curve” in practice, for four students, with four different grades, A, B, C, and D, whose values are 6, 2, 1, and 0, respectively. The associated prize vector thus equals  $(0, 1, 2, 6)$ . If we average the nonzero prizes by, for example, using a coarser grading scheme with three “Passes” and one “Fail” or assigning grades A, B, and C randomly among the three best-performing students, the new prize vector becomes  $(0, 3, 3, 3)$ , which is more equal than the original prize vector, according to definition 1.<sup>15</sup>

The next theorem presents the discouraging effect of prize inequality.

**THEOREM 2.** Suppose that  $w \in \mathcal{P}^n$  is more unequal than  $v \in \mathcal{P}^n$  and that  $w \neq v$ . For any concave, strictly increasing function,  $u$ , of individual contestant effort,

$$\mathbb{E}[u(X^v)] \geq \mathbb{E}[u(X^w)].$$

Moreover, if  $u$  is strictly concave and/or if the effort-cost function,  $c$ , is strictly convex, the above equation holds with strict inequality.

Theorem 2 shows that any risk-averse contest designer (with concave utility  $u$ ) prefers conducting contests with more equal prizes.<sup>16</sup> A well-known ordering of random variables based on the preferences of all risk-averse expected-utility maximizers is second-order stochastic dominance (SSD).<sup>17</sup> Theorem 2 can be equivalently interpreted as increasing prize inequality reducing each contestant’s effort in the sense of SSD. An immediate result of the theorem is the following.

**COROLLARY 2.** Suppose that  $w \in \mathcal{P}^n$  is more unequal than  $v \in \mathcal{P}^n$  and that  $w \neq v$ . Then the expected individual effort and the expected total effort are both strictly higher under  $v$  than under  $w$  if the effort-cost function,  $c$ , is strictly convex.

**EXAMPLE 1** (Motivating sales revenue through sales contests). Suppose that a firm recruits 15 homogeneous salespeople who are responsible for a selling task, with a quadratic cost function,  $c(x) = x^2/180,000$ , where

<sup>15</sup> We require averaging to be operated on a connected subset, because averaging over a disconnected subset, without reordering the prizes after averaging, may lead to a nonmonotone sequence of prizes. However, if we are allowed to reorder the prizes after averaging, we can always find a nondecreasing sequence of prizes after averaging, in which case averaging over a disconnected subset of ranks not containing the lowest rank does produce a prize vector with lower prize inequality (see Arnold 1990). This can be seen from averaging the values of A and C in our example, which produces a nonmonotone vector,  $(0, 7/2, 2, 7/2)$ . By reordering the prizes after averaging, we obtain a monotone vector,  $(0, 2, 7/2, 7/2)$ , which is more equal than the original prize vector  $(0, 1, 2, 6)$ , according to definition 1.

<sup>16</sup> In the proof of theorem 2, we show that the equilibrium effort cost under  $w$ ,  $c(X^w)$ , is a mean-preserving spread (MPS) of the equilibrium effort cost under  $v$ ,  $c(X^v)$ . Thus,  $X^w$  will be an MPS of  $X^v$  if costs are linear.

<sup>17</sup> SSD is known as dominance in the increasing concave order in the probability-and-statistics literature (see Shaked and Shanthikumar 2007, chap. 4). The term “SSD” and its use in ordering risky choices originated with Hadar and Russell (1969).

$x$  represents sales revenue in dollars. The firm has a budget of \$300 to allocate as a bonus to these salespeople to motivate them. For simplicity, assume that the bonus is of the same size for all salespeople who receive it. The firm then can decide only the number of salespeople,  $k$ , receiving a bonus (of  $300/k$  dollars). Higher  $k$  means that more salespeople receive a bonus but also that the bonus is smaller. As higher  $k$  makes the bonus scheme more equal, corollary 2 implies that expected sales increase in the number of salespeople,  $k$ , receiving a bonus. This effect is illustrated in figure 3A. In fact, expected sales are more than doubled when  $k = 14$  versus when  $k = 1$ . Figure 3B depicts the equilibrium density of individual sales with 15 salespeople and  $k = 1$  (red) and  $k = 5$  (blue). Notably, the equilibrium sales distribution is more spread out, with less weight placed on intermediate levels of sales, with  $k = 1$  bonus than with  $k = 5$  bonuses.

The prize vectors with  $k = 1$  and  $k = 5$  in example 1 coincide, respectively, with those of treatments 1W and GW of field experiment 1, an experimental sales contest conducted by Lim, Ahearne, and Ham (2009). Lim, Ahearne, and Ham consider their subjects to be relatively homogeneous, with similar sales experience. In their field experiment 1, they find that, with total rewards fixed, sales contests lead to higher average sales when  $k = 5$  identical rewards are offered than when a single reward is offered. The prize-inequality effect identified in corollary 2 is consistent with this finding, and, under the specific quadratic cost function we picked for example 1, the numbers of the expected individual sales predicted by our model are also consistent with their reported numbers of average sales (see table 2).<sup>18</sup>

The intuition for the discouraging effect of prize inequality is as follows. Tilting rewards toward top performers increases the benefit of reaching a top rank, as opposed to an intermediate rank, while it decreases the benefit of reaching an intermediate rank, as opposed to a bottom rank. Increased rewards to top performers, combined with decreased “punishment” of underperformance, induce contestants to choose a more

<sup>18</sup> Lim, Ahearne, and Ham (2009) conduct another field experiment, field experiment 2, to test a prediction from Kalra and Shi (2001) that the optimal sales contest should assign each of the nonzero prizes a distinct value. Since the two prize vectors in their field experiment 2 are not ordered by prize inequality according to definition 1, we refrain from reporting the results from their field experiment 2 here. Instead, in sec. S.1 of the online appendix, we show that the qualitative result from their field experiment 2 is consistent with our model prediction for all strictly convex power cost functions,  $c(x) = x^a$ , with power coefficients  $a \in (1, 19]$ . Lim, Ahearne, and Ham (2009) also perform two laboratory experiments with the experimental environment specifically controlled to match the theoretical framework of Kalra and Shi (2001), which adopts a symmetric, Lazear-Rosen tournament model (Lazear and Rosen 1981) with logistically distributed additive noise. As this specific experimental environment is clearly not compatible with our all-pay-contest framework, we do not report the results from their laboratory experiments.



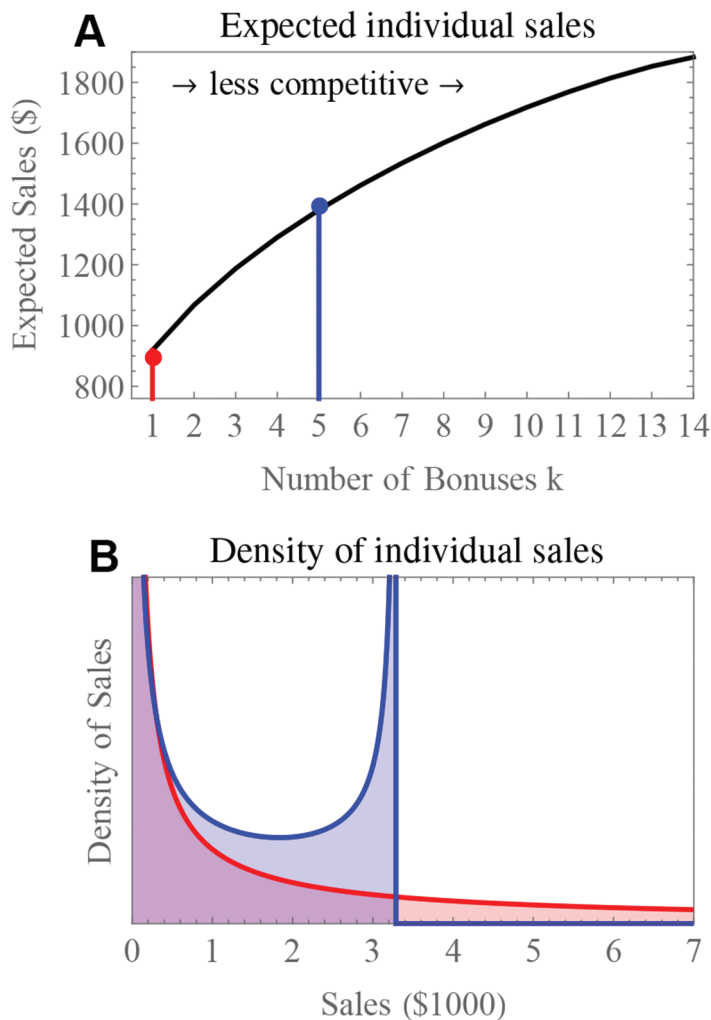


FIG. 3.—Expectation and density of sales in example 1. *A* illustrates expected individual sales from a sales contest with  $n = 15$  salespeople,  $c(x) = x^2/180,000$ , total rewards of \$300 and  $k$  equal bonuses,  $v = (0, \dots, 0, 300/k, \dots, 300/k)$ . Expected sales are more than doubled when  $k = 14$  versus when  $k = 1$  (1,883 vs. 919). The red and blue dots mark the results of field experiment 1 in Lim, Ahearne, and Ham (2009). *B* depicts the equilibrium density of sales with  $k = 1$  bonus (red) and  $k = 5$  bonuses (blue).

spread-out effort distribution. Because competition between homogeneous contestants fully dissipates rents and because total rewards are fixed, the expected cost a contestant incurs is fixed. Thus, given convex effort costs, a more spread-out effort distribution is accompanied by a

TABLE 2  
AVERAGE SALES IN FIELD EXPERIMENT 1 PERFORMED BY LIM, AHEARNE, AND HAM (2009)  
WITH  $n = 15$  SALESPeOPLE IN EACH CONTEST

Treatment	$v \in \mathcal{P}^{15}$	Average Sales (\$)	Predicted Sales (\$)
IW	(0, ..., 0, 300)	897.07	918.56
GW	(0, ..., 0, 60, 60, 60, 60, 60)	1,391.00	1,380.51

NOTE.—The prize vector is made more equal from treatment IW to treatment GW. The difference in average sales is statistically significant. The last column reports the value of expected individual sales predicted by our model for each treatment under quadratic costs  $c(x) = x^2/180,000$ .

decrease in expected effort; that is, contestants’ effort becomes smaller in the sense of SSD.

Corollary 2 implies that, when contestants have identical abilities and convex costs, “punishing” the worst performer by increasing the prize spread between intermediate performers and the worst performer is more effort encouraging than rewarding top performers. In fact, because, with the lowest-rank prize fixed at 0, giving all but the worst performer the same positive prize maximally “punishes” the worst performer (it minimizes prize inequality according to definition 1), when contestants are homogeneous, such a prize scheme maximizes expected total effort under convex costs.

In fact, the implication that “punishing” the worst performer is more effort encouraging than rewarding top performers extends to, and is even strengthened in, the case in which the lowest-rank prize can be reduced (and thus, the worst performer can be directly punished). To see this, suppose that we allow for a reduction of the lowest-rank prize by  $\sigma > 0$ . Because contestant payoff is additively separable in prize and effort costs, this prize reduction is strategically equivalent to adding to each of the prizes other than the lowest-rank one a value of  $\sigma$ . While, by corollary 1, adding these “extra rewards” clearly boosts effort (even when costs are linear), the very fact that, through punishment, these “extra rewards” will be split evenly between nonbottom performers rather than tilted toward top performers implies a much stronger boost in effort when costs are strictly convex than when costs are linear through the discouragement effect. For example, when  $n = 20$ , punishing the worst performer by  $\sigma > 0$  dollars produces a strictly higher expected total effort than rewarding the best performer with  $100 \times \sigma$  dollars under quadratic costs  $c(x) = x^2$ , whereas under linear costs  $c(x) = x$ , the effect of this punishment equals only the effect of rewarding the best performer with  $19 \times \sigma$  dollars.<sup>19</sup>

<sup>19</sup> See sec. S.2 of the online appendix for further illustrations of the relative effectiveness of rewards and punishments on motivating effort.

*B. Competitiveness Measured by Contest Scale*

We see from the example in section III that scaling up a contest by consolidating two identical contests into a larger contest reduces expected individual effort. Notably, scaling up a contest does not change the distribution of rewards. Thus, the effect of contest scale is orthogonal to the effect of prize inequality.<sup>20</sup> Because, after scaling, a contestant has to outperform more rivals to win a particular (nonzero) prize, scaling can arguably be considered as increasing competition.<sup>21</sup> Given the following formal definition of a scaling transformation, we show that the contest-scale effect illustrated in section III extends to contests with an arbitrary number of homogeneous contestants, an arbitrary prize vector, and an arbitrary convex cost function.

**DEFINITION 2 (Scaling).** Let  $s > 1$  be an integer;  $w \in \mathcal{P}^{n \cdot s}$  is a *scaling* of  $v \in \mathcal{P}^n$  if  $w_k = v_{\lfloor k/s \rfloor}$  for all  $k \in \{1, \dots, n \cdot s\}$ .

For example, the twofold ( $s = 2$ ) scaling of the prize vector  $(0, 1, 2)$  is  $(0, 0, 1, 1, 2, 2)$ . A scaling transformation arises when contestants receive grades based on performance percentiles, and thus an increase in the number of contestants does not change the distribution of the grades. Examples include an increase in classroom size when grading on a curve is employed, an increase in the workforce of a firm when the firm adopts a forced-ranking system for promotion and demotion, and an increase in the number of mutual funds under the Morningstar rating system, which assigns each rating to a fixed proportion of funds.<sup>22</sup>

Our next theorem shows that increasing contest scale produces a discouragement effect completely analogous to the discouragement effect produced by prize inequality.

**THEOREM 3.** Suppose that  $w \in \mathcal{P}^{n \cdot s}$ ,  $s > 1$ , is a scaling of  $v \in \mathcal{P}^n$ . For any concave, strictly increasing function,  $u$ , of individual contestant effort,

$$\mathbb{E}[u(X^v)] \geq \mathbb{E}[u(X^w)].$$

Moreover, if  $u$  is strictly concave and/or if the effort-cost function,  $c$ , is strictly convex, the above equation holds with strict inequality.

<sup>20</sup> Scaling does not change the Lorenz order we adopted to measure prize inequality, because scaling does not change the Lorenz curve of rewards, where the Lorenz curve of prize vector  $v \in \mathcal{P}^n$  is the graph of the continuous piecewise linear function,  $[0, 1] \rightarrow [0, 1]$ , connecting the points  $(q_i, L_i)$ ,  $i = 0, 1, \dots, n$ , where  $(q_0, L_0) = (0, 0)$  and for  $i = 1, \dots, n$ ,  $q_i = i/n$  and  $L_i = (\sum_{j=1}^i v_j) / (\sum_{j=1}^n v_j)$ . See Arnold (1990) for a further discussion of the Lorenz curve and the Lorenz order.

<sup>21</sup> For instance, in the example in sec. III, to win a nonzero prize, a contestant has to outperform only two rivals in the one-winner contest, but after a consolidation of the two identical one-winner contests, he has to outperform at least four rivals to win the same nonzero prize.

<sup>22</sup> Morningstar assigns a one-star rating to the 10% worst-performing funds, a two-star rating to the next 22.5% of funds, a three-star rating to the next 35% of funds, a four-star rating to the next 22.5% of funds, and a five-star rating to the 10% best-performing funds within a fund category.

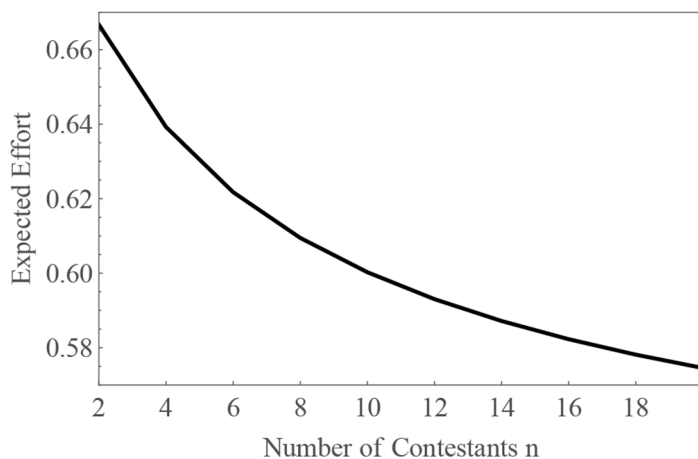


FIG. 4.—Expected individual effort for contests where half the contestants receive a prize of 1 and the rest get zero with quadratic costs  $c(x) = x^2$  and an even number of contestants.

Theorem 3 can be equivalently interpreted as increasing contest scale reducing individual effort in the sense of SSD.<sup>23</sup> By theorem 3, with the distribution of rewards fixed, any risk-averse contest designer (with concave utility  $u$ ) prefers splitting a large contest into smaller identical contests. The next result is immediate from theorem 3.

**COROLLARY 3.** Suppose that  $w \in \mathcal{P}^{n \cdot s}$ ,  $s > 1$ , is a scaling of  $v \in \mathcal{P}^n$  and that the effort-cost function,  $c$ , is strictly convex. Then the expected individual effort is strictly higher under  $v$  than under  $w$ . Moreover, the expected total effort is strictly higher when using  $s$  separate contests, each with prize vector  $v$ , than when consolidating these  $s$  contests into a single contest with prize vector  $w$ .

**EXAMPLE 2** (Effect of the scale of promotion contests). Suppose that an organization uses a promotion contest to motivate employees at the lowest rung of the organization's hierarchy. The best 50% of these employees are promoted to the next rung. The value of promotion to the next rung equals 1, and the value of staying in the bottom rung equals 0; effort costs are given by  $c(x) = x^2$ . Corollary 3 implies that, under strictly convex effort costs, the expected effort of each lowest-rung employee decreases as the scale of the contest increases. Figure 4 documents the predicted negative effect of scale on effort. The example also highlights that the effort reduction associated with increasing contest scale can be significant in magnitude.

<sup>23</sup> In the proof of theorem 3, we show that the equilibrium effort cost under  $w$ ,  $c(X^w)$ , is an MPS of the equilibrium effort cost under  $v$ ,  $c(X^v)$ . Thus,  $X^w$  will be an MPS of  $X^v$  if costs are linear.

The intuition for the discouraging effect of increased contest scale is perhaps easiest to understand by considering the simple contest in example 2, in which half of the contestants win and the other half lose. In this contest, a contestant obtains a winner prize if his effort is higher than the “sample” median effort of his rivals, where the sample is the vector of realized effort levels of all his rivals. With rivals’ mixed strategies fixed, scaling up the contest reduces sample variation. Diminishing sample variation makes it easier for the contestant to “target” the sample median, which is beneficial to the contestant but harmful to his rivals. To reach a new equilibrium, contestants choose more spread-out effort distributions, thereby adding strategic noise to effort choices to retain sample variation. Because scaling does not change the average reward, a more spread-out effort distribution induced by scaling reduces expected effort under convex costs in exactly the same way as when the “spread” is caused by increased prize inequality.

The contest-scale effect is consistent with the results from the experiments conducted by Garcia and Tor (2009), where individuals compete in the speed of solving a quiz (without compromising accuracy). Garcia and Tor find that, with the distribution of rewards fixed, individuals solve the quiz significantly slower in larger contests. Their explanation is based on social comparison. In contrast, we show that, even with purely selfish, rational, expected-utility preferences, it is optimal to exert higher effort in smaller contests, provided that contestants have identical abilities.

Because of the contest-scale effect, the prize-inequality effect we identified in section IV.A cannot be extended to prize vectors with different lengths. To see this, note that, in principle, the Lorenz order can be employed to measure prize inequality for prize vectors with different lengths.<sup>24</sup> For example, consider two prize vectors,  $v = (0, 1)$  and  $w = (0, 0, 0.1, 0.9, 1, 1)$ . Vector  $w$  can be transformed from  $v$  by a threefold scaling, followed by an inequality-decreasing transfer of 0.1 from the third-highest to the fourth-highest-rank prize. Because scaling does not change inequality measured by the Lorenz order,  $w$  is more equal than  $v$  under the Lorenz order. However, under the assumption that  $c(x) = x^3$ , it turns out that  $\mathbb{E}(X^w) \approx 0.70 < 0.75 = \mathbb{E}(X^v)$ , and thus expected effort is lower in the contest with more-equal prizes. Inspection of this example shows the driving force:  $w$ , while being more equal, has a larger scale. When the contest-scale effect dominates, increasing contest scale discourages effort more than reduced prize inequality encourages effort.

<sup>24</sup> In particular, one can compute the Lorenz curve associated with each prize vector. If the Lorenz curve of one prize vector never lies above that of the other and if the two Lorenz curves are not identical, then the former prize vector is more unequal than the latter, according to the Lorenz order (see Arnold 1990). See n. 20 for the definition of the Lorenz curve.

### C. Competitiveness Measured by Entry

Another simple way to increase competition is to add entrants to the competition without adding any new positive prize. How does contestant entry affect effort? To answer this question, we introduce the entry transformation—adding an entrant and, at the same time, a zero prize to the contest.

DEFINITION 3 (Entry). Prize vector  $w \in \mathcal{P}^{n+1}$  is an *entry transformation* of  $v \in \mathcal{P}^n$  if  $w = (0, v_1, v_2, \dots, v_n)$ .

The next result shows that entry reduces each contestant's effort in the sense of FSD while it increases expected total effort.

PROPOSITION 1. If  $w \in \mathcal{P}^{n+1}$  is an entry transformation of  $v \in \mathcal{P}^n$ , then the expected individual effort is strictly lower under  $w$  than under  $v$  and  $X^w \preceq_{\text{FSD}} X^v$ . If the effort-cost function,  $c$ , is (strictly) convex, then the expected total effort is (strictly) higher under  $w$  than under  $v$ .

Entry, by reducing the average reward, reduces the cost of effort a contestant is willing to incur in expectation, which discourages individual effort. Unlike when competition is increased by increasing prize inequality or consolidating identical contests, when competition is increased by adding an entrant, discouraged individual effort does not translate into reduced expected total effort. This is because the number of contestants becomes larger under contestant entry. Given that expected total costs are fixed by total rewards, which are unaffected by entry, one can think of adding an entrant as having a new agent “share” a fixed total workload (in expected-cost terms). Workload sharing increases total productivity under convex effort costs.

### D. Implications

Our analysis shows that, when effort costs are convex, increasing competition discourages individual effort. The key restriction imposed by our setting is contestant homogeneity. Thus, the implications of our analysis are most relevant for contests with preselected contestants, for example, employees who all passed employment screening tests, students in ability-track classes, and online gamers who have been assigned competitors on the basis of similar past performance. In research settings where preselection appears to substantially reduce contestant heterogeneity, such as marketing and personnel economics, homogeneous-contestant models are frequently developed and tested in laboratory and field experiments.<sup>25</sup>

Our analysis of competitiveness measured by prize inequality (theorem 2) predicts that reward systems are most effective in eliciting effort when they

<sup>25</sup> For examples in marketing research on sales contests, see Kalra and Shi (2001), Lim, Ahearne, and Ham (2009), and Lim (2010). For examples in personnel economics, see Krishna and Morgan (1998), Hvide (2002), Fu and Lu (2009), Chowdhury and Kim (2017), and Drugov and Ryvkin (2019).

impose severe consequences on the lowest performance levels and are insensitive to exceptionally high performance. Consistent with this prediction, Medoff and Abraham (1980) found that the promotion probability for employees with low ratings was significantly less than the probability for employees with middling ratings but that there was no significant difference between the promotion probabilities for middling and highly rated employees.<sup>26</sup> Bishop (1990) found that the effect of relative performance on the probability of dismissal is much larger than its effect on compensation.

In China, despite their illegality, “last-place elimination systems”—which dismiss any employee with last-place performance evaluations, even if the employee’s performance satisfies minimum job requirements—are pervasive (Xu 2016). In a survey conducted by *China Youth Daily* in 2016, 56.1% of respondents reported that their organizations use last-place elimination systems (*China Youth Daily* 2016).<sup>27</sup>

In educational contexts where students compete on a grading curve, coarsening a grading scheme (with no coarsening of the lowest grade) reduces prize inequality, and thus our analysis predicts that grade coarsening will increase the motivation of students with (nearly) identical abilities.<sup>28</sup>

With respect to competitiveness measured by contest scale, forced-ranking systems provide a potential test bed. Under a forced-ranking system, the distribution of rewards is fixed, while the scale of the contest is determined by the number of contestants in the comparison group.<sup>29</sup> Theorem 3 predicts that the motivational effect of forced-ranking systems will be decreasing in the size of the comparison group.

Empirical research has identified some comparison groups: for example, multidivision firms compare division managers (Cichello et al. 2009), and multibank holding companies compare branch-bank managers (Blackwell, Brickley, and Weisback 1994). In the setting of the Morningstar ratings of mutual fund performance, the forced-ranking system is public

<sup>26</sup> See table IV in Medoff and Abraham (1980).

<sup>27</sup> *China Youth Daily* is the official organ of the Communist Youth League of China.

<sup>28</sup> Olszewski and Siegel (2018b) show, in a large-contest framework with heterogeneous contestants (students), that coarsening grading schemes can improve students’ welfare in a Pareto sense by reducing students’ costly effort exerted to compete for better grades. In contrast, in our homogeneous-contestant setting, coarsening a grading scheme (with no coarsening of the lowest grade) increases students’ expected effort without affecting student welfare. See Dubey and Geanakoplos (2010) and Bodoh-Creed and Hickman (2018) for further discussions of how changes of grade structures affect students’ effort.

<sup>29</sup> For example, under Jack Welch’s famous “20-70-10” employee ranking system, employees are ranked on the basis of performance evaluations. The top 20%, the “A players,” are promoted, while the bottom 10%, the “C players,” are punished. Jennifer Alsever (CBS News) described the severe consequences of being labeled a “C player”: “They’re given no raises or bonuses and are either offered training, asked if they’d be happier elsewhere, or fired.” See Alsever (2008).

information, the competitors are mutual fund managers, and the comparison groups are fund categories.<sup>30</sup>

## V. Robustness and Extensions

### A. Contestant Heterogeneity

Abstracting from contestant heterogeneity has permitted us to sharply characterize the discouraging effects of increased prize inequality and contest scale as a consequence of effort-cost convexity. A natural extension of our analysis is to consider contests where contestants are asymmetric and information is complete. Unfortunately, this setup is known to be intractable, which makes a general characterization of the effect of cost convexity infeasible for contests with an arbitrary finite number of heterogeneous contestants.

However, the existing literature has provided results that suggest that introducing contestant heterogeneity does not always favor competitive contest designs.<sup>31</sup> Olszewski and Siegel (2018c) characterize optimal (effort-maximizing) contest designs by considering the limit of contestant behavior as the number of (ex post heterogeneous) contestants goes to infinity.<sup>32</sup> They show that, when effort costs are strictly convex, optimal contests feature many distinct prizes that only gradually vary with relative performance and that, if the marginal cost at zero effort is zero, it is optimal to give almost every contestant a positive prize. Their result suggests that in large contests, while contestant heterogeneity does push optimal designs away from extremely low prize inequality (the optimal design in our symmetric setting), the magnitude of such a “push” is limited in that optimal designs still feature many positive prizes or even a positive prize to almost every contestant.

Xiao (2018) characterizes optimal contest designs with three asymmetric contestants, complete information, and linear costs. He finds that the

<sup>30</sup> Morningstar assigns one of five different ratings to each fund within a fund category based on the percentile of the fund’s performance in that category. Rankings translate into rewards for managers: a high Morningstar rating has a positive effect on a management company’s assets under management (AUM; Del Guercio and Tkac 2008) and thus on its asset management fees. Fund manager compensation is frequently linked directly to the fund manager’s AUM. Moreover, fund managers, through profit sharing, also capture some of the increase in management fees received by the company (Ma, Tang, and Gómez 2019).

<sup>31</sup> Contestant heterogeneity is commonly introduced by assuming that contestant  $i$ , by exerting effort  $x_i$ , incurs effort cost  $c(x_i)/a_i$ , where  $a_i > 0$  represents  $i$ ’s ability.

<sup>32</sup> Olszewski and Siegel (2018a) show that all equilibria can be approximated by the behavior in a single-agent mechanism when the number of contestants goes to infinity, for very general contests that include as a special case the environment with symmetric contestants and complete information. Their analysis can be used to derive a result similar to theorem 2 for large contests.



most equal prize structure (i.e., offering two equal prizes in the three-contestant case) is optimal if the contestant with the greatest ability is sufficiently stronger than the other two contestants. Xiao's result suggests that the contestant-heterogeneity effect can sometimes be aligned with the cost-convexity effect when there is a single contestant with the greatest ability.

Moreover, as we show in section S.4 of the online appendix, while the directional alignment between the contestant-heterogeneity effect and the cost-convexity effect breaks when there are at least two contestants with the greatest ability, the cost-convexity effect can overwhelm the contestant-heterogeneity effect. In particular, we show that, if there are at least two contestants with the greatest ability, winner-take-all contests are always optimal when costs are linear, but when costs are strictly convex, winner-take-all contests can be suboptimal and will always be suboptimal if there are at least three contestants with the greatest ability. Thus, both the existing results and our analysis suggest that the discouragement effect we identified persists, at least to some extent, in contests with heterogeneous contestants.

### *B. Concave Effort Costs*

While concave effort costs are less relevant for economic applications, their effect is straightforward from our earlier analysis. Increasing prize inequality or contest scale induces each contestant to choose a more spread-out effort distribution without changing the expected effort cost. Thus, if costs are concave, increasing prize inequality or contest scale will induce a higher expected individual effort. Consequently, under concave costs, the optimal (effort maximizing) prize structure in our model is the winner-take-all structure. This result is consistent with the optimality of the winner-take-all structure under concave costs in symmetric incomplete-information all-pay contests (Moldovanu and Sela 2001) and in large contests (Olszewski and Siegel 2018c).

Now consider contestant entry. Entry, by reducing the average reward, reduces expected individual effort cost. In fact, as we show in the proof of proposition 1, entry reduces individual effort cost in the sense of FSD regardless of the effort-cost curvature. Thus, given the positive relationship between effort and effort cost, entry reduces individual effort in the sense of FSD even when costs are concave. However, the effect of entry on total effort is completely different when costs are concave rather than convex. Recall that, in section IV.C, we showed that entry, while discouraging individual effort, increases expected total effort under convex costs via "workload sharing." In contrast, workload sharing reduces total productivity

when costs are concave. Thus, entry reduces expected total effort under concave costs.<sup>33</sup>

### C. Risk-Averse Contestants

The discouragement effect identified in section IV was based on a setting with risk-neutral contestants. Introducing contestant risk aversion by assuming that contestant welfare is additively separable in the utility of prize,  $u^A(\cdot)$ , and the disutility of effort costs,  $c(\cdot)$ , where  $u^A: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a concave, strictly increasing function, does not affect any conclusion in section IV. To see this, note that, under risk aversion, we should distinguish between prize vector  $v \in \mathcal{P}^n$  and the associated “utility-prize” vector  $(u^A(v_1), u^A(v_2), \dots, u^A(v_n))$ . It is the utility-prize vector that directly governs contestant behavior. Because scaling up a contest or adding entrants to a contest changes the utility-prize vector in the same way as either changes the prize vector, it is clear that risk aversion does not change any result for the effects of contest scale or entry.

Now consider how contestant risk aversion might affect the analysis of the prize-inequality effect by considering an inequality-increasing transfer of value from the  $i$ th-lowest-rank prize to the  $j$ th-lowest-rank prize,  $1 < i < j$ . Under risk aversion, such a transfer reduces contestants’ valuation of the  $i$ th-lowest-rank prize by more than it increases their valuation of the  $j$ th-lowest-rank prize. Thus, reflected in the utility-prize vector, such a transfer can be considered as an inequality-increasing transfer operated on the utility-prize vector, followed by an absolute decrease in the utility value of the  $j$ th-lowest-rank prize. The former, by theorem 2, decreases effort in the second order and the latter, by corollary 1, decreases effort in the first order. This last effect is absent under contestant risk neutrality. This new effect makes increased prize inequality effort discouraging even under linear costs. Indeed, as shown by Glazer and Hassin (1988, their proposition 1), the optimal (effort-maximizing) prize structure for symmetric complete-information all-pay contests with risk-averse contestants and linear costs is the one that gives identical prizes to all but one contestant.<sup>34</sup> This prize structure coincides with the optimal prize structure in our model with risk-neutral contestants but convex costs.<sup>35</sup>

<sup>33</sup> This result is immediate from flipping the sign of the last inequality, labeled inequality (2), in the proof of proposition 1.

<sup>34</sup> While Glazer and Hassin (1988) allow for contest participation costs, their proposition 1 is derived under the assumption that there is no participation cost (i.e., an agent’s payoff from entering the contest and exerting zero effort is always no less than his payoff from not participating in the contest).

<sup>35</sup> See Olszewski and Siegel (2018c) for a discussion of the similarities and differences between the effects on the optimal prize structure of convex costs and concave prize valuations in large contests.

#### *D. Noisy Outcomes*

It is worth noting that the discouragement effect identified in this paper operates through the increased effort dispersion produced by increased competitiveness. Effort dispersion is a feature of equilibria in the noiseless-contest/homogeneous-contestant setting studied in this paper. However, in noisy contests (e.g., Tullock contests), where contestants' effort determines their performance ranking in a probabilistic way, equilibria that feature no effort dispersion exist if contests are sufficiently noisy (see Clark and Riis 1998; Schweinzer and Segev 2012). Absent effort dispersion, the discouragement effect we identified clearly does not operate.

However, the noisy-contest literature also provides results that suggest the presence of a discouragement effect when contests are not too noisy. For example, propositions 2 and 4 in Schweinzer and Segev (2012) show that, in Tullock contests with homogeneous contestants, when contests are fairly noisy but not too noisy, the winner-take-all structure induces positive effort dispersion, but multiprize structures exist that induce no effort dispersion and fully dissipate contestant rents. In such contests, when effort costs are strictly convex, for essentially the same reasons that the discouragement effect holds in our noiseless-contest model, these multiprize structures will yield greater contestant effort than the winner-take-all structure.<sup>36</sup>

## **VI. Conclusion**

We have shown that, in symmetric complete-information all-pay contests, increasing contest competitiveness, under three disparate notions of competitiveness (i.e., prize inequality, contest scale, and entry of new contestants), always discourages effort under convex costs. The discouragement effect produced by increased prize inequality or contest scale is an equilibrium phenomenon. Expecting other contestants' increased likelihood of exerting high effort in more competitive contests increases a contestant's incentive to exert low effort, leading to a lower expected per capita contestant equilibrium effort under convex costs. While other

<sup>36</sup> As shown by Fu and Lu (2012), Tullock contests are strategically equivalent to a Lazear-Rosen tournament (Lazear and Rosen 1981) with additive noise following a type I extreme-value (maximum) distribution. In a recent working paper, Drugov and Ryvkin (2019) study Lazear-Rosen tournaments with homogeneous contestants and convex costs under a more general noise distribution. Restricting to situations in which all prize structures support a symmetric equilibrium that features no effort dispersion (situations that require the noise distribution to be sufficiently dispersed), Drugov and Ryvkin identify "heavy tails" of noise as another factor that can render multiprize structures optimal (effort maximizing). Because the heavy-tails effect is identified in a setting absent equilibrium effort dispersion, while the discouragement effect we showed in this paper is identified in a setting absent exogenous noise (and, a fortiori, absent heavy-tailed noise), the results complement each other.

real-world factors, such as contestant heterogeneity, which we have abstracted from in our analysis, can sometimes counter the benefit of reduced competition resulting from the discouragement effect, given the discussion in section V.A and the significance of the effect as illustrated by examples 1 and 2, we believe that the discouragement effect will influence contestant behavior in real-world contests.

The analysis of this paper by no means exhausts the potentially useful and applicable results that our framework can deliver. For example, while we have largely focused on the effect of competition on expected effort, our analysis can also deliver useful implications for experimental and empirical research on how competition can affect the dispersion of effort. Because the discouragement effect scales down mean effort, measuring effort dispersion using variance or standard deviation of effort might not capture the dispersion effect we identified in theorems 2 and 3. In section S.5 of the online appendix, we provide sufficient conditions under which the dispersion effect can be identified and tested using standard statistical measures (e.g., variance, standard deviation, and coefficient of variation).<sup>37</sup>

## Appendix

### Proofs of Results

#### A1. Preliminaries

##### A1.1. Two Useful Stochastic Orders

In the proofs, we employ some well-known stochastic order results established in the probability-and-statistics literature. In this literature, an MPS is called *dominance in the convex order*, and SSD is called *dominance in the increasing concave order*. Their definitions can be found in Shaked and Shanthikumar (2007) and are given as follows.

DEFINITION A1. Let  $X$  and  $Y$  be two random variables.

- i. If  $\mathbb{E}[h(X)] \geq \mathbb{E}[h(Y)]$  for all convex functions  $h: \mathbb{R} \rightarrow \mathbb{R}$ , provided that the expectations exist, then  $X$  is said to be *larger than  $Y$  in the convex order*, denoted by  $X \succ_{\text{cx}} Y$ .  $X \succ_{\text{cx}} Y$  is equivalent to  $X$  being an MPS of  $Y$ .
- ii. If  $\mathbb{E}[h(X)] \geq \mathbb{E}[h(Y)]$  for all nondecreasing concave functions  $h: \mathbb{R} \rightarrow \mathbb{R}$ , provided that the expectations exist, then  $X$  is said to be *larger than  $Y$  in*

<sup>37</sup> For example, we show in proposition S-4, in sec. S.5 of the online appendix, that, if  $w \in \mathcal{P}^n$  is a value-preserving convexifying transformation of  $v \in \mathcal{P}^n$ , an inequality-increasing transformation that satisfies  $\sum_{i=1}^n w_i = \sum_{i=1}^n v_i$  and  $(w_{i+1} - w_i)(v_i - v_{i-1}) \geq (w_i - w_{i-1})(v_{i+1} - v_i)$  for all  $2 \leq i \leq n-1$  (which implies the satisfaction of the conditions in definition 1), then mean-scaled effort is more dispersed under  $w$  than under  $v$ , with  $X^w/\mathbb{E}(X^w)$  being an MPS of  $X^v/\mathbb{E}(X^v)$ , when the effort-cost function is geometrically convex (including all “standard,” convex effort-cost functions, e.g., power, exponential, quadratic). We then show in table S-3 in the online appendix that this prediction is consistent with the results of field experiment 1 performed by Lim, Ahearne, and Ham (2009).

the increasing concave order, denoted by  $X \succ_{\text{icv}} Y$ .  $X \succ_{\text{icv}} Y$  is equivalent to  $X \succ_{\text{SSD}} Y$ .

FACT A1. If  $X \succ_{\text{cx}} Y$ , then (i)  $\mu(X)$ , the expectation of  $X$ , equals  $\mu(Y)$ , and (ii)  $X \preccurlyeq_{\text{icv}} Y$ .

Such a relation is well known in the economics literature, as that  $X$  being an MPS of  $Y$  implies that  $\mu(X) = \mu(Y)$  and  $X \preccurlyeq_{\text{SSD}} Y$ .

## A1.2. Notation

In our proofs,  $F_v$  and  $Q_v$  denote, respectively, the cumulative distribution function (CDF) and the quantile function for  $X^v$ , the equilibrium random effort under prize vector  $v$ . The terms  $F_{v,c}$  and  $Q_{v,c}$  denote, respectively, the CDF and the quantile function for  $C^v$ , the equilibrium random effort cost under prize vector  $v$ . The terms  $F_w$ ,  $Q_w$ ,  $F_{w,c}$ , and  $Q_{w,c}$  are defined analogously for prize vector  $w$ . The term  $\Delta v_i = v_{i+1} - v_i$  is defined as the  $i$ th prize difference for prize vector  $v$ . The term  $\Delta w_i = w_{i+1} - w_i$  is defined analogously. The term  $\tilde{B}_{(m,p)}$  represents a binomial  $(m, p)$  random variable. Define  $b(k; m, p)$  and  $\hat{B}(k; m, p)$  as

$$b(k; m, p) := \mathbb{P}[\tilde{B}_{(m,p)} = k] = \binom{m}{k} p^k (1-p)^{m-k}, \quad (\text{A1})$$

$$\hat{B}(k; m, p) := \mathbb{P}[\tilde{B}_{(m,p)} \geq k] = \sum_{i=k}^m \binom{m}{i} p^i (1-p)^{m-i}. \quad (\text{A2})$$

## A1.3. Equilibrium Characterization (Theorem 1) and the Quantile Representations

Theorem 1 characterizes the unique symmetric equilibrium by presenting the CDF for the equilibrium effort. The proof of theorem 1 is given by the discussion in section II. In fact, for comparative statics, it is more convenient to employ the quantile functions for the equilibrium effort and effort-cost distributions. The following lemma gives such quantile representations.

LEMMA A1. Define  $b(k; m, p)$  and  $\hat{B}(k; m, p)$  by equations (A1) and (A2), respectively. The quantile function for the equilibrium effort cost distribution under  $v \in \mathcal{P}^n$ ,  $Q_{v,c} : [0, 1] \rightarrow [0, v_n]$ , has two equivalent representations:

$$Q_{v,c}(p) = \sum_{i=1}^n v_i b(i-1; n-1, p) = \sum_{i=1}^{n-1} \Delta v_i \hat{B}(i; n-1, p), \quad p \in [0, 1]. \quad (\text{A3})$$

The quantile density function for the equilibrium effort cost distribution is given by

$$Q'_{v,c}(p) = (n-1) \sum_{i=1}^{n-1} \Delta v_i b(i-1; n-2, p). \quad (\text{A4})$$

The associated effort distribution has its quantile function,  $Q_v : [0, 1] \rightarrow [0, c^{-1}(v_n)]$ , given by

$$Q_v(p) = c^{-1} \circ Q_{v,c}(p), \quad p \in [0, 1]. \quad (\text{A5})$$

*Proof.* As discussed in section II, the result in Barut and Kovenock (1998) implies that our game, when viewed as an all-pay auction with bids being effort costs, has a unique symmetric equilibrium and that in this equilibrium, every contestant continuously randomizes his effort-cost bid over  $[0, v_n]$  using the same cost distribution,  $F_{v,c}$ , and receives zero expected payoff. This implies the following indifference condition:

$$\pi_v \circ F_{v,c}(c) = c, \quad c \in [0, v_n], \quad (\text{A6})$$

where  $\pi_v$  is given by equation (1). Thus, over the support of  $F_{v,c}$ ,  $\pi_v$  composed with  $F_{v,c}$  leads to an identity function, implying  $Q_{v,c} = F_{v,c}^{-1} = \pi_v$ . The first equality in equation (A3) is then immediate from equations (1) and (A1).

Applying Abel summation to the expression for  $Q_{v,c}$  given by the first equality in equation (A3) gives the second equality in equation (A3). Note that binomial and beta distributions satisfy

$$\hat{B}(k; m, p) = \text{Beta}[k, m - k + 1](p), \quad (\text{A7})$$

where  $\text{Beta}[k, m - k + 1]$  represents the CDF of a Beta distribution with shape parameters  $k$  and  $m - k + 1$ . The probability density function of this Beta random variable satisfies

$$\partial_p \text{Beta}[k, m - k + 1](p) = m \binom{m-1}{k-1} p^{k-1} (1-p)^{m-k} = mb(k-1; m-1, p).$$

Thus, by equation (A7), we have

$$\partial_p \hat{B}(k; m, p) = mb(k-1; m-1, p). \quad (\text{A8})$$

Equation (A8) and the second equality in equation (A3) imply that

$$Q'_{v,c}(p) = \sum_{i=1}^{n-1} \Delta v_i \partial_p \hat{B}(i; n-1, p) = (n-1) \sum_{i=1}^{n-1} \Delta v_i b(i-1; n-2, p).$$

This establishes equation (A4). By theorem 1 and the fact that  $\pi_v = Q_{v,c}$ , we have  $Q_v = F_v^{-1} = c^{-1} \circ \pi_v = c^{-1} \circ Q_{v,c}$ . Equation (A5) thus follows. QED

## A2. Proofs

For proofs of theorem 1 and corollary 1, see the discussion in section II.

### A2.1. Proof of Theorem 2

We first consider a special form of a decrease in prize inequality, a P-D transfer: if two prize vectors,  $v \in \mathcal{P}^n$  and  $w \in \mathcal{P}^n$ , differ only in the prizes associated with two ranks, say the  $j$ th and the  $k$ th,  $1 < j < k \leq n$ , with  $w_k - v_k = -(w_j - v_j) > 0$ , we say that  $v$  differs from  $w$  by a P-D transfer. The next lemma is the key to the proof.

LEMMA A2. If  $v \in \mathcal{P}^n$  differs from  $w \in \mathcal{P}^n$  by a P-D transfer, then  $C^w$  is an MPS of  $C^v$ .

*Proof.* Given that  $v \in \mathcal{P}^n$  differs from  $w \in \mathcal{P}^n$  by a P-D transfer, suppose, without loss of generality, that  $w_k - v_k = -(w_j - v_j) = \delta > 0$ ,  $1 < j < k \leq n$ , and  $w_i = v_i$  for all  $i \neq j, k$ . Thus, by the first equality in equation (A3), we have

$$Q_{w,c}(p) - Q_{v,c}(p) = \delta(b(k-1; n-1, p) - b(j-1; n-1, p)), \quad p \in [0, 1]. \quad (\text{A9})$$

Note that  $b(k-1; n-1, p)/b(j-1; n-1, p) = t(p/(1-p))^{k-j}$ , where  $t = \binom{n-1}{k-1} / \binom{n-1}{j-1} > 0$  is a constant. Thus, given that, for  $1 < j < k \leq n$ ,  $p \mapsto (p/(1-p))^{k-j}$ ,  $p \in [0, 1]$ , is strictly increasing, tends to 0 as  $p \rightarrow 0$ , and tends to  $+\infty$  as  $p \rightarrow 1$ , there must exist  $p^* \in (0, 1)$  such that  $b(k-1; n-1, p^*)/b(j-1; n-1, p^*) = 1$  and  $b(k-1; n-1, p)/b(j-1; n-1, p) < (>)1$  if  $p < (>)p^*$ . Thus,  $b(k-1; n-1, p) - b(j-1; n-1, p)$  changes sign exactly once as  $p$  increases from 0 to 1, and the sign pattern is  $- +$ . Thus, by equation (A9) and the hypothesis that  $\delta > 0$ , the sign pattern of  $Q_{w,c} - Q_{v,c}$  is  $- +$ , implying that the sign pattern of  $F_{w,c} - F_{v,c}$  is  $+ -$ . Since the expected effort cost equals the average reward and the P-D transfer does not change the average reward, the expected cost is fixed. The lemma then follows immediately from the sign pattern,  $+ -$ , of  $F_{w,c} - F_{v,c}$  and theorem 3.A.44 in Shaked and Shanthikumar (2007). QED

Under our definition of prize inequality,  $w$  is more unequal than  $v$  if and only if  $w$  majorizes  $v$  with  $w_1 = v_1 = 0$ . Lemma 2 in Thon and Wallace (2004) shows that  $w$  majorizes  $v$  if and only if  $v$  can be obtained from  $w$  by a sequence of H-transfers, where an H-transfer is defined as a P-D transfer that preserves the original ranking of prizes. Thus, by lemma A2,  $C^w$  must differ from  $C^v$  by a sequence of MPSs and is hence an MPS of  $C^v$ .

By fact A1,  $C^w$  being an MPS of  $C^v$  implies that  $C^w \preccurlyeq_{\text{icv}} C^v$  (i.e.,  $C^w \preccurlyeq_{\text{SSD}} C^v$ ). Theorem 4.A.8 in Shaked and Shanthikumar (2007) shows that the increasing concave order is closed under increasing concave transformations. Thus, given that  $X^v = c^{-1}(C^v)$  and  $X^w = c^{-1}(C^w)$ , where  $c^{-1}$ , being the inverse function of  $c$ , is increasing and concave,  $C^w \preccurlyeq_{\text{icv}} C^v$  implies that  $X^w \preccurlyeq_{\text{icv}} X^v$  (i.e.,  $X^w \preccurlyeq_{\text{SSD}} X^v$ ). The definition of the increasing concave order then implies that  $\mathbb{E}[u(X^v)] \geq \mathbb{E}[u(X^w)]$  for any concave, strictly increasing function  $u$ .

To rule out the possibility of  $\mathbb{E}[u(X^v)] = \mathbb{E}[u(X^w)]$  when  $u$  is strictly concave and/or  $c$  is strictly convex, we can use a contradiction argument. Suppose, to the contrary, that  $\mathbb{E}[u(X^v)] = \mathbb{E}[u(X^w)]$  if  $u$  is strictly concave and/or  $c$  is strictly convex. Note that  $\mathbb{E}[u(X^v)] = \mathbb{E}[u(X^w)]$  implies that  $\mathbb{E}[u \circ c^{-1}(C^w)] = \mathbb{E}[u \circ c^{-1}(C^v)]$ . The hypothesis that  $u$  is strictly concave and/or  $c$  is strictly convex implies strict concavity of  $u \circ c^{-1}$ . Then, by theorem 4.A.49 in Shaked and Shanthikumar (2007),  $C^w \preccurlyeq_{\text{icv}} C^v$  and  $\mathbb{E}[u \circ c^{-1}(C^w)] = \mathbb{E}[u \circ c^{-1}(C^v)]$  imply that  $C^w \stackrel{\text{dist.}}{=} C^v$ . However, given that  $w \neq v$ , it is clear that  $C^w$  and  $C^v$  cannot be identically distributed, a contradiction. QED

The proof of corollary 2 follows immediately from that of theorem 2. QED

## A2.2. Proof of Theorem 3

We first show that  $C^w$  is an MPS of  $C^v$  (equivalently,  $C^w \succcurlyeq_{\text{cx}} C^v$ ). Note that scaling does not change the average reward and that the expected effort cost equals the average reward. Thus,  $w$  being a scaling of  $v$  implies that  $C^w$  and  $C^v$  have the

same mean. In this case, by theorem 3.A.5 in Shaked and Shanthikumar (2007), to show that  $C^w \succ_{\text{cx}} C^v$ , it suffices to show that

$$\int_0^p (Q_{w,c}(u) - Q_{v,c}(u)) du \leq 0, \quad p \in [0, 1]. \quad (\text{A10})$$

To show that condition (A10) holds under scaling, first note that  $v \in \mathcal{P}^n$  and its scaling  $w \in \mathcal{P}^{s \cdot n}$  are related as follows: for all  $i \in \{1, \dots, n-1\}$ , a prize difference that is at the  $i$ th rank of  $v$ ,  $\Delta v_i = v_{i+1} - v_i$ , is realized at the  $(si)$ th rank of  $w$ , that is,  $\Delta v_i = \Delta w_{si}$ ; and for any  $j$  that is not a multiple of  $s$ , that is, for any  $j \notin \{s, 2s, \dots, (n-1)s\}$ ,  $\Delta w_j = 0$ . This observation, combined with the second equality in equation (A3), implies that

$$\int_0^p (Q_{w,c}(u) - Q_{v,c}(u)) du = \sum_{i=1}^{n-1} \Delta v_i \mathbf{I}[\mathcal{D}_i](p), \quad \text{where } \mathbf{I}[\mathcal{D}_i](p) = \int_0^p \mathcal{D}_i(u) du, \quad (\text{A11})$$

$$\text{and } \mathcal{D}_i(u) = \hat{B}(si; sn-1, u) - \hat{B}(i; n-1, u), \quad (\text{A12})$$

with  $\hat{B}$  given by equation (A2).

Differentiate  $\mathcal{D}_i$ , using equation (A12) and apply equations (A8) and (A1). This yields

$$\begin{aligned} \mathcal{D}'_i(p) &= (sn-1)b(si-1; sn-2, p) - (n-1)b(i-1; n-2, p) \\ &= (sn-1)b(si-1; sn-2, p)(1 - Kv(p)^{-(s-1)}), \text{ where} \\ K &= \frac{(n-1) \binom{n-2}{i-1}}{(sn-1) \binom{sn-2}{si-1}} \text{ and } v(p) = p^i(1-p)^{n-i}. \end{aligned}$$

Because, for  $p \in (0, 1)$ ,  $(sn-1)b(si-1; sn-2, p) > 0$ , the above equation implies that

$$\mathcal{D}'_i(p) \stackrel{\text{sign}}{=} 1 - Kv(p)^{-(s-1)}, \quad p \in (0, 1). \quad (\text{A13})$$

Because  $s > 1$ , viewed as a function of  $v$ , the right-hand side of equation (A13) has at most one positive root,  $r_v$ . Since  $p \mapsto v(p)$  is proportional to a binomial probability, it is quasi-concave in  $p$ . Thus, the preimage of  $r_v$ ,  $\{p \in (0, 1) : v(p) = r_v\}$ , contains at most two elements. Thus, by equation (A13),  $\mathcal{D}'_i$  has at most two roots in  $(0, 1)$ . This implies that  $\mathcal{D}_i$  has at most two stationary points in the interval  $(0, 1)$ , which implies, by Rolle's theorem, that  $\mathcal{D}_i$  has at most three roots in the interval  $[0, 1]$ . Since two of these roots occur at 0 and 1,

$$\mathcal{D}_i \text{ has at most one root in the interval } (0, 1). \quad (\text{A14})$$

Note that  $\hat{B}(k; m, p)$ , defined in equation (A2), has the following property:

$$\begin{aligned} \int_0^1 \hat{B}(k; m, p) dp &= \int_0^1 \text{Beta}[k, m-k+1](p) dp \\ &= 1 - \text{Mean}[\text{Beta}[k, m-k+1]] = 1 - \frac{k}{m+1}, \end{aligned} \quad (\text{A15})$$



where the first equality follows from equation (A7), the second from integration by parts, and the last from the fact that the mean of a Beta $[k, m - k + 1]$  distributed random variable equals  $k/(m + 1)$ . Thus, by equations (A12) and (A15),

$$\mathbf{I}[\mathcal{D}_i](1) = \int_0^1 \hat{B}(si; sn - 1, u) - \hat{B}(i; n - 1, u) du = 1 - \frac{si}{sn} - \left(1 - \frac{i}{n}\right) = 0.$$

Given that  $\mathcal{D}_i$  is continuous and that  $\mathbf{I}[\mathcal{D}_i](1)$ , which represents the integral of  $\mathcal{D}_i$  over  $[0, 1]$ , equals 0,  $\mathcal{D}_i$  must have at least one root in the interval  $(0, 1)$ . This observation, combined with equation (A14), implies that  $\mathcal{D}_i$  has exactly one root,  $p^*$ , in  $(0, 1)$ . Since

$$\mathcal{D}_i(p) = -\binom{n-1}{i} p^i + o(p^i), \quad p \rightarrow 0,$$

$\mathcal{D}_i$  is nonpositive for  $p \leq p^*$  and nonnegative for  $p \geq p^*$ . Thus,  $\mathbf{I}[\mathcal{D}_i]$  is quasi-convex. Evaluated at its extreme points, 0 and 1,  $\mathbf{I}[\mathcal{D}_i]$  equals 0. Thus, for all  $p \in [0, 1]$ ,  $\mathbf{I}[\mathcal{D}_i](p) \leq 0$ . This implies, by equation (A11) and the fact that  $\Delta v_i \geq 0$  for  $i = 1, \dots, n - 1$ , that condition (A10) holds. Thus, given that  $C^w$  and  $C^v$  have the same mean,  $C^w \succ_{\text{cx}} C^v$ .

Theorem 3 then follows immediately from the fact that  $C^w \succ_{\text{cx}} C^v$  and the argument used in the last two paragraphs of the proof of theorem 2. QED

The proof of corollary 3 follows immediately from that of theorem 3. QED

### A2.3. Proof of Proposition 1

We first show that  $\mu(X^w) < \mu(X^v)$  and  $X^w \preccurlyeq_{\text{FSD}} X^v$ . Given  $v_1 = 0$ , when  $w = (0, v_1, v_2, \dots, v_n)$ , we have  $\Delta w_1 = 0$  and  $\Delta w_{i+1} = \Delta v_i$  for all  $i \in \{1, \dots, n - 1\}$ . The second equality in equation (A3) then implies that

$$Q_{w,c}(p) - Q_{v,c}(p) = \sum_{i=1}^{n-1} \Delta v_i (\hat{B}(i + 1; n, p) - \hat{B}(i; n - 1, p)), \quad (\text{A16})$$

with  $\hat{B}$  given by equation (A2). Since  $\tilde{B}_{(n,p)}$  and  $\tilde{B}_{(n-1,p)}$  represent binomial  $(n, p)$  and  $(n - 1, p)$  random variables, respectively,

$$\mathbb{P}[\tilde{B}_{(n,p)} \geq i + 1] = p\mathbb{P}[\tilde{B}_{(n-1,p)} \geq i] + (1 - p)\mathbb{P}[\tilde{B}_{(n-1,p)} \geq i + 1]. \quad (\text{A17})$$

Thus, for  $p \in (0, 1)$ ,

$$\begin{aligned} \hat{B}(i + 1; n, p) - \hat{B}(i; n - 1, p) &= \mathbb{P}[\tilde{B}_{(n,p)} \geq i + 1] - \mathbb{P}[\tilde{B}_{(n-1,p)} \geq i] \\ &= (1 - p)(\mathbb{P}[\tilde{B}_{(n-1,p)} \geq i + 1] - \mathbb{P}[\tilde{B}_{(n-1,p)} \geq i]) \\ &= -(1 - p)\mathbb{P}[\tilde{B}_{(n-1,p)} = i] \\ &= -(1 - p)b(i; n - 1, p) < 0, \end{aligned} \quad (\text{A18})$$

with  $b$  defined in equation (A1), where the first equality follows from equation (A2) and the second from equation (A17). Since  $v$  is a prize vector,  $\Delta v_i \geq 0$  for  $i = 1, \dots, n - 1$  with strict inequality for some  $i$ . Thus, equations (A16) and (A18) imply that

$$Q_{w,c}(p) - Q_{v,c}(p) = -(1-p) \sum_{i=1}^{n-1} \Delta v_i b(i; n-1, p) < 0, \quad (\text{A19})$$

$p \in (0, 1)$ . This, by Shaked and Shanthikumar (2007, 5; condition 1.A.12), implies that  $C^w \preccurlyeq_{\text{FSD}} C^v$ . Since  $c^{-1}$  is increasing and FSD is closed under increasing transformations (Shaked and Shanthikumar 2007, theorem 1.A.3), we must have  $X^w = c^{-1}(C^w) \preccurlyeq_{\text{FSD}} c^{-1}(C^v) = X^v$ . Clearly,  $X^w$  and  $X^v$  are not identically distributed. Thus,  $X^w \preccurlyeq_{\text{FSD}} X^v$  implies that  $\mu(X^w) < \mu(X^v)$ .

Next, we show that entry increases expected total effort. Applying the quantile density representation for the equilibrium effort cost given by equation (A4) to the entry vector,  $w$ , noting that  $w$  has  $n+1$  prizes, yields

$$\begin{aligned} Q'_{w,c}(p) &= n \sum_{i=1}^n \Delta w_i b(i-1; n-1, p) \\ &= n \sum_{i=2}^n \Delta v_{i-1} b(i-1; n-1, p) \\ &= n \sum_{j=1}^{n-1} \Delta v_j b(j; n-1, p), \end{aligned} \quad (\text{A20})$$

where the second equality follows from the fact that  $w$ , being an entry transformation of  $v$ , satisfies that  $\Delta w_i = \Delta v_{i-1}$  for  $i \in \{2, \dots, n\}$  and  $\Delta w_1 = 0$ . The third equality follows from reindexing, using  $j = i-1$ . Equations (A19) and (A20) then imply that

$$Q_{w,c}(p) - Q_{v,c}(p) = -\frac{(1-p)}{n} Q'_{w,c}(p). \quad (\text{A21})$$

The result that entry increases expected total effort is shown by noting that

$$\begin{aligned} (n+1)\mu(X^w) - n\mu(X^v) &= \mu(X^w) + n(\mu(X^w) - \mu(X^v)) \\ &= \mu(X^w) + n \int_0^1 Q_w(p) - Q_v(p) dp \\ &\stackrel{(1)}{=} \mu(X^w) + n \int_0^1 c^{-1} \circ Q_{w,c}(p) - c^{-1} \circ Q_{v,c}(p) dp \\ &\stackrel{(2)}{\geq} \mu(X^w) + n \int_0^1 (c^{-1})' \circ Q_{w,c}(p) (Q_{w,c}(p) - Q_{v,c}(p)) dp \\ &\stackrel{(3)}{=} \mu(X^w) - \int_0^1 (1-p)(c^{-1})' \circ Q_{w,c}(p) Q'_{w,c}(p) dp = \mu(X^w) - \int_0^1 (1-p) d(c^{-1} \circ Q_{w,c}(p)) \\ &\stackrel{(4)}{=} \mu(X^w) - \int_0^1 c^{-1} \circ Q_{w,c}(p) dp \stackrel{(5)}{=} \mu(X^w) - \int_0^1 Q_w(p) dp = 0, \end{aligned}$$

where equality (1) follows from equation (A5), inequality (2) from the concavity of  $c^{-1}$ , equality (3) from equation (A21), equality (4) from integration by parts, and equality (5) from equation (A5). Also note that inequality (2) is strict if  $c^{-1}$  is strictly concave (i.e., if  $c$  is strictly convex). QED

## References

- Alsever, Jennifer. 2008. "What is Forced Ranking?" CBS News, May 1. <https://www.cbsnews.com/news/what-is-forced-ranking>.
- Andreoni, James, and Andy Brownback. 2017. "All Pay Auctions and Group Size: Grading on a Curve and Other Applications." *J. Econ. Behavior and Org.* 137:361–73.
- Arnold, Barry C. 1990. "The Lorenz Order and the Effects of Taxation Policies." *Bull. Econ. Res.* 42 (4): 249–64.
- Baker, George P., Michael C. Jensen, and Kevin J. Murphy. 1988. "Compensation and Incentives: Practice vs. Theory." *J. Finance* 43 (3): 593–616.
- Barut, Yasar, and Dan Kovenock. 1998. "The Symmetric Multiple Prize All-Pay Auction with Complete Information." *European J. Polit. Econ.* 14 (4): 627–44.
- Bates, Steve. 2003. "Forced Ranking." *HR Magazine* 48 (6): 62–68.
- Bishop, John H. 1990. "Job Performance, Turnover, and Wage Growth." *J. Labor Econ.* 8 (3): 363–86.
- Blackwell, David W., James A. Brickley, and Michael S. Weisback. 1994. "Accounting Information and Internal Performance Evaluation: Evidence from Texas Banks." *J. Accounting and Econ.* 17 (3): 331–58.
- Bodoh-Creed, Aaron L., and Brent R. Hickman. 2018. "College Assignment as a Large Contest." *J. Econ. Theory* 175:88–126.
- China Youth Daily. 2016. "56.1% of the Respondents Reported That Their Organizations Use Last-Place Elimination Systems." [http://zqb.cyol.com/html/2016-12/13/nw.D110000zgqnb\\_20161213\\_1-07.htm](http://zqb.cyol.com/html/2016-12/13/nw.D110000zgqnb_20161213_1-07.htm) (in Chinese).
- Chowdhury, Subhasish M., and Sang-Hyun Kim. 2017. "'Small, yet Beautiful': Reconsidering the Optimal Design of Multi-winner Contests." *Games and Econ. Behavior* 104:486–93.
- Churchill, Gilbert A., Neil M. Ford, and Orville C. Walker. 1997. *Sales Force Management: Planning, Implementation, and Control*. Homewood, IL: Richard D. Irwin.
- Cichello, Michael S., C. Edward Fee, Charles J. Hadlock, and Ramana Sonti. 2009. "Promotions, Turnover, and Performance Evaluation: Evidence from the Careers of Division Managers." *Accounting Rev.* 84 (4):1119–43.
- Clark, Derek J., and Christian Riis. 1998. "Influence and the Discretionary Allocation of Several Prizes." *European J. Polit. Econ.* 14 (4):605–25.
- Del Guercio, Diane, and Paula A. Tkac. 2008. "Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow." *J. Financial and Quantitative Analysis* 43 (4): 907–36.
- Drugov, Mikhail, and Dmitry Ryvkin. 2019. "Tournament Rewards and Heavy Tails." Working paper.
- Dubey, Pradeep, and John Geanakoplos. 2010. "Grading Exams: 100, 99, 98, . . . or A, B, C?" *Games and Econ. Behavior* 69 (1): 72–94.
- Fu, Qiang, and Jingfeng Lu. 2009. "The Beauty of 'Bigness': On Optimal Design of Multi-winner Contests." *Games and Econ. Behavior* 66 (1): 146–61.
- . 2012. "Micro Foundations of Multi-prize Lottery Contests: A Perspective of Noisy Performance Ranking." *Social Choice and Welfare* 38 (3): 497–517.
- Garcia, Stephen M., and Avishalom Tor. 2009. "The N-Effect: More Competitors, Less Competition." *Psychological Sci.* 20 (7): 871–77.
- Gladwell, Malcolm. 2002. "The Talent Myth." *New Yorker*, July 22, 28–33.
- Glazer, Amihai, and Refael Hassin. 1988. "Optimal Contests." *Econ. Inquiry* 26 (1): 133–43.
- Hadar, Josef, and William R. Russell. 1969. "Rules for Ordering Uncertain Prospects." *A.E.R.* 59 (1): 25–34.

- Halac, Marina, Navin Kartik, and Qingmin Liu. 2017. "Contests for Experimentation." *J.P.E.* 125 (5): 1523–69.
- Hardy, Godfrey Harold, John Edensor Littlewood, and George Pólya. 1952. *Inequalities*. 2nd ed. Cambridge: Cambridge Univ. Press.
- Hvide, Hans K. 2002. "Tournament Rewards and Risk Taking." *J. Labor Econ.* 20 (4): 877–98.
- Kalra, Ajay, and Mengze Shi. 2001. "Designing Optimal Sales Contests: A Theoretical Perspective." *Marketing Sci.* 20 (2): 170–93.
- Krishna, Vijay, and John Morgan. 1998. "The Winner-Take-All Principle in Small Tournaments." In *Advances in Applied Microeconomics*, vol. 7, *Contests*, edited by Michael R. Baye, 61–74. Greenwich, CT: JAI.
- Lazear, Edward P. 1989. "Pay Equality and Industrial Politics." *J.P.E.* 97 (3): 561–80.
- Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *J.P.E.* 89 (5): 841–64.
- Lim, Noah. 2010. "Social Loss Aversion and Optimal Contest Design." *J. Marketing Res.* 47 (4): 777–87.
- Lim, Noah, Michael J. Ahearne, and Sung H. Ham. 2009. "Designing Sales Contests: Does the Prize Structure Matter?" *J. Marketing Res.* 46 (3): 356–71.
- Ma, Linlin, Yuehua Tang, and Juan-Pedro Gómez. 2019. "Portfolio Manager Compensation in the US Mutual Fund Industry." *J. Finance* 74 (2): 587–638.
- Medoff, James L., and Katharine G. Abraham. 1980. "Experience, Performance, and Earnings." *Q.J.E.* 95 (4): 703–36.
- Megidish, Reut, and Aner Sela. 2013. "Allocation of Prizes in Contests with Participation Constraints." *J. Econ. and Management Strategy* 22 (4): 713–27.
- Moldovanu, Benny, and Aner Sela. 2001. "The Optimal Allocation of Prizes in Contests." *A.E.R.* 91 (3): 542–58.
- Olszewski, Wojciech, and Ron Siegel. 2018a. "Large Contests without Single Crossing." Manuscript, Dept. Econ., Northwestern Univ.
- . 2018b. "Pareto Improvements in the Contest for College Admissions." Manuscript, Dept. Econ., Northwestern Univ.
- . 2018c. "Performance-Maximizing Contests." Manuscript, Dept. Econ., Northwestern Univ.
- Schweitzer, Paul, and Ella Segev. 2012. "The Optimal Prize Structure of Symmetric Tullock Contests." *Public Choice* 153 (1): 69–82.
- Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic Orders*. New York: Springer Sci. & Bus.
- Szymanski, Stefan, and Tommaso M. Valletti. 2005. "Incentive Effects of Second Prizes." *European J. Polit. Econ.* 21 (2): 467–81.
- Thon, Dominique, and Stein W. Wallace. 2004. "Dalton Transfers, Inequality and Altruism." *Social Choice and Welfare* 22 (3): 447–65.
- Xiao, Jun. 2018. "Equilibrium Analysis of the All-Pay Contest with Two Nonidentical Prizes: Complete Results." *J. Math. Econ.* 74:21–34.
- Xu, Mouquan. 2016. "Supreme Court: Dismissals of Least Effective Workers are Illegal." NewsChina, December 2. [http://newschinamag.com/newschina/print.do?article\\_id=981&section\\_id=26&magazine\\_id=](http://newschinamag.com/newschina/print.do?article_id=981&section_id=26&magazine_id=).