Does the cream rise to the top? Luck, talent, success, and merit*

Thomas Noe[†]
Balliol College/SBS, Oxford

Dawei Fang[‡]
University of Gothenburg

December 16, 2020

Abstract

Are the winners of highly competitive, high-stakes contests, talented or merely lucky? We address this question by characterizing the relationship between talent, luck, and top performance in the limit as the number of competitors grows without bound. We show that, for many standard statistical distributions of luck, increasing contest competitiveness never ensures that the top performing contestant is talented. In some cases, increasing competitiveness even favors the talentless. We apply our results to a number of economic problems: e.g., identifying ability, effecting social mobility, designing performance metrics, and assessing the effects of globalization on winner-take-all talent competitions.

*We thank Luis Corchon, Mikhail Drugov, Xavier Gabaix, Igor Letina, Marc Möller, Junjie Zhou and seminar audiences at University of Bern, University of Gothenburg, University of Nottingham, Research Institute of Industrial Economics (IFN), the Econometric Society World Congress, and the e-conference on "Contests: Theory and Evidence" for helpful comments and suggestions. Any remaining errors are our responsibility. Financial support by Browaldh Stiftelsen (BFh18-0010) and VINNOVA (2010-02449) is gratefully acknowledged.

[†]Saïd Business School, Park End Street, Oxford OX11HP UK, email:thomas.noe@sbs.ox.ac.uk

[‡]Department of Economics, P.O. Box 640, Vasagatan 1, Gothenburg 40530 Sweden, email:dawei.fang@economics.gu.se

Success frequently depends on winning high-stakes competitions. High-stakes tests, like the SAT and GRE in the US, and the A-levels in the UK, profoundly affect admissions decision at elite universities; National examinations (Gaokao) essentially determine college placement in China. Elite university degrees have significant effects on socio-economic status (Jia and Li, 2017; McDonough, 1997). In the presence of network externalities, financial constraints, and learning effects, a single decision conferring a slight advantage to a start-up firm over its rivals can lead to dominance in an emerging product market (Katz and Shapiro, 1994). A single successful prediction of stock market crash or large market rally can make a stock analyst's reputation (Denrell, 2013). Winning an early-career grant can significantly bolster a scholar's chance of winning future grants and a full professorship (Bol et al., 2018).

Few would doubt that random chance has some effect on performance in these high-stakes contests and thus on the future prospects of individual competitors. However, many high-stakes contests are also highly competitive, i.e., there are many more candidates worthy of selection than places. In such cases, can the force of competition ensure that those selected are worthy?

For example, suppose we establish a truly elite university, one that admits one student per year based on a single examination. Some fixed fraction of potential applicants are talented and worthy of admission and some are average and unworthy. Assuming the number of applicants, and thus talented applicants, increases without bound, can we be virtually assured that the one admitted student is talented? This paper will show that, generally, the answer to this question is no. Even when the luck component of performance is small and produced by draws from standard, innocuous statistical distributions, regardless of the number of competing students, the probability that an average ability student is admitted will frequently be bounded away from zero. In fact, in some cases, making the admissions contest more competitive by increasing the number of applicants *reduces* the probability that the selected student is talented.

We establish these results in a very simple framework: agents compete by submitting random performance. Some agents are *talented* and others are *average*. An agent's type determines the distribution of the agent's random performance. Talented agents have stochastically better performance than average agents. We fix the proportion of talented agents in the population between zero and one and ask the question of whether, when the number of competitors increases without bound, the probability that the best performer is talented converges to one, i.e., we consider the question of whether competition *identifies talent*.

Theorem 1 in Section 1 establishes a simple necessary and sufficient condition for talent identification: the ratio between the survival functions of the talented and average agents' performance distributions converges to infinity as performance converges to its (possibly infinite) upper bound.[1]

As well as being simple, the talent identification condition yields a surprising conclusion:

[1] The survival function of a distribution, F, is given by $1 - F$.

For quite innocuous specifications of performance distributions (e.g., Logistic, Laplace, Max Extreme Value), the talent identification condition fails.[2] For example, suppose that agent performance equals agent ability plus an i.i.d. unit variance, logistically distributed noise term, 5% of the competitors are talented, and the ability difference between talented and average competitors is one unit. Even when there are 2 million competitors, the best performer will be average about 75% of the time. Because, in this example, a talented agent's performance dominates an average agent's performance in the monotone likelihood ratio (MLR) order, any cutoff for elite performance less than best performance would lead to an even higher proportion of "elite performers" being lucky but untalented.[3]

When the talent identification condition is satisfied (e.g., when performance distributions are Normal), the effect of luck on selection can be made arbitrarily small by a sufficiently large increase in competitiveness. However, when the signal-to-noise ratio is low, "sufficiently large" can be practically infeasible. Application 2 considers the problem of identifying talented mutual fund managers. Assuming normally distributed managerial excess performance and using published estimates of the "signal," the excess return of a talented manager, and "noise," the standard deviations of excess returns, we show that if 5% of fund managers are talented, even when the assumed number of competing funds vastly exceeds the current worldwide count, at least 92% of top performing fund managers will have average ability. Consequently, even if a small fraction of fund managers can deliver exceptional performance, the expected excess mean performance of best performing "elite" fund managers will be almost indistinguishable from zero.

These results follow from the nature of rank-based selection of best performers in high-stakes contests. The probability that the best performer will be talented is increasing in a likelihood ratio: the ratio between the probability that a talented competitor submits the best performance and the probability that an average competitor submits the best performance. Consider the case where a talented agent's performance MLR dominates an average agent's performance. In this case, if "best performance" were defined as topping a fixed threshold performance level, increasing the threshold would indeed increase the likelihood ratio and thus the probability that the best performer is talented.

In our setting, "best performance" requires topping the random performance of rivals, not topping a fixed threshold. However, the map between the number of competitors and the best-performance threshold is increasing in the sense of first-order stochastic dominance.[4] Thus,

[2]The Max Extreme Value distribution is sometimes also termed the "Gumbel Distribution," the "Double Exponential Distribution," or the"Generalized Extreme Value Distribution Type-I."

[3]For any two distribution functions F_1 and F_2, F_1 is larger than F_2 under the MLR order if the ratio between the distributions' density functions, f_1/f_2, increases over the union of the interiors of the supports of F_1 and F_2 (Shaked and Shanthikumar, 2007, Section 1.C).

[4]For any two distribution functions F_1 and F_2, F_1 is larger than F_2 in the sense of first-order stochastic dominance if $F_1(x) \leq F_2(x)$ for all x.

it seems plausible that increased competitiveness will favor the talented. We call the effect increasing competition exerts through increasing the threshold for best performance, the *game-elevation effect*. The game-elevation effect appears to motivate common assertions that competition weeds out mediocre talents, e.g., “As global competition increases, ...only the best can stay in the game” (Jelinek and Adler, 1988).

However, for two reasons, our analysis does not generally support the tight association between ability and elite performance suggested by such assertions. First, as shown by Theorem 1, the force of the game elevation effect is, in many cases, quite weak, and unable to weed out average competitors even when the intensity of competition is unbounded.

Second, the game-elevation effect abstracts from the fact that best performance is defined relative to other competitors; thus, the threshold for best performance is random and the distribution of this random threshold is different for average and talented competitors.[5] MLR ordering only guarantees that increasing the best-performance threshold increases the likelihood ratio when talented and average competitors’ prospects are evaluated at the same threshold.

Consequently, as we illustrate in Section 2, because of the *differential-threshold effect*, even for standard statistical distributions of luck, increased competition can *reduce* the probability that the best performer is talented. Proposition 1 provides sufficient conditions that rule in and rule out such reversals. These conditions are based on qualitative features of the distribution of luck, not on “size” of the luck component relative to the ability component of performance.

The relationship between best performance and ability is particularly weak in the high-stakes contests that have a natural upper bound on performance. Examples include target-hitting contests, in which the best possible performance is making no error, and innovation races, in which the best possible performance is discovering the innovation immediately. We show in Section 3 that the talent identification condition always fails in these contests. As we illustrate with an example related to racing contests, even a small reduction in the stakes of the contest, effected by measuring performance based on the time to completion of two tasks rather than of one, produces a larger association between best performance and talent than an arbitrarily large increase in the number of competitors.

Finally, in Section 4, we extend our analysis of the effect of competition on talent identification in a number of directions. We consider endogenous contest participation and use our results to characterize the effect of globalizing local winner-take-all markets for artistic talent. We show that, although the opportunity cost of contest participation sometimes leads average ability agents to opt out of high-stakes contests, in general, endogenous participation does not lead to talent identification. Next, we expand our consideration of competition’s effect to consider the overall relationship between performance and ability rather than simply

[5]For example, in a contest with one talented and one average competitor, the random threshold for the talented competitor follows the distribution of the average competitor’s performance, whereas the random threshold for the average competitor follows the distribution of the talented competitor’s performance.

efficient identification of extraordinary talent. We provide conditions under which allowing agents to make multiple attempts (counting only their best attempt) strengthens and weakens the performance-ability relationship. We apply these results to evaluating a recent reform of Chinese university admissions examinations.

Our analysis has a number of implications for policy and research. First, the efficiency of task assignment based on rank-order competition in high-stakes meritocratic contests (e.g. examinations, employment tests) depends on the degree to which such contests actually identify talent. Our results can thus be thought of as characterizing an upper bound for the selection efficiency of high-stakes competitions.

Second, for elite competitions, the talent identification condition (Theorem 1) provides an alternative measure for the selection efficiency of performance metrics. For example, consider the debate over the role of standardized tests in admissions decisions at elite universities. Supporters of using standardized tests argue that these standardized test scores are unbiased and the most precise predictors of aptitude available (Payne, 2018).[6] Opponents of standardized tests argue for relying solely on "holistic metrics" such as grades, recommendation letters, personal statements, etc. The opponents raise various objections related to bias, stress, financial burden, etc (Ledford et al., 2020).

This debate elides a rather important question: is precision, measured by signal-to-noise ratio, the best measure of selection efficiency in elite contests? Signal precision measures the extent to which performance reveals ability differences between two randomly selected agents of differing ability. What matters for selection efficiency in elite contests is the ability of performance to reveal ability differences *conditioned* on performance being drawn from the extreme right tails of the competitors' performance distributions. In the limit, as the number of competitors grows without bound, this "conditional precision" will depend only on the tail ratio. If performance distributions under holistic and standardized metrics differ by more than scale and location, there is no a priori reason to believe that the lower precision metric will yield a lower probability of selecting apt students than the higher precision metric.

Third, empirical research frequently uses extraordinary relative performance as a proxy variable for unobservable talent. The "talent" variable is then employed as an independent variable in a prediction regression whose dependent variable is some aspect of agent behavior (Ding and Wermers, 2004; Franck and Nüesch, 2012; Leone and Wu, 2007). In such studies, the extent to which extraordinary performance is a noisy proxy for talent will affect the power of these tests as well as the bias introduced by errors in variables. Weak identification can lead to the spurious conclusion that talent does not matter when, in fact, talent does matter but the pool of "top performers" is only marginally more talented than a randomly selected contestant

[6] Hoffman et al. (2018) empirically find, in the context of managers making hiring decisions, that the average quality of hired workers improves if managers are not allowed to overrule job-test recommendations.

pool.

Related literature

Our paper belongs to the literature on contests as selection mechanisms. Most of this literature focuses on designing contests that maximize the chance of selecting the most able agents when selection outcomes are affected by agents' effort choice (Clark and Riis, 2001, 2007; Kawamura and Moreno de Barreda, 2014), risk taking strategies (Fang and Noe, 2019; Hvide and Kristiansen, 2003), sabotage strategies (Münster, 2007), or participation decisions (Hizen and Okui, 2009). In contrast, the focus of our paper is not on optimal designs of selection contests, but on evaluating the extent to which exceptional contest performance reveals talent. For most of our analysis, we abstract from agent strategies.[7] Thus, the limited power of contests for selecting the talented identified in our paper can be considered an intrinsic feature of selection contests, provided the performance distributions produced by agent strategies satisfy the weak regularity conditions of our framework.[8]

The framing of our analysis is closely related to the framing of noisy contest models in which agent performance is affected by luck, i.e., exogenous noise, that is independent of the characteristics or efforts of the contestants. In a noisy contest setting, Ryvkin and Ortmann (2008) and Ryvkin (2010) show that increasing competition, which, as in our paper, is defined as increasing the number of contestants, increases winner quality. However, they do not consider the question we consider, the conditions under which, in the limit, competition can eliminate the effects of luck on selection.

Like our paper, Drugov and Ryvkin (2019) considers the effect of the shape of the luck distribution and competition on the outcomes of tournaments. In contrast to our analysis, Drugov and Ryvkin (2019) analysis centers on the effects of luck distributions on contestant effort. We focus on their effects on selection. Because of this fundamental difference between our analysis and theirs, the distributional restrictions required to characterize the effect of competition on effort in Drugov and Ryvkin (2019) are very different than the conditions we require to characterize the relationship between competition and selection efficiency.

Our paper also contributes to prior work on the role of luck in explaining success. In this literature, "success" is defined as submitting a *high level* of performance. Hoffman (2010) and Taleb (2005) point out that, if a large number of untalented agents try their luck, some will succeed. Anecdotal evidence (Frank, 2016) and simulation results (Pluchino et al., 2018)

[7]We consider agent participation decisions in Section 4.1.

[8]In fact, in noisy contests, with the winner's prize fixed, when contests become too competitive, all contestants, talented and average, will exert little effort. The reduction in the effort difference between talented and average agents will make exceptional performance a noisier signal of ability. Thus, introducing effort into our framework would further strengthen our results that increasing competition may favor luck rather than talent and, in many cases, no degree of competitiveness can ensure that the best performer has talent.

suggest that, if success is defined as reaching an elite social and economic status, the primary driver of success is luck, not talent. Unlike these analyses, our paper considers the effect of luck on success when success is defined as reaching the *highest* performance *rank*. Moreover, our analysis precisely characterizes the roles of luck and talent in determining success when the intensity of competition is unbounded.

Denrell and Liu (2012) and Denrell et al. (2017) consider the question of whether high observed levels of performance signal high ability. They show that, when more able contestants' performance dominates less able contestants' in the sense of first-order stochastic dominance, but not in the sense of hazard rate or MLR dominance, higher performance can signal lower ability. In contrast, we consider the question of how increasing the number of competitors affects the expected quality of top-ranked performers in a setting where ability differences can engender MLR dominance. In fact, we show that, even when ability differences do lead to MLR (and thus also hazard rate) dominance, because of a subtle differential-threshold effect, increased competitiveness can reduce the probability that top-ranked performers have talent.

The research question we address is also related to recent empirical research. Although most empirical research focused on the relationship between talent and performance either attempts to predict performance based on proxies of talent or, as discussed above, uses performance as a proxy for talent, recently empiricists have turned their attention to an empirical question closely related to our analysis—the extent to which past contest performances reveal contestant ability (e.g., Cuthbertson et al., 2008; Migheli, 2019). Our analysis identifies the distributional properties of luck which affect the informativeness of past elite performance in large competitions.

1 Model

1.1 Setup

Consider a collection of C agents, indexed by $i \in \{1,2,\ldots C-1,C\}$, who we will term *competitors*. Competitors submit *performance*. Performance is random. However, the distribution of performance depends on the competitors' type. There are two types of competitors: *talented* competitors and *average* competitors.

The distribution of performance for talented and average competitors are represented by F_T and F_A, respectively. We assume that these distributions have the nice properties associated with textbook distribution functions, no level of performance can identify the type of a competitor with certainty, and talented competitors' performance distribution is better than the performance distribution of average competitors.

Assumption 1. The performance distributions of average and talented competitors, denoted by

F_A and F_T, respectively, satisfy the following conditions:

(i) F_A and F_T are absolutely continuous, and have probability density functions, f_A and f_T, which are continuous and strictly positive over the interior of the supports of their respective distributions.

(ii) The support of F_A and F_T is $[\ell, h]$, where $-\infty \leq \ell < h \leq \infty$.

(iii) F_T strictly stochastically dominates F_A, i.e. for all $x \in (\ell, h)$, $F_T(x) < F_A(x)$.

Assumption 1.i for the most part represents regularity conditions for the compared performance distributions. These assumptions could be considerably relaxed at the cost of increased technical complications. However, the scope of application for our results is not significantly expanded by relaxing these conditions. The one implication of Assumption 1.i that is crucial for our analysis is that the continuity of the types' distribution functions ensures that the probability that two competitors will submit the same performance is zero. Thus, the best performing competitor can be unambiguously identified.

Assumption 1.ii implies that the upper bound on performance for talented and average competitors is the same. Absent this assumption, the problem we consider—whether talent can be identified by performance in the limit as the number of competitors increases without bound—becomes trivial. With probability 1, eventually a talented competitor will submit performance that exceeds the upper bound of the performance of average competitors. Since the best performance must exceed this performance level, the best performing competitor, in the limit, will always be talented. The lower bound on performance plays almost no role in our analysis.[9] However, permitting different lower bounds for talented and average competitors would complicate some of our subsequent definitions because we would need to avoid division by 0.

Assumption 1.iii embeds the notion that talented competitors, on average, perform better than average competitors under the weakest possible definition of distributional dominance, first-order stochastic dominance. In most of the applications we subsequently consider, even stronger notions of dominance are imposed, such as MLR dominance.

We assume that (a) competitors' performances are jointly independent, (b) all competitors have the same ex ante probability of being talented and thus a competitor's index provides no information about the competitor's type, and (c) the total number of average competitors is fixed and given by $a \in \{1, 2, \ldots C-1\}$.

Assumption 2. (i) For all $i \in \{1, 2, \ldots C\}$, $X_A^i \overset{d}{\sim} F_A$ and $X_T^i \overset{d}{\sim} F_T$.

(ii) Let $\{X_A^i\}_{i=1}^C$ and $\{X_T^i\}_{i=1}^C$ be collections of random variables such that $\{X_A^i\}_{i=1}^C \cup \{X_T^i\}_{i=1}^C$ are jointly independent.

(iii) The performance of competitor i, X_i, is given by $X_i = I^i X_A^i + (1 - I^i) X_T^i$, where $\{I^i\}_{i=1}^C$ is a collection of random indicator functions that satisfies the following conditions:

[9] The only exception is Application 1, where we investigate the expected ability of the worst performer and the assumption of a common lower bound makes the analysis nontrivial.

a. $\{I^i\}_{i=1}^C$ is independent of $\{X_A^i\}_{i=1}^C \cup \{X_T^i\}_{i=1}^C$
b. $\mathbb{P}[I^i = 1] = \mathbb{P}[I^j = 1]$, for all i, j and
c. there exists an integer $a \in \{1, 2, \ldots C-1\}$ such that $\sum_{i=1}^C I^i = a$ with probability 1.

In our context, independent performance, Assumption 2.ii, is a fairly innocuous assumption. Dependence between the performance levels of the competitors induced by a common component would have no effect on the ranking of performance, the object of our analysis. Parts (a) and (b) of Assumption 2.iii restrict our analysis to type (average or talented) assignment mechanisms under which a competitor's index has no effect on the probability that the competitor is talented. This assumption is not necessary to derive our result. However, index independence does ensure that the probability that the best performing competitor is average equals the probability that a competitor is average conditioned on being the best performer. Thus, it permits us to interpret our results as statements about how performance affects inferences about competitor talent.

Part (c) of Assumption 2.iii implies that there is no aggregate uncertainty regarding the number of talented competitors. The assumption of no aggregate uncertainty simplifies our analysis somewhat. In addition, as we show in Appendix C.2, introducing aggregate uncertainty increases the probability that the best performer is average. Because the thrust of our analysis is that frequently the probability that the best performer is average is quite large even in highly competitive environments, the no aggregate uncertainty assumption does not inflate the magnitude of the effects we identify.

1.2 Performance and ability

Our goal is to characterize the relationship between best performance and underlying ability. To do this, we need to compute the probability that the best performer is average. The best performer will be average if and only if the best performance of average competitors exceeds the best performance of talented competitors. Thus, if we define best performance for average and talented competitors by

$$\bar{X}_A = \max[X_i : I^i = 1],\ \bar{X}_A \stackrel{d}{\sim} \bar{F}_A \quad \text{and} \quad \bar{X}_T = \max[X_i : I^i = 0],\ \bar{X}_T \stackrel{d}{\sim} \bar{F}_T,$$

then, if the best performance is submitted by an average competitor, it must be the case that $\{\bar{X}_A > \bar{X}_T\}$. Because performance distributions are continuous, this event has the same probability as $\{\bar{X}_A \geq \bar{X}_T\}$. Thus, by Assumptions 1 and 2, the probability that the best performer is average, which we denote by Π_A, is given by

$$\Pi_A = \mathbb{P}[\bar{X}_A \geq \bar{X}_T] = \int_\ell^h \mathbb{P}[\bar{X}_T \leq x]\, d\bar{F}_A(x) = \int_\ell^h \bar{F}_T(x)\, d\bar{F}_A(x). \tag{1}$$

Define the function $u : [0, 1] \to [0, 1]$, by

$$u(s) = \begin{cases} 0 & s = 0 \\ F_T \circ F_A^{-1}(s) & s \in (0,1) \\ 1 & s = 1 \end{cases}. \tag{2}$$

Note that u is simply the unique continuous extension of $t \hookrightarrow F_T \circ F_A^{-1}(t)$, which is defined over $(0,1)$, to $[0,1]$.

Next, let $t = C - a$ denote the number of talented competitors. Note that the best performance for average (talented) competitors is the maximum order statistic for a (t) independent draws from F_A (F_T). Thus, for $x \in [\ell, h]$, $\bar{F}_A(x) = (F_A(x))^a$ and $\bar{F}_T(x) = (F_T(x))^t$. The inverse function for $\bar{F}_A$, i.e., its quantile function, is given by $\bar{F}_A^{-1}(s) = F_A^{-1}(s^{1/a})$, $s \in (0,1)$. Using equation (2) and change of variables, we can express equation (1) as follows:

$$\Pi_A = \int_0^1 \left(u(s^{1/a})\right)^t ds. \tag{3}$$

As equation (3) implies, what is relevant for determining Π_A is how the quantiles of performance for the two types are related. For a given $s \in (0,1)$, $\bar{F}_A^{-1}(s)$ represents the s-quantile of the distribution of the best performance of average competitors. The probability that the best performance of average competitors is no greater than the s-quantile equals the probability that the performance of all average competitors is no greater than the s-quantile, i.e., $\mathbb{P}[X_A \leq \bar{F}_A^{-1}(s)]^a$. Because $\bar{F}_A^{-1}(s)$ is the s-quantile of the best performance of average competitors, $\mathbb{P}[X_A \leq \bar{F}_A^{-1}(s)]^a = s$; so $\mathbb{P}[X_A \leq \bar{F}_A^{-1}(s)] = s^{1/a}$. Thus, the $s^{1/a}$ quantile of an average competitor's performance equals the s-quantile of the best performance of average competitors. This performance quantile, $F_A^{-1}(s^{1/a})$, is sufficient to top the performance of all talented competitors with probability $\mathbb{P}[X_T \leq F_A^{-1}(s^{1/a})]^t = (F_T \circ F_A^{-1}(s^{1/a}))^t$. Inspection of the definition of u provided by equation (2), shows that this is exactly the expression integrated on the right-hand side of equation (3).

1.3 The effect of competition on talent identification

To capture the effect of competition, assume that $a = n a_o$ and $t = n t_o$, where t_o and a_o are positive integers and n is an integer that affects the scale (i.e., number of competitors) in the contest. Increased competition is modeled as an increase in n. Using equation (3), we express the probability that the best performer is average, Π_A, given the scale factor, n, as follows

$$\begin{aligned} \Pi_A(n) &= \int_0^1 U_n(s)\, ds, \text{ where} \\ U_n(s) &= \left(u(s^{1/(a_o n)})\right)^{t_o n}. \end{aligned} \tag{4}$$

When $\lim_{n\to\infty} \Pi_A(n) = (>)0$, we say that talent is (not) identified by competition.

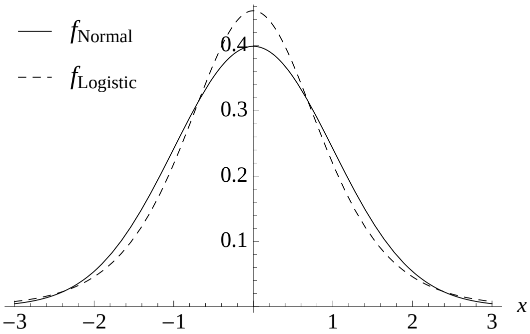
Example 1. *Talent identification: Normal vs. Logistic.* What are the implications of equation (4) for the effect of competition on talent identification? We initiate our discussion of this question with a simple example. In this example, we compare the probability that the best performer is average under two scenarios. In both scenarios, the standard deviation of performance equals 1, the difference between the mean performance of talented and average competitors equals 1, and the ratio between the number of talented and average competitors, t_o/a_o, also equals 1.

The scenarios differ only with respect to the "shape" of the distribution of performance. In the first scenario, the performance of both types of competitors is Normally distributed. In the second, performance is Logistically distributed. The Logistic distribution is quite innocuous—like the Normal distribution, the Logistic distribution has a log-concave density function, is not heavy tailed, and has moments of all orders.[10] In latent variable specifications of the Logit and Probit qualitative choice models, the Logistic distribution plays the same role in the Logit model as the Normal distribution plays in the Probit model. It is often asserted that the results of Logit and Probit regressions are usually "quite similar."[11] Thus, one might expect that the effect of competition under the two scenarios will be quite similar. However, as the results in Figure 1 suggest, the power of competition for identifying talent is quite different in the two scenarios.

Example 1 shows that, when the number of competitors is small, the probability that the best performer is average, Π_A, is quite similar under the Normal and Logistic specifications. However, as the number of competitors increases, Π_A falls at a much more rapid rate under the Normal specification and becomes negligible when the number of competitors is very large. In contrast, under the Logistic specification, although Π_A falls with the number of competitors, the rate of descent is quite slow. Even when there are 2 million competitors, the probability that the best competitor is average is approximately 14%. The example thus suggests that under the Normal specification, talent will be identified but under the Logistic specification, it will not.

[10] A function is log-concave if the logarithm of the function is concave. By Shaked and Shanthikumar (2007, Theorem 1.C.66), when the compared distributions have log-concave densities and differ only in location, first-order stochastic dominance implies MLR dominance.

[11] "Estimates from both [Logit and Probit] models produce similar results, and using one or the other is a matter of habit or preference" (https://blog.stata.com/2016/01/07/probit-or-logit-ladies-and-gentlemen-pick-your-weapon/).



A. PDFs: Error distributions

	Π_A	
# Competitors	Normal	Logistic
2	0.2398	0.2273
4	0.1951	0.1860
8	0.1550	0.1629
200	0.0493	0.1410
2,000	0.0242	0.1403
2,000,000	0.0048	0.1402

B. Prob. best performer is average

Figure 1: In the figure, the expected performance of talented competitors and the expected performance of average competitors satisfy $\mu_T - \mu_A = 1$. Half of the competitors are talented. Realized performance is distributed $\mu_i + \tilde{\varepsilon}$, $i = T, A$ and $\tilde{\varepsilon}$ is an iid, zero mean, unit variance error term that is either Normally or Logistically distributed.

Our next result, Theorem 1 shows that this conjecture is correct. In fact, under a very mild hypothesis—that the limit of the tail ratio between the talented and average competitors' performance distributions, $(1-F_T)/(1-F_A)$, exists—Theorem 1 provides general conditions for talent identification and, when talent is not identified, provides the asymptotic probability that the best performer is average. The proof of this result and all subsequent proofs are provided in Appendix A.[12]

Theorem 1. *If* $\lim_{x \to h}(1-F_T(x))/(1-F_A(x)) = L$ *exists (possibly infinite), then*

(i) if $L = \infty$*, talent is identified by competition, i.e.,* $\lim_{n\to\infty} \Pi_A(n) = 0$.

(ii) If $L < \infty$*, talent is not identified by competition and the asymptotic probability that the best performer is average is given by*

$$\lim_{n\to\infty} \Pi_A(n) = \frac{a_o}{a_o + L t_o}.$$

The intuition for the identification condition is as follows. The probability that the best performer is average is decreasing in a likelihood ratio: the ratio between the probability that a given talented competitor submits the best performance and the probability that a given average competitor submits the best performance. As competition increases unboundedly, in the limit, this likelihood ratio becomes the tail ratio between the talented and average competitors' performance distributions due to a *game-elevation effect*: increasing competition increases the threshold performance for being the best. As competition increases without bound, this threshold approaches the (possibly infinite) upper bound on performance. Thus, as implied by Theorem 1, talent identification is determined by the limit of the tail ratio, L, and the asymptotic

[12]Theorem 1 also holds when the assignment of talent to competitors does not feature a fixed number of talented and average competitors (part (c) of Assumption 2.iii) provided that with probability 1, the fraction of average competitors converges to $a_o/(a_o + t_o)$ as $n \to \infty$. For a detailed discussion, see Appendix C.1.

probability that the best performer is average is decreasing in L.

The condition for talent identification is not easy to satisfy. For example, suppose that competitor performance equals ability plus i.i.d. noise. When the noise distribution is Laplace, Logistic, or Hyperbolic Secant, the talent identification condition is never satisfied. Thus, talent identification can fail even when the talented and average competitors' performance distributions are symmetric and unimodal, have log-concave densities, moments of all orders and are not heavy tailed. When the performance distributions are more irregular, identification can fail in an even more spectacular fashion. For example, when the additive noise is Student's-t or Cauchy, neither of which has log-concave densities or moments of all orders, $L = 1$, and thus best performance provides *no* information about ability.[13] The talent identification condition is satisfied when the additive noise is Normal. Consequently, talent identification through competition depends on a rather subtle condition restricting the relationship between distributions of talented and average competitors' performance. In many economic applications, there is no a priori reason to believe that this condition is satisfied.

Finally, note that the only condition required to apply Theorem 1 is that the tail ratio has a limit as performance approaches its (possibly infinite) upper bound.[14] This is a very easy condition to satisfy. A sufficient, but by no means necessary, condition for the existence of a tail ratio limit, is that F_T hazard-rate dominates F_A.[15] Hazard rate dominance is a weaker restriction than MLR dominance (Shaked and Shanthikumar, 2007, Theorem 1.C.1).

Theorem 1 resolves the conjectures posed in Example 1. First, consider the Logistic scenario and note that the scale parameter of Logistic distributions with unit variance is $\sqrt{3}/\pi$. When both F_T and F_A are Logistic with scale parameter $\sqrt{3}/\pi$, and the mean difference between talented and average competitors' performance equals 1, the limit of the tail ratio equals

$$L = \lim_{x \to \infty} \frac{1 - F_T(x)}{1 - F_A(x)} = e^{\pi/\sqrt{3}} \approx 6.134.$$

Thus, in the Logistic scenario, $L < \infty$ and, hence, Theorem 1.ii implies that talent is not identified by competition. Based on the formula in Theorem 1.ii, the asymptotic probability that the

[13] Because the Student's-t and the Cauchy distributions do not have log-concave densities, by Shaked and Shanthikumar (2007, Theorem 1.C.66), when the additive noise is Student's-t or Cauchy, F_T does not MLR dominate F_A, in which case better performance may not be a stronger signal of talent.

[14] In some cases, it is possible to determine if talent is identified even when the hypothesis of Theorem 1, that the tail ratio converges as $x \to \infty$, is not satisfied. Because, in applications, nonconvergent tail ratios are unlikely to arise, we defer discussion of these conditions to Appendix B.

[15] For any two random variables, X and Y, with distribution functions denoted by F_X and F_Y respectively and the upper bound of the support denoted by h_X and h_Y respectively, X hazard-rate dominates Y if $x \hookrightarrow (1 - F_X(x))/(1 - F_Y(x))$ is nondecreasing for $x \in (-\infty, \max[h_X, h_Y])$ (here $a/0$ is taken to equal ∞ whenever $a > 0$).

best performer is average is given by

$$\frac{a_o}{a_o + Lt_o} = \frac{1}{1+L} = \frac{1}{1+e^{\pi/\sqrt{3}}} \approx 0.14, \tag{5}$$

which is consistent with the numerical results in Panel B of Figure 1. Performing the same computation in the Normal scenario shows that $L = \infty$. Thus, Theorem 1.i implies that competition identifies talent in the Normal scenario. This result is consistent with the negligible probability that the best performer is average when the number of competitors equals 2 million, computed in Panel B of Figure 1.

In fact, when competitor performance equals ability plus i.i.d. noise, the talent identification condition boils down to a condition on the right tail of the noise distribution.

Corollary 1. *Suppose that, for $t = T, A$, $X_t \overset{dist}{=} \mu_t + \varepsilon$, $\mu_T > \mu_A$, and $\varepsilon \overset{d}{\sim} F$, where F is supported by the real line and has a differentiable probability density function, f, for x sufficiently large. If $\lim_{x\to\infty} f'(x)/f(x)$ exists (possibly infinite), then talent is identified by competition if and only if $\lim_{x\to\infty} f'(x)/f(x) = -\infty$.*

Corollary 1 implies that talent identification does not depend on the ability difference between talented and average competitors; nor does it depend on the variance of the noise term. To see the latter, suppose that the additive noise term is "scaled up," i.e., it becomes $\lambda\,\varepsilon$, $\lambda > 1$, in which case the variance of noise is scaled up by λ^2. Let G denote the CDF for $\lambda\,\varepsilon$. When $\varepsilon \overset{d}{\sim} F$, G satisfies that $G(x) = F(x/\lambda)$ for all x. Applying Corollary 1 then shows that in the noisier environment, talent is identified by competition if and only if $\lim_{x\to\infty}(1/\lambda)f'(x/\lambda)/f(x/\lambda) = -\infty$, a condition that is equivalent to $\lim_{x\to\infty} f'(x)/f(x) = -\infty$. Thus, the talent identification condition does not depend on the signal-to-noise ratio.

Because talent identification does not depend on the precision of a performance metric, measured by the signal-to-noise ratio, when the number of competitors is sufficiently large, performance metrics that satisfy the talent identification condition but have low precision are better at selecting elite talent compared to performance metrics that have high precision but do not satisfy the talent identification condition, even though the former may be remarkably inferior to the latter when the number of competitors is small.

For example, suppose that we increase the variance of the noise term from 1 to 2 in the Normal scenario of Example 1, while keeping the other assumptions. In this case, the Normal scenario has twice as large variance of noise as the Logistic scenario. Numerical computation shows that, when there are only two competitors, the probability that the best performer is average under the noisier, Normal scenario is about 0.31, much larger than the probability under the less noisy, Logistic scenario (≈ 0.23). However, when there are two thousand competitors, the probability that the best performer is average under the Normal scenario is only about half of the probability under the Logistic scenario (0.07 vs. 0.14).

Application 1. *What matters: the pond or the fish?* In practice, agents competing for status in the same environment do not always perform in the same contest. Frequently, there are many contest venues within a given environment and agents' status and economic rewards are based on the ranking of their venue as well as the ranking of their performance. Legal, economic, social, and information constraints frequently circumscribe venue choice and thereby push agents toward different venues.[16] There is no a priori reason to suspect that the proportion of talented and average competitors in these contests is the same in all of these venues. Some venues are likely to be "better" than others because they feature, on average, better competitors. In a highly competitive world, to what extent can superlative performance by a talented competitor competing in an inferior venue overcome the stigma attached to the venue?

Consider a simple example: agents are assigned to one of two venues—a patrician venue and a plebeian venue. Agents assigned to the patrician venue, whom we term "patricians," are more likely to be talented. Agents assigned to the plebeian venue, whom we term "plebeians," are less likely to be talented. However, a positive fraction of plebeians are talented and a positive fraction of patricians are average. Social status is determined by rational assessments of ability, i.e., the probability of being talented, conditioned on competitors' performance rank and the rank of their venue. We consider the question of whether competition can effect social mobility, i.e., whether it is possible for the assessed ability of any plebeian to outrank the assessed ability of any patrician. Given MLR ordering of performance distributions, competition can effect social mobility if and only if the assessed ability of the best-performing plebeian exceeds the assessed ability of the worst-performing patrician.

We consider this question in the limit, as competitiveness increases without bound, by comparing the assessed ability of the best-performing plebeian and the worst-performing patrician. The asymptotic assessed ability of the best-performing plebeian can be derived from Theorem 1. The asymptotic assessed ability of the worst-performing patrician can be found by simply noting that the worst performer has the highest *negative performance*, $-X$. Thus, by Theorem 1, fixing the proportion of talented competitors, the asymptotic probability that the competitor with the highest negative performance (i.e., the worst performer) is average is determined by the limit of the tail ratio between the distributions of negative performance. Because the tail ratio between the distributions of negative performance equals the *complementary* tail ratio between the performance distributions, F_T/F_A, we obtain the following result.

Corollary 2. *Let $\pi_A(n)$ be the probability that the worst performer is average when the scale-*

[16] Many authors discuss barriers to entry for the unconnected, and elite pathways for the connected, in a variety of contexts, e.g., elite education in the U.S. and France (Buisson-Fenet and Draelants, 2013; McDonough, 1997), catchments for elite public schools in the U.S. (Bergman et al., 2019), and desirable high-status careers in China (Jia and Li, 2017).

of-competition factor equals n. If $\lim_{x\to\ell} F_T(x)/F_A(x) = l$ *exists, then*

$$\lim_{n\to\infty} \pi_A(n) = \frac{a_o}{a_o + l\,t_o}.$$

Corollary 2 implies that, if the limit of the complementary tail ratio, l, equals 0, the asymptotic probability that the worst performer is average equals 1. In this case, competition *identifies mediocrity*. Whenever competition identifies talent and/or mediocrity, the assessed ability of the best performer at both venues will be greater than the assessed ability of the worst performer at either venue.

However, it is possible that competition identifies neither talent nor mediocrity. For example, when competitor performance equals ability plus i.i.d. noise and the noise distribution is symmetric, the talent and mediocrity identification conditions either both hold or both fail. By Theorem 1 and Corollary 2, when both conditions fail, asymptotically, the best-performing plebeian is more likely to be average than the worst-performing patrician if

$$\frac{a_o}{a_o + L t_o} > \frac{a_o^*}{a_o^* + l\,t_o^*}, \tag{6}$$

where the ratio between the number of talented and average competitors equals t_o/a_o at the plebeian venue and t_o^*/a_o^* at the patrician venue.

For example, suppose that the mean difference between talented and average competitors' performance equals one and that the noise distribution is Logistic with location parameter 0 and scale parameter 1. If 30% of patricians are talented (i.e., $t_o^*/a_o^* = 3/7$) and 5% of plebeians are talented (i.e., $t_o/a_o = 1/19$), condition (6) will hold. Moreover, because under Logistic noise, F_T MLR dominates F_A, at both venues, the best performer has the highest assessed ability and the worst performer the lowest. Thus, satisfaction of condition (6) implies that, even when the number of plebeians, and thus talented plebeians, is arbitrarily large, and exceptional performance is defined as topping the performance of all other plebeians, exceptional performance at the plebeian venue can never bridge the status gap between plebeians and patricians.[17]

Application 2. *Mutual fund managers: Luck or skill?* A significant body of research has developed around the question of whether some mutual fund managers are talented and thus can earn positive expected abnormal returns, "alpha" (e.g., Grinblatt and Titman, 1992; Cuthbertson et al., 2008). A common strategy for identifying talent is to estimate the abnormal performance of a sample of funds, identify "elite managers," the managers with the highest alphas (typically 95th or 99th percentile) and then test to see if the subsequent performance of these elite managers is better than average. Since there is little evidence supporting posit-

[17]In fact, in this example, within each venue, the assessed ability of the worst performer is decreasing in the number of competitors while the assessed ability of the best performer is increasing in the number of competitors. Thus, our conclusion holds even if each venue has an arbitrary number of competitors.

ive abnormal performance for the majority of fund managers, most research concentrates on comparing the following alternative hypotheses: (a) a small fraction of managers have talent versus (b) abnormal performance is simply the result of luck. The logic underlying these tests is that, if there are talented managers, the pool of elite managers will contain a sufficient fraction of such managers and therefore, under hypothesis (a), elite managers will generate significant out-of-sample positive abnormal returns.

Suppose that alpha is normally distributed, an assumption which per Theorem 1 favors talent identification through competition. Based on estimates in Cuthbertson et al. (2008), assume that the standard deviation of annual alpha is 0.25. Also assume that talented managers have an expected annual alpha of 3% and that average managers' expected annual alpha is zero. The 3% assumption for talented managers' expected annual alpha is a high estimate relative to the literature and thus favors talent identification because it postulates a very large (relative to mutual fund literature) difference between talented and average managers' expected performance. Because Normal random variables with the same variance and different means are MLR ordered, the probability that a manager is average conditioned on elite performance is no less than the probability that a manager is average conditioned on best performance. Thus, our estimates of the probability that a manager is average conditioned on best performance, Π_A, are lower bounds on the probability that a manager is average conditioned on elite performance. Assume that 5% of the managers are talented. In Table 1, using these assumptions and equation (4), we compute Π_A.

# Fund Managers	Π_A
1,000	0.9283
2,000	0.9267
100,000	0.9186
160,000	0.9177

Table 1: Mutual funds: Probability that the best performer is average, Π_A.

In the # Fund Managers column, the first two numbers encompass the number of funds included in representative empirical studies. 100,000 corresponds with the estimated number of mutual funds in the world (https://www.statista.com/topics/1441/mutual-funds). Even if we posit that the alphas of 160,000 managers are compared, the probability that the world's best performing mutual fund manager has only average ability is approximately 92%. Since 92% represents a lower bound on the probability that an elite manager has only average ability, if talent is persistent, the expected future alpha of elite managers is at most $0.08 \times 0.03 + 0.92 \times 0.00 = 0.0024$. Thus, for any imaginable number of competing fund managers, the out-of-sample performance of elite managers will be almost indistinguishable from the performance of average managers even if 5% of fund managers are extremely talented.

In this application, the fact that noise in portfolio alpha (standard deviation) is quite large relative to the expected performance advantage of talented managers implies that exceptional performance is primarily the product of luck. Because there are many more average managers than talented managers, an average manager is much more likely than a talented manager to be the most lucky manager. Although the signal-to-noise ratio and the fraction of average competitors do not matter for the talent identification condition (i.e., the condition in Theorem 1.i), they matter for how quickly Π_A approaches zero conditioned on the talent identification condition being satisfied. Because the Normal distribution satisfies the talent identification condition, ultimately, as the number of competing managers increases without bound, the best performing manager will have talent almost surely. However, for this application, because of the low signal-to-noise ratio and the high fraction of average managers, the number of competitors required to weed out untalented managers is implausibly large relative to practical constraints on the size of the competition.

2 Does increased competition favor luck or ability?

In the applications discussed thus far, although Theorem 1 showed that for many common distributional specifications, no level of competitiveness can eliminate the effect of luck, increasing competition nevertheless favored the talented, i.e., reduced the probability that the best performer is average. However, even if talented competitors' performance and average competitors' performance are MLR or hazard-rate ordered, competition need not favor the talented. The reason is that, while under MLR (or hazard rate) ordering, topping any fixed threshold performance level becomes relatively more difficult for an average competitor than for a talented competitor as the threshold increases, in rank-order competitions, the threshold for best performance is random and the distribution of this random threshold is different for average and talented competitors. Because of this *differential-threshold effect*, the extent to which the threshold is pushed up through the game-elevation effect of competition is not the same for the two types of competitors.

Our next result shows that the differential-threshold effect can dominate the game elevation effect and produce a negative total effect of competition on the expected ability of the best performer, and provides general conditions for competitions favoring or disfavoring talented competitors.[18]

Proposition 1. *For $x \in (\ell, h)$, let $r_T(x) = f_T(x)/F_T(x)$ and $r_A(x) = f_A(x)/F_A(x)$ represent the*

[18] In contrast to the talent identification result in Theorem 1, our characterization in Proposition 1, of the direction from which $\lim_{n\to\infty} \Pi_A$ is approached, does depend on our assumption of no aggregate uncertainty about the number of talented competitors (part (c) of Assumption 2.iii). This dependence results because, when the number of competitors is fairly small, aggregate uncertainty can greatly increase the probability that the best performer is average. Because we are primarily concerned with large competitions, we defer discussion of the effects of aggregate uncertainty to Appendix C.3.

reversed hazard rates of the talented and average competitors' performance distributions, F_T and F_A. Let n be the scale-of-competition factor.

(i) If $r_T(x)/r_A(x)$ is nondecreasing (increasing) in x, then the probability that the best performer is average, Π_A, is nonincreasing (decreasing) in n.

(ii) If $r_T(x)/r_A(x)$ is nonincreasing (decreasing) in x, then the probability that the best performer is average, Π_A, is nondecreasing (increasing) in n.

Proposition 1 shows that the monotonicity of the reversed hazard-rate ratio, $x \hookrightarrow r_T(x)/r_A(x)$, implies monotone convergence of Π_A to the limiting value provided by Theorem 1. The condition in Proposition 1.i of increasing reversed hazard-rate ratio is quite similar in appearance to the well-known MLR dominance condition, which requires the density ratio, f_T/f_A, to be increasing. However, MLR dominance and the reversed hazard-rate ratio being increasing are by no means equivalent conditions. In fact, Noe (2020) develops an order relation termed *geometric dominance*. Under our continuity and smoothness conditions, F_T (strictly) geometrically dominating F_A is equivalent to $x \hookrightarrow r_T(x)/r_A(x)$ being nondecreasing (increasing). Noe (2020) shows that geometric dominance neither implies nor is implied by MLR dominance.

When the reversed hazard-rate ratio is increasing, Proposition 1.i implies that increasing the scale of competition reduces the probability that the best performer is average. In this case, although increased competition might not identify talent, it does favor talented competitors. In fact, this case is quite commonly encountered: when both F_T and F_A are members of *the same* family of distributions, and differ only with respect to either scale or location, for most text-book distribution families (Normal, Gamma, Logistic, etc.), the reversed hazard-rate ratio will be increasing.[19] However, it is possible to construct location/scale families of distributions in which scale and location shifts that produce MLR dominance do not produce geometric dominance. Section 3.2 provides such an example.

Interestingly, as the next example shows, if F_T and F_A are members of *different* standard location/scale distribution families, even when F_T MLR dominates F_A, the reversed hazard-rate ratio can be decreasing, and thus, as Proposition 1.ii implies, increased competition favors the average. Because this result is admittedly a bit nonintuitive, we develop the logic behind it in the following example.

Example 2. Suppose that F_T is Max Extreme Value with location parameter 6 and scale parameter 1, and that F_A is Logistic with location parameter 5 and scale parameter 1. Then,

$$F_T(x) = \exp[-e^{-(x-6)}], \quad F_A(x) = (1+e^{-(x-5)})^{-1}, \quad x \in (-\infty, \infty). \tag{7}$$

[19] I.e., F_T, will geometrically dominate F_A. See Table 1 in Noe (2020). When the distribution family is not a location-scale family, cases exist where the compared distributions within a family are ordered by MLR but not by geometric dominance. For example, when both F_T and F_A are Beta, the β shape parameter (common to both distributions) is less than 1, and the α shape parameter is larger for F_T; or when both F_T and F_A are Kumaraswamy, the b shape parameter (common to both distributions) is less than 1, and the a shape parameter is larger for F_T.

Under these assumptions, the mean performance by talented competitors equals approximately 6.58; mean performance by average competitors equals 5. As we verify in Section D.1, F_T MLR dominates F_A. The reversed hazard-rate ratio is given by

$$\frac{r_T(x)}{r_A(x)} = \frac{e^{-(x-6)}}{(1+e^{x-5})^{-1}} = e + e^{6-x},$$

which is clearly decreasing in x. Thus, Proposition 1.ii implies that increasing competition will increase the probability that the best performer is average, Π_A. Assuming $a_o = t_o = 1$, we illustrate in Panel B this effect by varying the scale-of-competition factor, n, between 1 and 60.

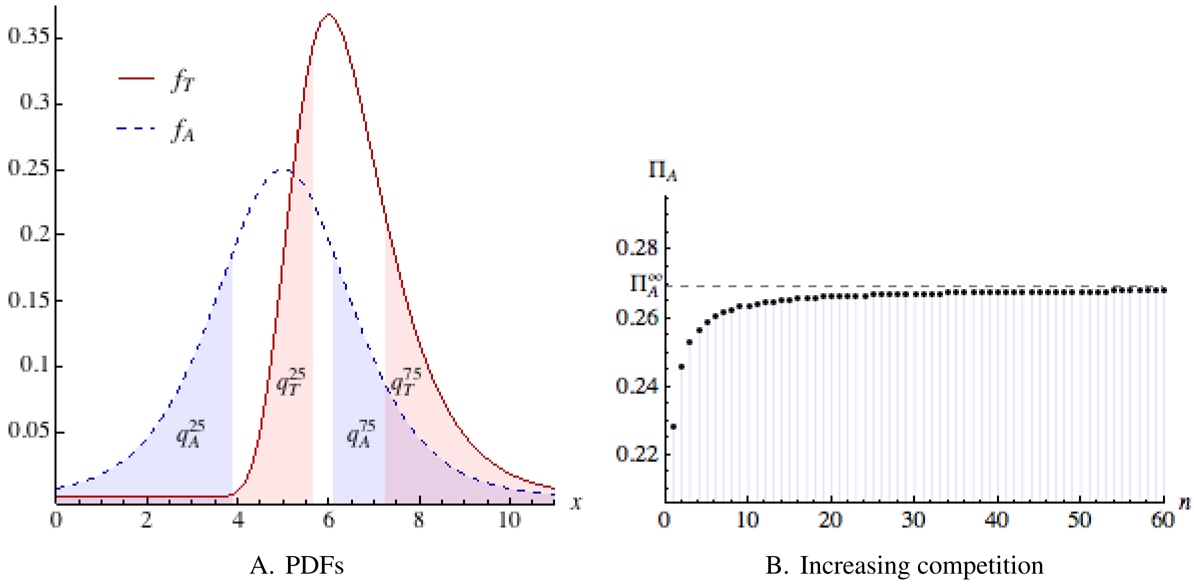


Figure 2: *When increased competition favors the average.* This figure depicts a scenario in which there are an equal number of talented and average competitors, $a_o = t_o = 1$, and the talented and average competitors' performances are distributed as specified in equation (7). In Panel A, the regions shaded in red (blue) represent probability mass in the first, q^{25}, and fourth, q^{75}, quartiles of F_T (F_A). In Panel B, the horizontal axis represents the scale of competition. The vertical axis represents the probability that the best performer is average, Π_A. The limiting value of Π_A is represented by Π_A^∞.

In this example, the increased competition favors average competitors. The intuition is perhaps easiest to understand by considering the relationship between the quantiles of performance for the two types of competitors. As Panel A of the figure reveals, the relative advantage of talented competitors over average competitors is greater in the lowest quantiles. For example, if we condition on performance being drawn from the lowest quartiles of F_T and F_A, the expected performance of talented competitors is about 90% greater than the expected performance of average competitors (5.20 vs. 2.75). In contrast, conditioned on the top quartile, the expected performance of talented competitors is only about 15% greater (8.32 vs. 7.25).

Increasing the number of competitors reduces the likelihood that the best performance will be drawn from low quantiles and hence reduces the advantage of talented competitors. Thus, the probability that an average competitor will be the best performer, Π_A, increases with the number of competitors. However, as the number of competitors increases, the probability that competition for best performance will be fought over the lower quantiles of the distributions becomes smaller and smaller and thus the differential-threshold effect peters out. In the limit, Π_A converges to the value specified in Theorem 1.

3 Racing and target-hitting contests

In this section, we consider two classes of competitions, racing contests and target-hitting contests. As we demonstrate below, in such competitions, talent can never be identified and increasing the scale of competition, in many cases, does not even uniformly reduce the probability of the best performer being average.

3.1 Racing contests

In *racing contests*, e.g., patent races, hacker competition, problem-solving contests, competitors compete in the speed of completing a task. The quality standard for completion is fixed while time required for completion is random but correlated with ability. Performance is measured by time to completion. The competitor who completes the task first "wins the race," i.e., is the best performer. A formal definition of the racing contests we consider is given below:

Definition 1. A racing contest is a contest in which performance, X, is given by

$$X_T = -\tau_T, \quad X_A = -\tau_A,$$

where

(i) $\tau_T \stackrel{\text{dist}}{=} \tau/\lambda_T$, $\tau_A \stackrel{\text{dist}}{=} \tau/\lambda_A$, and $\lambda_T > \lambda_A > 0$;

(ii) $\tau \stackrel{d}{\sim} F$, where

- a. F is supported by the nonnegative real line, and
- b. F varies regularly at 0, i.e., there exists $\nu > 0$ such that for any $a > 0$, $\lim_{x \to 0^+} F(ax)/F(x) = a^\nu$, where ν is termed the *index of variation* of F.

This definition encompasses the distributions typically used to model innovation adoption and used in theoretical models of R&D races, as well as empirical studies of innovation, e.g., Exponential, Weibull, Erlang, Gamma distributions. Condition i ensures that the talented and average competitors' performance distributions differ only with respect to a scaling factor, τ_j, $j = T, A$, applied to a common *racing distribution*, F, and that, on average, talented competitors complete faster than average competitors. Condition ii.a is quite natural given that the metric for performance is time. The regular variation condition, condition ii.b, is a standard condition in probability and mathematics (§1.4.2 Bingham et al., 1989). Regular variation of F at 0 implies that the racing distribution, F, behaves like a power function in sufficiently small neighborhoods of 0. A simple sufficient condition for F to be regularly varying at 0 with index of variation $\nu > 0$ is $F(x) \sim \kappa x^{\nu}$, as $x \to 0^{+}$, where $\kappa > 0$.[20] If the density of F satisfies $f(0) > 0$ and $\lim_{x\to 0^{+}} f(x) = f(0)$, then $F(x) \sim \kappa x$ as $x \to 0^{+}$, where $\kappa = f(0) > 0$, and thus $\nu = 1$. If F is analytic in the neighborhood of 0, then ν equals the index of the first non-vanishing derivative of F's power series expansion.

Our next result shows that talent is never identified in racing contests and that the index of variation, ν, and the ratio of expected completion speeds, λ_T/λ_A, provide a complete characterization of the asymptotic probability that the race winner is average.

Proposition 2. *In any racing contest,*

(i) talent is never identified and the asymptotic probability that the race winner is average is given by $\lim_{n\to\infty} \Pi_A(n) = a_o/(a_o + Lt_o)$, *where* $L = (\lambda_T/\lambda_A)^{\nu}$.

(ii) If the hazard rate function of the racing distribution, $h = f/(1-F)$, *is (strictly) geometrically convex, then* $n \hookrightarrow \Pi_A(n)$ *is nondecreasing (increasing).* [21]

(iii) If the hazard rate function of the racing distribution, $h = f/(1-F)$, *is (strictly) geometrically concave, then* $n \hookrightarrow \Pi_A(n)$ *is nonincreasing (decreasing).* [22]

Proposition 2 has a number of immediate implications. Part (i) shows that the only relevant characteristic for the shape of the racing distribution is its index of variation, ν. For many commonly employed racing distributions (e.g, Weibull, Gamma, Erlang), ν equals the shape parameter of the distribution. For other distributions used to model time to completion (e.g., Exponential, Gompertz), $\nu = 1$. In the patent-race models pioneered by Loury (1979) and Dasgupta and Stiglitz (1980), the discovery time of an innovation is Exponential. Proposition 2 implies that, in these patent-race models, talent is never identified and the asymptotic probability that the race winner is average will be the same as under any distribution F that has $\nu = 1$.

The most interesting implication of parts (ii) and (iii) follows from noting that the hazard rate of the Weibull distribution, the most commonly used statistical model for the length of time between events, is log-linear and thus geometrically linear. Thus, Proposition 2 implies

[20] $F(x) \sim \kappa x^{\nu}$, as $x \to 0^{+}$ means that F is asymptotically equivalent to a power function, i.e. $\lim_{x\to 0^{+}} F(x)/(\kappa x^{\nu}) = 1$.

[21] A continuous, differentiable function, $\phi : (0,+\infty) \to (0,+\infty)$, is (strictly) geometrically convex if $x \hookrightarrow x\phi'(x)/\phi(x)$ is nondecreasing (increasing); this condition is equivalent to $\log g(x)$ being a (strictly) convex function of $\log x$.

[22] A continuous, differentiable function, $\phi : (0,+\infty) \to (0,+\infty)$, is (strictly) geometrically concave if $x \hookrightarrow x\phi'(x)/\phi(x)$ is nonincreasing (decreasing); this condition is equivalent to $\log g(x)$ being a (strictly) concave function of $\log x$.

that, when F is Weibull, increased competition has *no* effect on the expected ability of the race winner.

Application 3. *Need for speed vs. need for talent: High stakes vs. low stakes.* High-stakes racing contests, where many competitors compete and their speed of completing a single task is decisive for winning, are widely used to speed up innovation, e.g., crowdsourcing contests. In these contests, quite often, the contest designer only cares about the highest speed but not the ability of the race winner. However, there also exist real-world situations in which the designer cares about the ability of the race winner but not the speed per se. Such situations are most common when the designer uses racing contests for recruitment. In what follows, we show that low-stakes racing contests, where few competitors compete and no single task is decisive for winning, is more likely to produce talented race winners than high-stakes racing contests.

Consider an m-stage racing contest, $m \geq 1$, where in each stage, competitors compete in the speed of finishing a given task. The race winner is the competitor who produces the highest total speed over the m tasks. For each task, the time to completion is Exponential with rate parameter λ_t for a type-t competitor, $t = T, A$. Because the sum of m i.i.d. Exponential random variables is Gamma with shape parameter m, the racing distribution, F, for the m-stage racing contest is Gamma with shape parameter m.

High-stakes racing contests correspond to the case of $m = 1$ (i.e., one-stage performance is decisive) and many competitors competing. When $m = 1$, the racing distribution, F, is Exponential (or equivalently, Gamma with shape parameter 1). In this case, by Proposition 2, increasing the scale of competition has no effect on the expected ability of the race winner and the race winner is average with probability $a_o/(a_o + (\lambda_T/\lambda_A)t_o)$. For example, suppose that half of the competitors are talented (i.e., $a_o = t_o$) and talented competitors' speed is in expectation twice as fast as average competitors' speed (i.e., $\lambda_T/\lambda_A = 2$). Then, under a high-stakes racing contest, $\Pi_A = 1/3$.

Now consider a low-stakes racing contest with $m = 2$ and with only two competitors, one talented and one average. Keep the assumption of $\lambda_T/\lambda_A = 2$. Numerical integration, using equation (4), shows that, in this case, $\Pi_A \approx 0.259$. Thus, even with only two competitors, a low-stakes contest produces a winner with significantly higher expected ability than high-stakes contests with any scale of competition.

3.2 Target-hitting contests

In *target-hitting contests*, e.g., shooting sports and forecasting competitions, competitors' performance is measured by the accuracy of their "shot" at a "target." The best performer is the competitor whose shot is closest to the target, e.g., the best macroeconomic forecaster is the forecaster whose prediction of GDP is closest to actual GDP.

It is natural to assume that the error made by competitors in a target-hitting task (i.e., the difference between a competitor's "shot" and the target) has a zero-mean, symmetric, unimodal distribution, and that the smaller the absolute value of the error, the better the competitor's performance. We thus define target-hitting contests as follows:

Definition 2. A target-hitting contest is a contest in which performance, X, is given by

$$X_T = -g(|\varepsilon_T|), \quad X_A = -g(|\varepsilon_A|), \tag{8}$$

where

(i) $\varepsilon_T \overset{\text{dist}}{=} \varepsilon/\lambda_T$, $\varepsilon_A \overset{\text{dist}}{=} \varepsilon/\lambda_A$, and $\lambda_T > \lambda_A > 0$.

(ii) $\varepsilon \overset{d}{\sim} F$, where the *target-hitting distribution*, F, is symmetric about 0, unimodal, supported by the real line, and has a continuous probability density function, f.

(iii) $g : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing, continuous, $g(0) = 0$, and $\lim_{x\to\infty} g(x) = \infty$.

The target-hitting distribution in our definition captures all text-book distributions that are symmetric and unimodal. These distributions include not only Normal distributions, which have been frequently adopted to model measurement and prediction error, but also other regular distributions (e.g., Logistic, Laplace, Hyperbolic Secant) that have log-concave densities and moments of all orders, as well as more irregular ones (e.g., Cauchy, Student's-t, Slash) that do not have log-concave densities or moments of all orders. According to our definition, λ_T/λ_A equals the ratio between precisions of the error distribution (whenever exists) of talented and average competitors. Thus, λ_T/λ_A measures the accuracy asymmetry between talented and average competitors. The next proposition characterizes the probability that the best performer is average in target-hitting contests.

Proposition 3. *In any target-hitting contest,*

(i) talent is never identified and the asymptotic probability that the best performer is average equals $\lim_{n\to\infty} \Pi_A(n) = \frac{a_o}{a_o+Lt_o}$, *where* $L = \lambda_T/\lambda_A$.

(ii) If the hazard rate function of the target-hitting distribution, $h = f/(1-F)$, *is (strictly) geometrically convex over* $(0,\infty)$ *(e.g., F is Normal), then* $n \hookrightarrow \Pi_A(n)$ *is nondecreasing (increasing).*

(iii) If the hazard rate function of the target-hitting distribution, $h = f/(1-F)$, *is (strictly) geometrically concave over* $(0,\infty)$*, then* $n \hookrightarrow \Pi_A(n)$ *is nonincreasing (decreasing).*

Proposition 3 has two interesting implications. First, by part (i), talent is never identified in target-hitting contests and the asymptotic probability that the best performer is average depends only on accuracy asymmetry, λ_T/λ_A, but not on the shape of the target-hitting distribution, F. For example, this probability is the same under Cauchy and Normal laws despite the vast differences in the properties of these two laws. Part (i) also imposes rather tight limits on the power of competition for selecting talented competitors. For example, if half of the competitors are talented, to ensure that 95% of the time, the competitor submitting the most accurate shot is talented in the limit, the accuracy asymmetry between the competitors, λ_T/λ_A, must be at least 19:1. Thus, in large target-hitting contests, whenever accuracy asymmetry is not extreme, the competitor making the most accurate shot will frequently be average.[23]

Second, part (ii) provides a sufficient condition under which increased competition always favors average competitors and shows that Normally distributed errors satisfy this condition. Thus, under the most commonly adopted error law, the Normal error law, increased competition increases the probability that the competitor making the most accurate shot is average. Consequently, under the Normal error law, best performance in a target-hitting (e.g., forecasting) contest featuring many competitors is less informative about the best-performer's ability than best performance in a contest with few competitors.

Part (iii) shows a sufficient condition, the geometric concavity of the hazard rate function of F, under which increased competition favors talented competitors. However, this condition can only be satisfied by error laws that also satisfy a very standard regularity condition for error laws—having a log-concave distribution—if the hazard rate function of the error law is geometrically linear (see Appendix D.2). When the hazard rate function is geometrically linear, increased competition has no effect on the expected ability of the best performer. Thus, in target-hitting competitions, increased competition favoring talented competitors is atypical.

4 Extensions

4.1 Endogenous contest participation

In Section 2, we studied whether increased competition favors luck or talent. Our analysis was based on the assumption that increased competition does not change the composition of the contestant pool, i.e., we fixed the ratio between the number of talented and average competitors. In what follows, we allow for the possibility that the composition of the contestant pool varies with the degree of competition by endogenizing contest participation. We restrict attention to winner-take-all contests (with an entry cost), in which only the best performer receives a prize. We consider situations in which competition is increased by *globalizing* identical winner-take-all contests, i.e., merging n winner-take-all contests, each having a prize of value v into a single

[23]The limited effect of competition on the expected ability of the best performer in target-hitting contests is not caused by the target-hitting specification (Definition 2) failing to engender MLR or hazard-rate dominance relations between talented and average performance. The necessary and sufficient condition for X_T to MLR dominate X_A in a target-hitting contest is that the PDF for the target-hitting distribution, F, is geometrically concave over $(0, \infty)$. This condition is satisfied by all error distributions with log-concave densities (e.g., Normal, Logistic, Laplace, Hyperbolic Secant) and also by many other error distributions without log-concave densities (e.g., Cauchy, Student's-t, Slash).

winner-take-all contest, having a prize of value nv. These situations, as will be elaborated in more detail in our application below, resemble the change of the reward structure for artists caused by globalization.

We aim to address the question of how globalization affects the composition of the contestant pool. To answer this question, it is useful to first analyze how globalization affects talented and average competitors' expected prize when the composition of the contestant pool is fixed. To this end, we define *replication* of a contest following Wärneryd (2001):

Definition 3. Let $\mathscr{C}$ be a contest with t_o talented competitors, a_o average competitors, and a prize of value v. The *n-fold replication* of $\mathscr{C}$ is the contest $\mathscr{C}_n$ with nt_o talented competitors, na_o average competitors, and a prize of value nv.

The next result is straightforward from Proposition 1.

Corollary 3. *Let $\mathscr{C}_n$ be an n-fold replication of contest $\mathscr{C}$, $n \geq 2$. For $x \in (\ell, h)$, define $r_T(x)$ and $r_A(x)$ as in Proposition 1.*

(i) If $r_T(x)/r_A(x)$ is nondecreasing (increasing) in x, then the expected prize received by an average competitor is (strictly) lower in $\mathscr{C}_n$ than in $\mathscr{C}$.

(ii) If $r_T(x)/r_A(x)$ is nonincreasing (decreasing) in x, then the expected prize received by an average competitor is (strictly) higher in $\mathscr{C}_n$ than in $\mathscr{C}$.

Corollary 3 shows that an n-fold replication of a contest reduces (increases) the expected prize received by an average competitor if the reversed hazard-rate ratio, $r_T(x)/r_A(x)$, is increasing (decreasing). Intuitively, replication increases competitiveness without changing the composition of the contestant pool or per capita reward. As shown by Proposition 1, increased competitiveness favors talented (average) competitors if $r_T(x)/r_A(x)$ is increasing (decreasing). Thus, replication makes average competitors worse (better) off if $r_T(x)/r_A(x)$ is increasing (decreasing).

Application 4. *Competition for superstardom: How does globalization affect the supply of untalented artists?* A prominent feature of the world of arts is the winner-take-all compensation structure for artists. A handful of superstars reap a disproportionate share of economic rewards while the vast majority earn very little (Menger, 1999). Quite often, talent is not the deciding factor for winning the competition for superstardom; luck matters (Adler, 1985; Ginsburgh and Van Ours, 2003). The high degree of uncertainty in prize assignment, combined with the lottery type of reward structure, has allured many individuals, even untalented ones, to give up regular jobs to become an artist (Frank and Sohn, 2011), causing a misallocation of labor (Frank and Cook, 2010).

Globalization, by replacing local stars by global superstars, intensifies the competition (Adler, 2006). Does globalization also exacerbate the misallocation of labor? Some economists and sociologists argue that the misallocation of labor is a consequence of untalented artists' tendency to overestimate their chance of winning (Frank and Cook, 2010; Frank and Sohn, 2011). Given the fact that globalization tremendously increases the top performers' prize, this behavioral argument would predict globalization as a stimulus of misallocation of labor.

In this application, we use a rational-choice model to show that globalization exacerbates the misallocation of labor under certain conditions. When these conditions fail, it is possible that globalization alleviates the misallocation of labor and we provide a sufficient condition under which globalization drives out untalented artists.

Consider a world consisting of $n \geq 2$ regions. Each region has $t_o > 0$ talented agents and an infinite supply of average agents. All agents are risk neutral. Each agent has two choices in life: taking a regular job, or becoming an artist and participating in a winner-take-all art contest. All agents are equally able to take a regular job. The talented agents are better artists, in the sense that the distribution of a talented agent's art performance, F_T, MLR dominates the distribution of an average agent's art performance, F_A. Assume that F_T and F_A satisfy the regularity conditions in Assumption 1. Becoming an artist requires an agent to forego the wage for a regular job, $w > 0$. The only reward for an artist is the prize for an art contest, which is assigned only to the best performer of the contest.

Before globalization, each region has its own local winner-take-all art contest, which can only be attended by artists in the region and rewards the best local art performer with a prize of value v. We assume that v is sufficiently large such that, even when all t_o local talented agents choose to be an artist, it is still profitable for some local average agents to be an artist.

After globalization, the n local art contests are merged into a single global art contest. This global contest has only one prize, assigned to the best global art performer. We abstract from the possible effect of globalization on the total rewards to artists by assuming that the value of the prize of the global contest equals $n\,v$, the sum of the rewards provided by all local contests before globalization. Moreover, we assume that globalization does not change the wage for a regular job, w.

We ignore the integer problem by allowing the number of competitors to be any nonnegative real number. We focus on pure-strategy equilibria and adopt Pareto dominance for equilibrium refinement. Because of the infinite supply of average agents, in any equilibrium, the expected payoff to an average agent always equals w. In contrast, given that F_T MLR dominates F_A, implying that a talented artist's probability of winning is higher than an average artist's, in any equilibrium in which some average agents become an artist, a talented artist's expected payoff must be greater than w, the wage the talented agent would have received had she chosen not to be an artist. Thus, a Pareto-dominant equilibrium must have no average artists unless all talented agents become an artist.

The next result shows the effect of globalization on participation in artistic markets.

Result 1. For $x \in (\ell, h)$, define $r_T(x)$ and $r_A(x)$ as in Proposition 1. The following results hold:

(i) Globalization does not affect the total number of talented artists.

(ii) In contrast,

  (a) if $r_T(x)/r_A(x)$ is decreasing in x, globalization increases the total number of average artists.

  (b) If $r_T(x)/r_A(x)$ is increasing in x, globalization reduces the total number of average artists. In this case, the fraction of average artists converges to 0 as the number of local contests being merged, n, tends to infinity if and only if

$$L \geq v/(w t_o),$$

  where $L = \lim_{x \to h} (1 - F_T(x))/(1 - F_A(x))$ (possibly infinite).

As shown by Corollary 3, when the reversed hazard-rate ratio, $r_T(x)/r_A(x)$, is decreasing, with agents' participation decision fixed, globalization increases an average competitor's expected payoff. Part ii.a of Result 1 shows that, in this case, when participation is endogenous, globalization increases the supply of untalented artists, exacerbating the misallocation of labor. In contrast, part ii.b shows that, if $r_T(x)/r_A(x)$ is increasing, globalization reduces the supply of untalented artists. In the limit, i.e., when the number of local contests being merged tends to infinity, globalization drives out untalented artists, in the sense that the fraction of untalented artists converges to 0, if the limit of the tail ratio, L, is sufficiently large.

4.2 Global rank dominance

The scale-of-competition effect presented in Proposition 1 can be applied to *best-shot contests*, contests where competitors are ranked by their most successful attempt out of a fixed number of attempts (e.g., Olympic weightlifting), to study whether allowing competitors to take more "shots" favors luck or talent. We denote by F_T (F_A) the distribution of a talented (average) competitor's performance from a single shot. Note that the probability that the best performer is average in a best-shot contest with a_o average competitors and t_o talented competitors, each making n independent attempts, equals the probability that the best performer is average in a single-shot contest with $n a_o$ average competitors and $n t_o$ talented competitors. Thus, the effect of increasing the number of shots on the expected ability of the best performer in a best-shot contest is equivalent to the scale-of-competition effect in a single-shot contest. Thus, by Proposition 1, if the reversed hazard-rate ratio, $r_T(x)/r_A(x)$, is increasing (decreasing), increasing the number of shots in a best-shot contest will reduce (increase) the probability that the best performer is average.

In fact, as the next result shows, if the reversed hazard-rate ratio, $r_T(x)/r_A(x)$, is increasing (decreasing), increasing the number of shots in a best-shot contest not only reduces (increases) an average competitor's probability of attaining the highest rank but, more generally, lowers (improves) an average competitor's rank in the sense of first-order stochastic dominance.

Proposition 4. *For $x \in (\ell, h)$, let $r_T(x) = f_T(x)/F_T(x)$ and $r_A(x) = f_A(x)/F_A(x)$ represent the reversed hazard rates of the talented and average competitors' (one-shot) performance distributions, F_T and F_A. In a best-shot contest with K shots for each competitor, competitors are ranked in ascending order based on their highest draw out of K independent draws (from F_T for talented competitors and from F_A for average competitors). The following results hold for any $K'' > K' \geq 1$:*

(i) if $r_T(x)/r_A(x)$ is nondecreasing (increasing) in x, then an average competitor's performance rank in a best-shot contest with K'' shots for each competitor is (strictly) lower than her performance rank in a best-shot contest with K' shots for each competitor in the sense of first-order stochastic dominance.

(ii) If $r_T(x)/r_A(x)$ is nonincreasing (decreasing) in x, then an average competitor's performance rank in a best-shot contest with K'' shots for each competitor is (strictly) higher than her performance rank in a best-shot contest with K' shots for each competitor in the sense of first-order stochastic dominance.

Proposition 4 implies that, if the reversed hazard-rate ratio, $r_T(x)/r_A(x)$, for the one-shot performance distributions is increasing (decreasing), allowing competitors to take more "shots" in a best-shot contest will make performance a less (more) noisy signal of talent.

Application 5. *Education reform in China: Furthering meritocracy?* The annually held China's National College Entrance Examination, known as *Gaokao*, serves the function of allocating high-school graduates to China's highly stratified higher-education system based on students' Gaokao scores. Because of the high-stakes nature of Gaokao, students undergo tremendous pressure when preparing for and taking Gaokao (Altback, 2016). To alleviate students' stress, China initiated a set of Gaokao reforms in 2014. One reform is to allow students to take the test of English language, a compulsory Gaokao subject, *twice* and use the higher score for consideration in college admissions (Gu and Magaziner, 2016).

Obviously, this reform tends to increase a student's Gaokao score. However, because college admissions are based on relative performance, what is not obvious is whether this reform furthers meritocracy by better aligning admission outcomes with students' ability. Proposition 4 implies that the answer depends on the behavior of the ratio between the reversed hazard rates of the talented and average students' (one-shot) score distributions, $r_T(x)/r_A(x)$. If this ratio is increasing in x, then the Gaokao reform will disfavor average students, lowering their Gaokao rank and, consequently, making Gaokao rank a better signal of ability. When this happens, the Gaokao reform, which intends to ease students' Gaokao stress, also makes the Gaokao system more meritocratic. In contrast, if the ratio, $r_T(x)/r_A(x)$, is decreasing in x, the reform will favor average students, in which case the stress-easing benefit provided by the reform will come at a cost of less meritocratic selection.

5 Conclusion

Few would doubt that luck (i.e., exogenous noise that garbles the relationship between ability and performance) plays a non-negligible role in determining an individual's success (Frank, 2016). What is less clear is whether the role of luck will be negligible when success is defined as being the best performer and when the contest environment is sufficiently competitive. By deriving a necessary and sufficient condition under which competition identifies talent (i.e., as the number of competitors tends to infinity, the probability that the best performer is average tends to zero), we showed that, in a wide range of settings, such as forecasting competitions and innovation races, talent cannot be identified by competition. In fact, increasing competition can sometimes favor luck rather than talent because of a subtle, differential-threshold effect.

These results have many implications, more than we could explore explicitly through the applications we developed. For example, in many economic environments, losers imitate winners: unsuccessful researchers imitate successful researchers (Heesen, 2017), firms imitate their most profitable rivals (Vega-Redondo, 1997), investors imitate the most successful investors (Banerjee, 1992; Bernardo and Welch, 2001; Peltz, 2001). When the imitated elite are following superior strategies, imitation can lead to convergence to socially desirable outcomes (Vega-Redondo, 1997). In contrast, when the imitated elite are merely lucky, imitation can be destabilizing (Bikhchandani and Sharma, 2000). Thus, in highly competitive markets characterized by imitation, the talent identification condition provides insight into whether imitation will lead to socially desirable outcomes.

The effect of high-stakes competitions on the relationship between luck, talent, and elite status is also likely to have substantial effects on public policy. Based on survey data, Fong (2001) reports that support for redistributive policies is significantly higher when respondents believe that inequality is the result of luck rather than unequal ability. Philosophers also distinguish between differential outcomes produced by luck and ability (Nagel, 1979).[24] In some economic models, the optimal level of wealth redistribution is positively related to the degree to which wealth differentials are caused by luck (e.g., Varian, 1980). Thus, it is quite plausible to assert that intellectual and popular support for redistribution policies will be positively related to the degree to which success depends on luck. Given this assertion, our model—which shows that the outcomes of high-stakes contests, even in highly competitive environments, are frequently the product of luck—predicts that allocations of status and rewards through high-

[24] In this literature, what we call "luck" is called "resultant luck;" what we call "ability" is called "constitutional luck."

stakes competitions will increase demands for wealth redistribution.

Finally, in Section 1, we showed that, even when the effects of luck are moderate, and produced by standard textbook statistical distributions, performance, no mater how extraordinary, in an inferior but highly competitive venue (e.g., large comprehensives or charter schools) may never produce as much status as performance, no matter how undistinguished, at a superior venue (e.g., elite universities, wealthy school catchments). Thus, our analysis suggests that, for implementing social mobility, competition and high standards at inferior venues cannot generally substitute for increased access to superior venues.

References

Adler, Moshe, "Stardom and talent," *American Economic Review*, 1985, *75* (1), 208–212.

_ , "Stardom and talent," *Handbook of the Economics of Art and Culture*, 2006, *1*, 895–906.

Altback, Philip G, "China's glass ceiling and feet of clay," `https://www.universityworldnews.com/post.php?story=20160217143711361` 2016. Accessed: 2019-08-09.

Ash, Robert, *Real Analysis and Probability*, Academic Press, 1972.

Banerjee, Abhijit V, "A simple model of herd behavior," *Quarterly Journal of Economics*, 1992, *107* (3), 797–817.

Baricz, Árpád, "Geometrically concave univariate distributions," *Journal of Mathematical Analysis and Applications*, 2010, *363* (1), 182–196.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F Katz, and Christopher Palmer, "Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice," Technical Report, National Bureau of Economic Research 2019.

Bernardo, Antonio E and Ivo Welch, "On the evolution of overconfidence and entrepreneurs," *Journal of Economics & Management Strategy*, 2001, *10* (3), 301–330.

Bikhchandani, Sushil and Sunil Sharma, "Herd behavior in financial markets," *IMF Staff Papers*, 2000, *47* (3), 279–310.

Bingham, Nicholas H, Charles M Goldie, and Jef L Teugels, *Regular Variation*, Cambridge University Press, 1989.

Birkel, Thomas, "A note on the strong law of large numbers for positively dependent random variables," *Statistics & Probability Letters*, 1988, *7* (1), 17–20.

Bol, Thijs, Mathijs de Vaan, and Arnout van de Rijt, "The Matthew effect in science funding," *Proceedings of the National Academy of Sciences*, 2018, *115* (19), 4887–4890.

Buisson-Fenet, Hélène and Hugues Draelants, "School-linking processes: describing and explaining their role in the social closure of French elite education," *Higher Education*, 2013, *66* (1), 39–57.

Clark, Derek J. and Christian Riis, "Rank-Order Tournaments and Selection," *Journal of Economics*, 2001, *73* (2), 167–191.

Clark, Derek J and Christian Riis, "Contingent payments in selection contests," *Review of Economic Design*, 2007, *11* (2), 125–137.

Cuthbertson, Keith, Dirk Nitzsche, and Niall O'Sullivan, "UK mutual fund performance: Skill or luck?," *Journal of Empirical Finance*, 2008, *15* (4), 613–634.

Dasgupta, Partha and Joseph Stiglitz, "Uncertainty, industrial structure, and the speed of R&D," *Bell Journal of Economics*, 1980, *11* (1), 1–28.

Denrell, Jerker, "" Experts" Who Beat the Odds Are Probably Just Lucky," *Harvard Business Review*, 2013, *91* (4), 28–29.

_ and Chengwei Liu, "Top performers are not the most impressive when extreme performance indicates unreliability," *Proceedings of the National Academy of Sciences*, 2012, *109* (24), 9331–9336.

_ , _ , and Gaël Le Mens, "When more selection is worse," *Strategy Science*, 2017, *2* (1), 39–63.

Ding, Bill and Russ Wermers, "Mutual fund "stars": The performance and behavior of US fund managers," 2004. University of Maryland Working Paper.

Drugov, Mikhail and Dmitry Ryvkin, "The shape of luck and competition in tournaments," 2019. Working paper, U. of Florida.

Fang, Dawei and Thomas H Noe, "Less competition, more meritocracy?," SSRN Working Paper 2019.

Fong, Christina, "Social preferences, self-interest, and the demand for redistribution," *Journal of Public Economics*, 2001, *82* (2), 225–246.

Franck, Egon and Stephan Nüesch, "Talent and/or popularity: What does it take to be a superstar?," *Economic Inquiry*, 2012, *50* (1), 202–216.

Frank, Joshua and Saeyoon Sohn, "A behavioral economic analysis of excess entry in arts labor markets," *The Journal of Socio-Economics*, 2011, *40* (3), 265–273.

Frank, Robert H, *Success and Luck: Good Fortune and the Myth of Meritocracy*, Princeton University Press, 2016.

_ **and Philip J Cook**, *The Winner-Take-All Society: Why the Few at the Top Get So Much More than the Rest of Us*, Random House, 2010.

Ginsburgh, Victor A and Jan C Van Ours, "Expert opinion and compensation: Evidence from a musical competition," *American Economic Review*, 2003, *93* (1), 289–296.

Grinblatt, Mark and Sheridan Titman, "The persistence of mutual fund performance," *Journal of Finance*, 1992, *47* (5), 1977–1984.

Gu, Mini and Jessica Magaziner, "The *Gaokao*: History, reform, and rising international significance of China's National College Entrance Examination," `https://wenr.wes.org/2016/05/the-gaokao\-history-reform\-and-international-significance\-of-chinas-national-college-entrance-examination` 2016. Accessed: 2019-08-09.

Heesen, Remco, "Academic superstars: competent or lucky?," *Synthese*, 2017, *194* (11), 4499–4518.

Hizen, Yoichi and Ryo Okui, "Olympic athlete selection," *The BE Journal of Economic Analysis & Policy*, 2009, *9* (1).

Hoffman, Greg, "Paul the Octopus proves Buffett was right," *Sydney Morning Herald*, 2010, *14*.

Hoffman, Mitchell, Lisa B Kahn, and Danielle Li, "Discretion in hiring," *The Quarterly Journal of Economics*, 2018, *133* (2), 765–800.

Hvide, Hans K. and Eirik G. Kristiansen, "Risk taking in selection contests," *Games and Economic Behavior*, 2003, *42* (1), 172–179.

Jelinek, Mariann and Nancy J Adler, "Women: World-class managers for global competition," *Academy of Management Perspectives*, 1988, *2* (1), 11–19.

Jia, Ruixue and Hongbin Li, "The value of elite education in China," 2017. 21st Century China Center Research Paper 2017-07, available on SSRN.

Katz, Michael L and Carl Shapiro, "Systems competition and network effects," *Journal of Economic Perspectives*, 1994, *8* (2), 93–115.

Kawamura, Kohei and Inés Moreno de Barreda, "Biasing selection contests with ex-ane identical agents," *Economics Letters*, 2014, *123* (2), 240–243.

Ledford, Sarah H., Minda M. Monteagudo, Alejandro N. Flores, Jennifer B. Glass, and Kim M. Cobb, "#GeoGRExit: Why Geosciences Programs Are Dropping the GRE," *Eos*, 2020, *101.*

Leone, Andrew J and Joanna S Wu, "What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts," 2007. Simon School of Business Working Paper.

Loury, Glenn C, "Market structure and innovation," *Quarterly Journal of Economics*, 1979, *93* (3), 395–410.

Matuła, Przemysław, "A note on the almost sure convergence of sums of negatively dependent random variables," *Statistics & Probability Letters*, 1992, *15* (3), 209–213.

McDonough, Patricia M, *Choosing Colleges: How Social Class and Schools Structure Opportunity*, SUNY Press, 1997.

Menger, Pierre-Michel, "Artistic labor markets and careers," *Annual Review of Sociology*, 1999, *25* (1), 541–574.

Migheli, Matteo, "Competing for promotion: Are "the best" always the best?," *Research in Economics*, 2019, *73* (2), 149–161.

Münster, Johannes, "Selection tournaments, sabotage, and participation," *Journal of Economics & Management Strategy*, 2007, *16* (4), 943–970.

Nagel, Thomas, "Moral luck," in Paul Russell and Oisín Deery, eds., *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*, Oxford University Press, 1979.

Noe, Thomas, "Comparing the chosen: Selection bias when selection is competitive," *Journal of Political Economy*, 2020, *128* (1), 342–390.

Payne, David G, "The Value of Testing in Graduate Admissions," *Inside Higher Ed*, 2018.

Peltz, Lois, *The New Investment Superstars: 13 Great Investors and Their Strategies for Superior Returns*, John Wiley & Sons, 2001.

Pluchino, Alessandro, Alessio Emanuele Biondo, and Andrea Rapisarda, "Talent versus luck: The role of randomness in success and failure," *Advances in Complex Systems*, 2018, *21* (3–4), 1850014:1–31.

Ryvkin, Dmitry, "The selection efficiency of tournaments," *European Journal of Operational Research*, 2010, *206* (3), 667–675.

_ **and Andreas Ortmann**, "The predictive power of three prominent tournament formats," *Management Science*, 2008, *54* (3), 492–504.

Shaked, Moshe and J George Shanthikumar, *Stochastic Orders*, Springer Science & Business Media, 2007.

Taleb, Nassim, *Fooled by randomness: The hidden role of chance in life and in the markets*, Vol. 1, Random House Incorporated, 2005.

Varian, Hal R, "Redistributive taxation as social insurance," *Journal of Public Economics*, 1980, *14* (1), 49–68.

Vega-Redondo, Fernando, "The evolution of Walrasian behavior," *Econometrica*, 1997, *65* (2), 375–384.

Wärneryd, Karl, "Replicating contests," *Economics Letters*, 2001, *71* (3), 323–327.

For Online Publication

Online Appendices to "Does the cream rise to the top"

by Thomas Noe and Dawei Fang

A Proofs of results

Proof of Theorem 1 This proof will be initiated by establishing some simple preliminary results.

The first result shows that talent will be identified if and only if the function integrated in equation (4) converges to zero almost surely.

Lemma A-1. $\lim_{n\to\infty}\Pi_A(n) = 0$ *if and only if* $\lim_{n\to\infty} U_n(s) = 0$*, for almost all* $s \in [0,1]$.

Proof. First note that $s \hookrightarrow u(s^{1/(a_o n)}))^{t_o n}$ is positive and bounded by the constant function, 1, over the interval $(0,1)$. Thus, Lebesgue's dominated convergence theorem (LDCT) implies

$$\lim_{n\to\infty}\Pi_A(n) = \int_0^1 \left(\lim_{n\to\infty}\left(u(s^{1/(a_o n)})\right)^{t_o n}\right) ds$$

Thus, $\lim_{n\to\infty}\left(u(s^{1/(a_o n)})\right)^{t_o n} = 0$ almost surely implies $\lim_{n\to\infty}\Pi_A(n) = 0$.

To prove the converse, let $\zeta_n(s) = \left(u(s^{1/(a_o n)})\right)^{t_o n}$. $\lim_{n\to\infty}\Pi_A(n) = 0$ is equivalent to

$$\lim_{n\to\infty}\int_0^1 |\zeta_n(s)|\, ds = 0.$$

Thus, $\lim_{n\to\infty}\Pi_A(n) = 0$ is equivalent to $\{\zeta_n\}_n \to 0$ in mean, which implies that $\{\zeta_n\}_n \to 0$ in probability, which, in turn, implies convergence in distribution. Thus, $\zeta_n \xrightarrow{\text{dist}} F_0$, where F_0 is the distribution function for a point mass at 0. Next, note that the distribution function of ζ_n is given by $F_n(x) = m\{s : \zeta_n(s) \le x\}$, where m denotes Lebesgue measure. Because, for all n, ζ_n is increasing and continuous, each ζ_n has an increasing continuous inverse, ζ_n^{-1}. Thus,

$$m\{s : \zeta_n(s) \le x\} = m\{s : s \le \zeta_n^{-1}(x)\} = \zeta_n^{-1}(x).$$

Because $F_0(x) = 1$ for $x \in (0,1)$ and F_0 is continuous over $(0,1)$, convergence in distribution implies that,

$$\text{for all } x \in (0,1), \quad \lim_{n\to\infty}\zeta_n^{-1}(x) = 1. \tag{A-1}$$

Fix any $s \in (0,1)$. Let $\varepsilon \in (0,1)$. Equation (A-1) implies that for n sufficiently large,

$$\zeta_n^{-1}(\varepsilon) > s. \tag{A-2}$$

Equation (A-2) implies that $\zeta_n(s) < \varepsilon$, for n sufficiently large. Thus, for all $s \in (0,1)$, $\lim_{n\to\infty} \zeta_n(s) = 0$, which implies that $\zeta_n(s) \to 0$ almost surely. □

Definition A-1. If $\phi : [0,1] \to [0,1]$ is continuous, increasing, and $\phi(1) = 1$, let $\hat{\phi} : (-\infty, 0] \to (-\infty, 0]$ represent the *conjugate function* to ϕ, defined as follows:

$$\hat{\phi}(y) = \log \circ \phi \circ \exp(y), \quad y \leq 0.$$

Lemma A-2. *Let $\hat{u}$ be the conjugate function to u (defined in equation* (2)*), then*

$$\lim_{y\to 0} \frac{\hat{u}(y)}{y} = \lim_{x\to h} \frac{1 - F_T(x)}{1 - F_A(x)},$$

i.e., if either limit exists, so does the other, in which case, both limits are the same.

Proof. First note that, by Definition A-1,

$$\frac{\hat{u}(y)}{y} \equiv \frac{\log \circ u(e^y)}{y}. \tag{A-3}$$

Next, let $y = \log(t)$, $t \in (0,1)$. Then because, log is an increasing continuous function,

$$\lim_{y\to 0} \frac{\log \circ u(e^y)}{y} = \lim_{t\to 1} \frac{\log \circ u(t)}{\log t}. \tag{A-4}$$

Then let $t = F_A(x)$, $t \in (\ell, h)$. Because, F_A is strictly increasing and continuous, and because by definition, $u \circ F_A = F_T$,

$$\lim_{t\to 1} \frac{\log \circ u(t)}{\log t} = \lim_{x\to h} \frac{\log F_T(x)}{\log F_A(x)}. \tag{A-5}$$

Now note that, for $j = T, A$, $\log F_j = \log(1 - (1 - F_j))$ Because $1 - F_j(x) \to 0$ as $x \to h$, and because $\log(1 - z) \sim -z$ as $z \to 0$,

$$\lim_{x\to h} \frac{\log F_T(x)}{\log F_A(x)} = \lim_{x\to h} \frac{1 - F_T(x)}{1 - F_A(x)}. \tag{A-6}$$

The lemma follows directly from equations (A-3), (A-4), (A-5), and (A-6). □

Proof of Theorem 1 (cont.) Proof of part i. First note that, by Definition A-1 and equation (4), the conjugate function to U_n is given by

$$\hat{U}_n(y) = t_o\, n \log \circ u\left((e^y)^{1/(a_o n)}\right) = t_o\, n\, \hat{u}\left(\tfrac{y}{a_o n}\right), \quad y \leq 0.$$

Now make the change of variable $\lambda = a_o n$. This yields

$$\hat{U}_n(y) = \frac{t_o}{a_o} \lambda \hat{u}\left(\lambda^{-1} y\right), \quad y \leq 0. \tag{A-7}$$

Hence,

$$\lim_{n\to\infty} \hat{U}_n(y) = \frac{t_o}{a_o} \lim_{\lambda\to\infty} \lambda \hat{u}\left(\lambda^{-1} y\right). \tag{A-8}$$

Now note that

$$\lim_{\lambda\to\infty} \lambda \hat{u}\left(\lambda^{-1} y\right) = \lim_{\lambda\to\infty} y \frac{\hat{u}\left(\lambda^{-1} y\right)}{\lambda^{-1} y}.$$

Because, for all $y < 0$, $\lambda^{-1} y \to 0$ as $\lambda \to \infty$, Lemma A-2 and the hypothesis of the theorem that $\lim_{x\to h} \frac{1-F_T(x)}{1-F_A(x)} = L$ imply

$$\lim_{\lambda\to\infty} \lambda\, \hat{u}\left(\lambda^{-1} y\right) = yL, \quad y < 0. \tag{A-9}$$

Thus, $\hat{U}_n(y) \to -\infty$ for all $y < 0$ almost surely if and only if $L = \infty$. Because, by Definition A-1,

$$U_n(s) = \exp \circ \hat{U}_n \circ \log(s), \quad s \in (0,\, 1], \tag{A-10}$$

$U_n(s) \to 0$ for all $s \in (0,\, 1)$ almost surely if and only if $L = \infty$. Lemma A-1 thus implies that $\lim_{n\to\infty} \Pi_A(n) = 0$ if and only if $L = \infty$.

Proof of part ii.

If L, as defined in the theorem, satisfies $L < \infty$, then (A-8) and (A-9) imply

$$\lim_{n\to\infty} \hat{U}_n(y) = y \frac{t_o}{a_o} L. \tag{A-11}$$

Now, note that

$$U_n(s) = \exp \circ \hat{U}_n \circ \log(s).$$

Because exp and log are continuous functions, equations (A-10) and (A-11) imply

$$\lim_{n\to\infty} U_n(s) = \exp[L \tfrac{t_o}{a_o} \log(s)] = s^{L t_o/a_o}, \quad s \in (0, 1].$$

Because U_n is uniformly bounded and defined over a compact set $[0,1]$, LDCT implies

$$\lim_{n\to\infty} \Pi_A(n) \;=\; \lim_{n\to\infty}\int_0^1 U_n(s)\,ds \;=\; \int_0^1 \lim_{n\to\infty} U_n(s)\,ds \;=\; \int_0^1 s^{Lt_o/a_o}\,ds \;=\; \frac{a_o}{a_o+Lt_o}.$$

□

Proof of Corollary 1. Let $\mu \equiv \mu_T - \mu_A > 0$. Without loss of generality, assume $\mu_A = 0$. Thus, $F_A = F$ and, for all x, $F_T(x) = F(x-\mu)$. Thus,

$$(1-F_T(x))/(1-F_A(x)) = (1-F(x-\mu))/(1-F(x)), \quad \text{for all } x.$$

To proceed, we first establish a technical lemma.

Lemma A-3. *If* $\lim_{x\to\infty} f'(x)/f(x)$ *exists (possibly infinite), then for all* $\mu > 0$, $\lim_{x\to\infty} f(x-\mu)/f(x)$ *exists and*

$$\lim_{x\to\infty}\frac{f(x-\mu)}{f(x)} = 1-\mu\,\lim_{x\to\infty}\frac{f'(x)}{f(x)}.$$

Proof. Noting that $-\mu < 0$ and using the Mean Value Theorem shows that

$$-\mu\,\sup\{f'(\xi):\xi\in(x-\mu,x)\} \le f(x-\mu)-f(x) \le -\mu\,\inf\{f'(\xi):\xi\in(x-\mu,x)\}.$$

Taking limits as $x\to\infty$, using the Squeeze Theorem, and noting that, by hypothesis, $\lim_{x\to\infty} f'(x)/f(x)$ exists (possibly infinite), shows that

$$\lim_{x\to\infty}\frac{f(x-\mu)-f(x)}{f(x)} = -\mu\,\lim_{x\to\infty}\frac{f'(x)}{f(x)}. \tag{A-12}$$

The result then follows immediately from equation (A-12). □

By L'Hopital's rule, if $\lim_{x\to\infty} f(x-\mu)/f(x)$ exists (possibly infinite),

$$\lim_{x\to\infty}\frac{1-F_T(x)}{1-F_A(x)} = \lim_{x\to\infty}\frac{f(x-\mu)}{f(x)}. \tag{A-13}$$

Equation (A-13) and Lemma A-3 then imply that, if $\lim_{x\to\infty} f'(x)/f(x)$ exists (possibly infinite),

$$\lim_{x\to\infty}\frac{1-F_T(x)}{1-F_A(x)} = 1-\mu\,\lim_{x\to\infty}\frac{f'(x)}{f(x)},$$

in which case, given $\mu>0$, $\lim_{x\to\infty}(1-F_T(x))/(1-F_A(x)) = \infty$ if and only if $\lim_{x\to\infty} f'(x)/f(x) = -\infty$. The result then follows immediately from Theorem 1. □

Proof of Corollary 2. If G_j denotes the distribution of a type-$j = T, A$ competitor's negative performance, then the support of G_j is $[-h, -\ell]$, $-\infty \leq -h < -\ell \leq \infty$, and $G_j(x) = 1 - F_j(-x)$ for all $x \in [-h, -\ell]$. Thus,

$$\lim_{x \to -\ell} (1 - G_T(x))/(1 - G_A(x)) = \lim_{x \to -\ell} (1 - (1 - F_T(-x)))/(1 - (1 - F_A(-x))) = \lim_{x \to \ell} F_T(x)/F_A(x).$$

Note that the worst performer is the competitor who has the highest negative performance. Thus, by Theorem 1, the asymptotic probability that the worst performer is average is determined by $\lim_{x \to -\ell}(1 - G_T(x))/(1 - G_A(x))$. The corollary thus follows from Theorem 1 and the above equation. □

Proof of Proposition 1. The proof requires a couple of preliminary results.

Lemma A-4.

$$\hat{u}'(y) = \frac{\exp(y)\, u'(\exp(y))}{u(\exp(y))}.$$

Proof. This result follows from Definition A-1 and differentiation. □

Lemma A-5.

$$\frac{F_A(x) u'(F_A(x))}{u(F_A(x))} = \frac{r_T(x)}{r_A(x)}, \quad x \in (\ell, h). \tag{A-14}$$

Proof. By the inverse function theorem and equation (2),

$$u'(s) = (F_T \circ F_A^{-1}(s))' = \frac{f_T(F_A^{-1}(s))}{f_A(F_A^{-1}(s))}, \quad s \in (0, 1).$$

Thus,

$$\frac{su'(s)}{u(s)} = \frac{s\, f_T(F_A^{-1}(s))}{f_A(F_A^{-1}(s))\, F_T(F_A^{-1}(s))}, \quad s \in (0, 1).$$

Hence,

$$\frac{F_A(x) u'(F_A(x))}{u(F_A(x))} = \frac{F_A(x)\, f_T(x)}{f_A(x)\, F_T(x)} = \frac{r_T(x)}{r_A(x)}, \quad x \in (\ell, h).$$

□

Lemma A-6. *$\hat{u}$ is (strictly) convex if and only if $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is nondecreasing (increasing). $\hat{u}$ is (strictly) concave if and only if $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is nonincreasing (decreasing).*

Proof. This result follows from Lemmas A-4 and A-5 and the fact that F_A is increasing. □

Proof of Proposition 1 (cont.) For any given $y \le 0$, define $\hat{u}_y(\lambda)$ by

$$\hat{u}_y(\lambda) = \lambda\, \hat{u}(\lambda^{-1} y). \tag{A-15}$$

Note that

$$\hat{u}_y'(\lambda) = \hat{u}(\lambda^{-1} y) - (\lambda^{-1} y)\, \hat{u}'(\lambda^{-1} y). \tag{A-16}$$

First consider the case where $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is nondecreasing. By Lemma A-6, $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is nondecreasing if and only if $\hat{u}$ is convex. By the definitions of u (equation (2)) and the conjugate function (Definition A-1), $\hat{u}(0) = 0$, and by the facts that log and exp are increasing functions, $y \hookrightarrow \hat{u}(y)$ is increasing for $y \le 0$. Given that $\hat{u}(0) = 0$ and $\hat{u}$ is increasing, the convexity of $\hat{u}$ implies that $\hat{u}(y) \le y\, \hat{u}'(y)$ for all $y \le 0$. Hence, by equation (A-16), for all $y \le 0$, $\hat{u}_y'(\lambda) \le 0$. Thus, for all $y \le 0$, $\lambda \hookrightarrow \hat{u}_y(\lambda)$ is nonincreasing. Because $\lambda = a_o\, n$, equations (A-7) and (A-15) imply that, for all $y \le 0$, $\hat{U}_n$ is nonincreasing in n. Because log and exp are increasing functions, the definition of the conjugate function, Definition A-1, implies that U_n is nonincreasing in n for all $s \in (0, 1)$. Because Π_A is the integral of U_n, Π_A is nonincreasing in n.

If $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is increasing, by Lemma A-6, $\hat{u}$ is strictly convex, in which case, for all $y \le 0$, $\lambda \hookrightarrow \hat{u}_y(\lambda)$ is decreasing. By essentially the same argument as just used, it follows that U_n is decreasing in n for all $s \in (0, 1)$ and, thus Π_A is decreasing in n.

Now consider the case where $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is nonincreasing. By Lemma A-6, $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is nonincreasing if and only if $\hat{u}$ is concave. Given that $\hat{u}(0) = 0$ and $\hat{u}$ is increasing, the concavity of $\hat{u}$ implies that $\hat{u}(y) \ge y\, \hat{u}'(y)$ for all $y \le 0$, in which case, for all $y \le 0$, $\lambda \hookrightarrow \hat{u}_y(\lambda)$ is nondecreasing. The rest of the argument, including the one for the case in which $x \hookrightarrow \frac{r_T(x)}{r_A(x)}$ is decreasing, follows in exactly the same fashion as the case of a convex $\hat{u}$ above. □

Proof of Proposition 2. (i): In a racing contest, a type-j competitor's performance distribution, F_j, $j = T, A$, is given by

$$F_j(x) = \mathbb{P}[-\tau_j \le x] = \mathbb{P}[-\tau/\lambda_j \le x] = \mathbb{P}[\tau \ge -\lambda_j x] = 1 - F(-\lambda_j x), \quad x \le 0. \tag{A-17}$$

By equation (A-17), if $\lim_{x \to 0^-} (1 - F_T(x))/(1 - F_A(x))$ exists,

$$\lim_{x \to 0^-} \frac{1 - F_T(x)}{1 - F_A(x)} = \lim_{x \to 0^-} \frac{F(-\lambda_T x)}{F(-\lambda_A x)} = \lim_{x \to 0^+} \frac{F((\lambda_T/\lambda_A)x)}{F(x)}. \tag{A-18}$$

Because, by Definition 1, in a racing contest, F varies regularly at 0, with ν representing the index of variation,

$$\lim_{x \to 0^+} \frac{F((\lambda_T/\lambda_A)x)}{F(x)} = (\lambda_T/\lambda_A)^{\nu}. \tag{A-19}$$

Equations (A-18) and (A-19) imply

$$\lim_{x\to 0^-} \frac{1-F_T(x)}{1-F_A(x)} = (\lambda_T/\lambda_A)^{\nu}. \tag{A-20}$$

Part (i) then follows from Theorem 1, equation (A-20), and the fact that, in a racing contest, the upper bound of the support of any competitor's performance distribution is 0.

(ii): By equation (A-17),

$$F_A^{-1}(s) = -\frac{1}{\lambda_A} F^{-1}(1-s), \quad s \in (0,1). \tag{A-21}$$

By the definition of u, given in equation (2), and by equations (A-17) and (A-21), in a racing contest, for $s \in (0,1)$,

$$u(s) = F_T \circ F_A^{-1}(s) = 1 - F(\lambda F^{-1}(1-s)), \tag{A-22}$$

where $\lambda = \lambda_T/\lambda_A > 1$. By the last part of the proof of Proposition 1, to establish part (ii), it suffices to show that the conjugate function to u, $\hat{u}$, is (strictly) concave if $h = f/(1-F)$ is (strictly) geometrically convex. By Lemma A-4, to show that $\hat{u}$ is (strictly) concave, it suffices to show that $s \hookrightarrow s u'(s)/u(s)$, $s \in (0,1)$, is nonincreasing (decreasing). Thus, to establish part (ii), it suffices to show that

$$\begin{aligned} &h = f/(1-F) \text{ is (strictly) geometrically convex} \\ &\quad \Longrightarrow s \hookrightarrow s u'(s)/u(s),\ s \in (0,1), \text{ is nonincreasing (decreasing).} \end{aligned} \tag{A-23}$$

Note that, by equation (A-22) and the inverse function theorem,

$$u'(s) = \frac{\lambda f(\lambda F^{-1}(1-s))}{f(F^{-1}(1-s))}, \quad s \in (0,1). \tag{A-24}$$

Equations (A-22) and (A-24) imply

$$\frac{s u'(s)}{u(s)} = \lambda s \frac{f(\lambda F^{-1}(1-s))}{f(F^{-1}(1-s))\,(1-F(\lambda F^{-1}(1-s)))}. \tag{A-25}$$

Let

$$\psi(x) = \frac{\lambda h(\lambda x)}{h(x)}, \quad x \geq 0, \tag{A-26}$$

where $h = f/(1-F)$ is the hazard rate function of F. By the definition of h,

$$\frac{\lambda h(\lambda x)}{h(x)} = \lambda \left(\frac{1-F(x)}{f(x)}\right)\left(\frac{f(\lambda x)}{1-F(\lambda x)}\right),$$

whose right hand side coincides with the right hand side of equation (A-25) when $x = F^{-1}(1 - s)$. Thus, by equations (A-25) and (A-26),

$$\frac{su'(s)}{u(s)} = \psi \circ F^{-1}(1-s), \quad s \in (0,1). \tag{A-27}$$

The next technical lemma shows that the direction of monotonicity of $h(\lambda x)/h(x)$ is determined by whether h is geometrically convex or concave.

Lemma A-7. *Let $\phi : (0,+\infty) \to (0,+\infty)$ be a continuous, differentiable function. If $a > 1$, then*

(i) $x \hookrightarrow \frac{\phi(ax)}{\phi(x)}$, $x > 0$, *is nondecreasing (increasing) if ϕ is (strictly) geometrically convex;*

(ii) $x \hookrightarrow \frac{\phi(ax)}{\phi(x)}$, $x > 0$, *is nonincreasing (decreasing) if ϕ is (strictly) geometrically concave.*

Proof. Note that

$$\frac{\partial}{\partial x}\left(\frac{\phi(ax)}{\phi(x)}\right) \geq (>)(\leq)(<)0$$
$$\implies x \hookrightarrow \frac{\phi(ax)}{\phi(x)} \text{ is nondecreasing (increasing) (nonincreasing) (decreasing).} \tag{A-28}$$

$$x\frac{\partial}{\partial x}\left(\frac{\phi(ax)}{\phi(x)}\right) = \left(\frac{\phi(ax)}{\phi(x)}\right)\left(\frac{xa\phi'(xy)}{\phi(xa)} - \frac{x\phi'(x)}{\phi(x)}\right). \tag{A-29}$$

Because a continuous, differentiable function, $\phi : (0,+\infty) \to (0,+\infty)$, is (strictly) geometrically convex if $x \hookrightarrow x\phi'(x)/\phi(x)$ is nondecreasing (increasing), and is (strictly) geometrically concave if $x \hookrightarrow x\phi'(x)/\phi(x)$ is nonincreasing (decreasing), the lemma follows from inspection of (A-28) and (A-29). □

By Lemma A-7 and the fact that $\lambda = \lambda_T/\lambda_A > 1$, $x \hookrightarrow h(\lambda x)/h(x)$ is nondecreasing (increasing) for $x > 0$ if h is (strictly) geometrically convex. Thus, by equation (A-26), ψ is nondecreasing (increasing) in $x > 0$ if h is (strictly) geometrically convex. Because $s \hookrightarrow F^{-1}(1-s)$, $s \in (0,1)$, is decreasing, it follows that $s \hookrightarrow \psi \circ F^{-1}(1-s)$, $s \in (0,1)$, is nonincreasing (decreasing) if h is (strictly) geometrically convex. This result, combined with equation (A-27), establishes (A-23), and part (ii) follows.

(iii): By Lemma A-7 and the fact that $\lambda > 1$, $x \hookrightarrow h(\lambda x)/h(x)$ is nonincreasing (decreasing) for $x > 0$ if h is (strictly) geometrically concave. The rest of the argument follows in exactly the same fashion as the geometrically convex case above. □

Proof of Proposition 3. (i): In a target-hitting contest, a type-j competitor's performance distribution, F_j, $j = T, A$, is given by

$$\begin{aligned} F_j(x) = \mathbb{P}[-g(|\varepsilon_j|) \le x] = \mathbb{P}[-g(|\varepsilon|/\lambda_j) \le x] = \mathbb{P}[|\varepsilon| \ge \lambda_j\, g^{-1}(-x)] \\ = 2\mathbb{P}[\varepsilon \le -\lambda_j\, g^{-1}(-x)] = 2F(-\lambda_j\, g^{-1}(-x)), \quad x \le 0, \end{aligned} \tag{A-30}$$

where the first equality in the second line follows from the symmetry of the target-hitting distribution. Because $g : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing, continuous, and $g(0) = 0$, its inverse function, $g^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$, exists and is increasing, continuous, and $g^{-1}(0) = 0$. By equation (A-30) and the definition of L in Theorem 1, we have

$$\begin{aligned} L = \lim_{x \to 0^-} \frac{1 - F_T(x)}{1 - F_A(x)} = \lim_{x \to 0^-} \frac{1 - 2F(-\lambda_T\, g^{-1}(-x))}{1 - 2F(-\lambda_A\, g^{-1}(-x))} \\ = \lim_{y \to 0^-} \frac{1 - 2F(\lambda_T\, y)}{1 - 2F(\lambda_A\, y)} = \frac{\lambda_T}{\lambda_A} \lim_{y \to 0^-} \frac{f(\lambda_T\, y)}{f(\lambda_A\, y)} = \frac{\lambda_T}{\lambda_A}, \end{aligned} \tag{A-31}$$

where the first equality in the second line follows from substitution, using $y = -g^{-1}(-x)$, and the facts that $g^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and $g^{-1}(0) = 0$. The second equality in the second line follows from L'Hopital's rule and the last equality from the hypothesis that the target-hitting distribution, F, is unimodal and symmetric about 0, which implies that $f(0) > 0$. Part (i) then follows immediately from equation (A-31) and Theorem 1.

(ii) and (iii): By equation (A-30),

$$F_A^{-1}(s) = -g\left(-\frac{1}{\lambda_A} F^{-1}\left(\frac{s}{2}\right)\right). \tag{A-32}$$

By the definition of u, given in equation (2), and by equations (A-30) and (A-32), in a target-hitting contest, for $s \in (0, 1)$,

$$u(s) = F_T \circ F_A^{-1}(s) = 2F\left(\lambda\, F^{-1}\left(\frac{s}{2}\right)\right), \tag{A-33}$$

where $\lambda = \lambda_T/\lambda_A > 1$.

Below, we show that

$$\begin{aligned} h = f/(1-F) \text{ is (strictly) geometrically convex over } (0, \infty) \\ \implies s \hookrightarrow s\,u'(s)/u(s),\ s \in (0, 1), \text{ is nonincreasing (decreasing).} \end{aligned} \tag{A-34}$$

By equation (A-33) and the inverse function theorem, we have

$$\frac{s\,u'(s)}{u(s)} = \frac{\lambda\, s}{2}\, \frac{f\left(\lambda\, F^{-1}\left(\frac{s}{2}\right)\right)}{f\left(F^{-1}\left(\frac{s}{2}\right)\right) F\left(\lambda\, F^{-1}\left(\frac{s}{2}\right)\right)}, \quad s \in (0,\,1). \tag{A-35}$$

Let

$$\zeta(x) = \lambda\, \frac{h(-\lambda\, x)}{h(-x)}, \quad x \leq 0, \tag{A-36}$$

where $h = f/(1-F)$ is the hazard rate function of F. By the definition of h and the symmetry of the target-hitting distribution,

$$\zeta(x) = \lambda\, \frac{(1-F(-x))}{f(-x)}\, \frac{f(-\lambda\, x)}{(1-F(-\lambda\, x))} = \lambda\, \frac{F(x)}{f(x)}\, \frac{f(\lambda\, x)}{F(\lambda\, x)}, \quad x \leq 0,$$

where the last expression coincides with the right hand side of equation (A-35) when $x = F^{-1}(s/2)$. Thus, by equation (A-35),

$$\frac{s\,u'(s)}{u(s)} = \zeta \circ F^{-1}(s/2), \quad s \in (0,\,1). \tag{A-37}$$

By Lemma A-7, equation (A-36), and the facts that $\lambda > 1$ and $x \hookrightarrow -x$ is decreasing, ζ is nonincreasing (decreasing) in $x < 0$ if h is (strictly) geometrically convex over $(0,\,\infty)$. Thus, given that $s \hookrightarrow F^{-1}(s/2)$, $s \in (0,\,1)$, is increasing, equation (A-34) follows from equation (A-37).

Given the satisfaction of (A-34), by a similar argument as used in the proof of Proposition 2, it is clear that $n \hookrightarrow \Pi_A(n)$ is nondecreasing (increasing) if h is (strictly) geometrically convex. The result that $n \hookrightarrow \Pi_A(n)$ is nonincreasing (decreasing) if h is (strictly) geometrically concave follows in exactly the same fashion as the geometrically convex case.

To complete the proof of the proposition, it suffices to show that h is strictly geometrically convex if the target-hitting distribution, F, is Normal. Baricz (2010, Section 5.1) shows that the Mills ratio, $m = (1-F)/f$, of the standard Normal distribution is strictly geometrically concave over $(0,\,\infty)$. Because geometric concavity of m is invariant to a scale transformation of the standard Normal distribution, the Mills ratio of any zero-mean Normal distribution is strictly geometrically concave over $(0,\,\infty)$. Because, by the definition of the hazard rate, $h = f/(1-F) = 1/m$ and because $1/m$ is strictly geometrically convex if m is strictly geometrically concave, if F is Normal, its hazard rate, h, is strictly geometrically convex over $(0,\,\infty)$. $\square$

Proof of Corollary 3. Let $p_A(n)$ be an average competitor's probability of being the best performer when the scale-of-competition factor is n. Note that

$$a_o\, n\, p_A(n) = \Pi_A(n), \tag{A-38}$$

where $a_o > 0$ is a constant. By Proposition 1, if $x \hookrightarrow r_T(x)/r_A(x)$ is nondecreasing, $\Pi_A(n) \leq \Pi_A(1)$ for all $n > 1$, in which case, by equation (A-38), $n\,p_A(n) \leq p_A(1)$ for all $n > 1$. Thus, given that the expected prize received by an average competitor is $n v p_A(n)$ in contest $\mathscr{C}_n$, $n > 1$, and is $v p_A(1)$ in contest $\mathscr{C}$, if $x \hookrightarrow r_T(x)/r_A(x)$ is nondecreasing, an average competitor's expected prize is weakly lower in $\mathscr{C}_n$ than in $\mathscr{C}$. The other cases can be proved analogously. □

Proof of Result 1. Let $p_j(t,a)$ be the probability of winning for a type-j artist, $j = T, A$, when there are t talented and a average artists in a contest. The hypothesis that in each local contest before globalization, when all t_o talented agents become an artist, it is still profitable for some average agents to be an artist, implies

$$\lim_{a \to 0} v\,p_A(t_o, a) > w. \tag{A-39}$$

Because $a \hookrightarrow p_A(t_o, a)$ is decreasing and $\lim_{a\to\infty} p_A(t_o,a) = 0$, equation (A-39) implies the existence of $a_o > 0$ such that

$$v\,p_A(t_o, a_o) = w. \tag{A-40}$$

Because, by hypothesis, F_T MLR dominates F_A, $p_T(t_o, a_o) > p_A(t_o, a_o)$. Thus, by equation (A-40), before globalization, when each region has t_o talented and a_o average artists, all talented agents prefer being an artist and all average agents are indifferent between being an artist and taking a regular job. Hence, given that a region has an infinite number of average agents but only t_o talented agents, an equilibrium exists in which, before globalization, each region has t_o talented and a_o average artists. Because this equilibrium satisfies the condition for a Pareto dominant equilibrium (i.e., no average artists unless all talented agents become an artist), this equilibrium is a Pareto dominant equilibrium.

In what follows, let t_n and a_n be, respectively, the number of talented and average artists in each region when globalization merges the art contests in $n \geq 2$ regions. We first argue that any Pareto dominant equilibrium must have $t_n = t_o$. Suppose, to the contrary, that there exists a Pareto dominant equilibrium in which $t_n \neq t_o$. Because each region has t_o talented agents, $t_n \neq t_o$ implies $t_n < t_o$. Because, in a Pareto dominant equilibrium, there are no average artists unless all talented agents become an artist, $t_n < t_o$ implies $a_n = 0$. Thus, in this equilibrium, after globalization, a talented artist's probability of winning satisfies

$$p_T(n t_n, n a_n) = p_T(n t_n, 0) > p_T(n t_o, 0) = p_T(t_o, 0)/n > p_T(t_o, a_o)/n > p_A(t_o, a_o)/n, \tag{A-41}$$

where the first inequality and the second equality follow from the fact that $p_T(t, 0) = 1/t$ for all $t > 0$, the second inequality from $a \hookrightarrow p_T(t_o, a)$ being decreasing, and the last inequality from the hypothesis that F_T MLR dominates F_A. Because a talented artist's expected payoff after globalization equals $n v p_T(n t_n, n a_n)$, equations (A-40) and (A-41) imply that a talented artist's expected payoff after globalization is strictly greater than w. Thus, the talented agents who take a regular job would have been better off had they chosen to be an artist, a contradiction. This contradiction implies that in any Pareto dominant equilibrium, $t_n = t_o$. Part (i) thus follows.

By Corollary 3 and part (i), if globalization had no effect on an average agent's decision, then after globalization, average artists would receive an expected payoff strictly larger (smaller) than w if $x \hookrightarrow r_T(x)/r_A(x)$ is decreasing (increasing). Thus, to attain a new Pareto dominant equilibrium, the number of average artists after globalization must increase (decrease) if $x \hookrightarrow r_T(x)/r_A(x)$ is decreasing (increasing). Part (ii).(a) and the first half of part (ii).(b) thus follow.

To establish the second half of part (ii).(b), let $a^* = \lim_{n\to\infty} a_n$. Let $\Pi_A(nt, na)$ be the probability that the best performer is average when there are nt talented artists and na average artists. Note that, by Theorem 1,

$$\lim_{n\to\infty} \Pi_A(n t_o, n a_n) = \frac{a^*}{a^* + L t_o}, \tag{A-42}$$

where $L = \lim_{x\to h}(1 - F_T(x))/(1 - F_A(x))$ (possibly infinite). Also note that, for all $a_n > 0$,

$$p_A(n t_o, n a_n) = \Pi_A(n t_o, n a_n)/(n a_n).$$

Thus, given that the prize value for the global contest equals nv, whenever $a_n > 0$, the expected prize received by an average competitor is given by

$$n v\, p_A(n t_o, n a_n) = v \Pi_A(n t_o, n a_n)/a_n. \tag{A-43}$$

Equations (A-42) and (A-43), combined with the fact that $a^* = \lim_{n\to\infty} a_n$, imply that, whenever $a^* > 0$, in the limit, the expected payoff to an average artist after globalization equals $v/(a^* + L t_o)$. Thus, an average agent has an incentive to be an artist after globalization if and only if $v/(a^* + L t_o) \geq w$. This implies that, in the limit, there is a positive fraction of average artists after globalization if and only if there exists $a^* > 0$ such that $v/(a^* + L t_o) \geq w$. This condition is equivalent to

$$L < v/(w t_o).$$

The second half of part (ii).(b) thus follows. □

Proof of Proposition 4. Let $F_{T,K}$ and $F_{A,K}$ be, respectively, a talented and an average competitor's performance distribution in a best-shot contest with K shots for each competitor. Note that

$$F_{T,K}(x) = (F_T(x))^K \quad \& \quad F_{A,K}(x) = (F_A(x))^K, \quad x \in [\ell, h]. \tag{A-44}$$

For any given $s \in (0,1)$, $F_{A,K}^{-1}(s)$ represents the s-quantile of the distribution of an average competitor's performance in a best-shot contest with K shots for each competitor. If an average competitor's performance equals $F_{A,K}^{-1}(s)$, $s \in (0,1)$, her probability of besting any given average rival, whose performance is denoted by $X_{A,K}$, is given by

$$\mathbb{P}[X_{A,K} \leq F_{A,K}^{-1}(s)] = F_{A,K} \circ F_{A,K}^{-1}(s) = s, \tag{A-45}$$

and her probability of besting any given talented rival, whose performance is denoted by $X_{T,K}$, is given by

$$\mathbb{P}[X_{T,K} \leq F_{A,K}^{-1}(s)] = F_{T,K} \circ F_{A,K}^{-1}(s) = (F_T \circ F_A^{-1}(s^{(1/K)}))^K = (u(s^{(1/K)}))^K, \tag{A-46}$$

where the second equality follows from equation (A-44) and the last from the definition of u given in equation (2).

Note that $(u(s^{(1/K)}))^K$ coincides with U_n defined in equation (4) with $n = K$ and $a_o = t_o = 1$. In the proof of Proposition 1, we showed that, for $a_o, t_o \geq 1$, U_n is nonincreasing (decreasing) in n if $x \hookrightarrow r_T(x)/r_A(x)$ is nondecreasing (increasing), and U_n is nondecreasing (increasing) in n if $x \hookrightarrow r_T(x)/r_A(x)$ is nonincreasing (decreasing). Thus, for any $s \in (0,1)$ and any $K'' > K' \geq 1$,

$$x \hookrightarrow r_T(x)/r_A(x) \text{ is nondecreasing (increasing)} \implies (u(s^{(1/K'')}))^{K''} \leq (<) (u(s^{(1/K')}))^{K'} \tag{A-47}$$

$$x \hookrightarrow r_T(x)/r_A(x) \text{ is nonincreasing (decreasing)} \implies (u(s^{(1/K'')}))^{K''} \geq (>) (u(s^{(1/K')}))^{K'} \tag{A-48}$$

Suppose $K'' > K' \geq 1$. Equations (A-46) and (A-47) imply that, if $x \hookrightarrow r_T(x)/r_A(x)$ is nondecreasing (increasing), for any given $s \in (0, 1)$, an average competitor has a (strictly) lower probability of besting any given talented rival if (a) the average competitor is in a K''-shot contest and has performance equal to $F_{A,K''}^{-1}(s)$ than if (b) she is in a K'-shot contest and has performance equal to $F_{A,K'}^{-1}(s)$. Note that, by equation (A-45), an average competitor's probability of besting any given average rival is the same no matter whether she is in situation (a) or (b). Thus, if $x \hookrightarrow r_T(x)/r_A(x)$ is nondecreasing (increasing), an average competitor's performance rank in situation (a) is (strictly) lower than her rank in situation (b) in the sense of first-order stochastic dominance. Also note that, by equation (A-45), for any given $s \in (0, 1)$, the probability (density) that an average competitor's performance equals $F_{A,K''}^{-1}(s)$ in a K''-shot contest is the same as the probability (density) that an average competitor's performance equals $F_{A,K'}^{-1}(s)$ in a K'-shot contest. Therefore, an average competitor's performance rank must be (strictly) lower in the K''-shot contest than in the K'-shot contest in the sense of first-order stochastic dominance if $x \hookrightarrow r_T(x)/r_A(x)$ is nondecreasing (increasing). This establishes part (i).

Equations (A-46) and (A-48) imply that, if $x \hookrightarrow r_T(x)/r_A(x)$ is nonincreasing (decreasing), for any given $s \in (0, 1)$, an average competitor has a (strictly) higher probability of besting any given talented rival in situation (a) than (b). Part (ii) then follows from this fact and an argument analogous to the one used above. □

B No limit to the tail ratio and talent identification

Suppose $\lim_{x\to h}(1-F_T(x))/(1-F_A(x))$ does not exist. Following almost exactly the steps in the proof of Lemma A-2, it is easy to show that

$$\limsup_{y\to 0}\frac{\hat{u}(y)}{y}=\limsup_{x\to h}\frac{1-F_T(x)}{1-F_A(x)}.$$

Hence, if $\limsup_{x\to h}(1-F_T(x))/(1-F_A(x))<\infty$, then

$$\limsup_{y\to 0}\frac{\hat{u}(y)}{y}<\infty.$$

Then, using an argument completely analogous to the argument in the proof of Theorem 1, it is easy to show that

$$\limsup_{y\to 0}\frac{\hat{u}(y)}{y}<\infty \Longrightarrow \text{for any } s\in(0,1),\ U_n(s) \text{ does not converge a.s. to } 0.$$

Thus, $\limsup_{x\to h}(1-F_T(x))/(1-F_A(x))<\infty$, implies that, for any $s\in(0,1)$, $U_n(s)$ does not converge a.s. to 0. This implies, by an argument analogous to the one used in the proof of Lemma A-1, that $\Pi_A(n)$ does not converge to 0. Thus, if $\limsup_{x\to h}(1-F_T(x))/(1-F_A(x))<\infty$, talent is not identified by competition. Of course, because, by supposition, $\lim_{x\to h}(1-F_T(x))/(1-F_A(x))$ does not exist, the characterization of the limit of Π_A provided in part (ii) of Theorem 1 generally will not be valid.

Because stochastic dominance implies that $(1-F_T(x))/(1-F_A(x))\geq 1$ for $x<h$,

$$\liminf_{x\to h}\frac{1-F_T(x)}{1-F_A(x)}\geq 1>-\infty.$$

Thus,

$$\limsup_{x\to h}(1-F_T(x))/(1-F_A(x))<\infty \iff \limsup_{x\to h}\frac{1-F_T(x)}{1-F_A(x)}-\liminf_{x\to h}\frac{1-F_T(x)}{1-F_A(x)}<\infty.$$

Hence, if the tail ratio function, $x\hookrightarrow(1-F_T(x))/(1-F_A(x))$, exhibits bounded oscillation over every upper interval (x_o,h), $x_o<h$, talent cannot be identified by competition.

Result B-1. If $\limsup_{x\to h}(1-F_T(x))/(1-F_A(x))<\infty$, which is equivalent to $x\hookrightarrow(1-F_T(x))/(1-F_A(x))$ exhibiting bounded oscillation as $x\to h$, talent cannot be identified by competition.

In contrast, if

$$\limsup_{x\to h}\frac{1-F_T(x)}{1-F_A(x)}=\infty>\liminf_{x\to h}\frac{1-F_T(x)}{1-F_A(x)},$$

in which case,

$$\limsup_{x \to h} \frac{1 - F_T(x)}{1 - F_A(x)} - \liminf_{x \to h} \frac{1 - F_T(x)}{1 - F_A(x)} = \infty,$$

i.e., $x \hookrightarrow (1 - F_T(x))/(1 - F_A(x))$ exhibits unbounded oscillation over every upper interval (x_o, h), $x_o < h$, the asymptotic behavior of the tail ratio does not determine whether talent is identified. Tail ratios with unbounded oscillation can be constructed (example available upon request). However, producing unbounded oscillation under the restrictions imposed by Assumption 1, which require that the talented and average competitors' performance distributions are fairly smooth and ordered by first-order stochastic dominance, does require rather artful constructions and non text-book distributions. Thus, we believe that this limitation of Theorem 1 has no practical import.

C Effect of alternative talent assignment mechanisms

C.1 Robustness of Theorem 1 to alternative assignment mechanisms

In our baseline model, we assumed that the mechanism for assigning ability to competitors was a collection of identically distributed indicator functions, $\{I^i\}$, satisfying

$$\sum_{i=1}^{C} I^i = a,$$

where C represents the total number of competitors and a represents the number of average competitors. Under this ability assignment mechanism, although the ability of each individual competitor is random, there is no aggregate uncertainty about the total number of average and talented competitors.

In this appendix, we demonstrate that our asymptotic results are quite robust to alternative ability assignment mechanisms including assignment mechanisms under which $\{I^i\}$ are jointly independent (i.e., ability is determined by a sequence of independent draws from a Bernoulli distribution) as well as many specifications of positive and negative dependence between the assignment variables, I^i. Our result shows that as long as the hypothesis of Theorem 1 is satisfied, and the strong law of large numbers holds for the sequence of assignment variables, $\{I^i\}$, the asymptotic probability that the best performer is average equals the probability specified in Theorem 1. Later, we remark that, for bounded identically distributed collections of random variables such as $\{I^i\}$, the strong law of large numbers holds under fairly weak conditions even absent the assumption of independence.

Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let $\{I^i : i \in \mathbb{N}\}$ be a countable set of identically distributed random indicator functions, i.e., identically distributed random variables that assume values in $\{0,1\}$, P-almost surely. Let $\bar{\pi} \in (0,1)$ represent the probability that a competitor is average, i.e, for all i, $\mathbb{P}[I^i = 1] = \bar{\pi}$. Let

$$A_C(\omega) = \sum_{i=1}^{C} I^i$$

represent the number of average competitors given that the total number of competitors is C. Using equation (3), we see that, under this specification, the unconditional probability that an average competitor will be the best performer is given by

$$\Pi_A(C) = \int\int_{\Omega \times [0,1]} (u(s^{1/A_C(\omega)}))^{C - A_C(\omega)} \, d(P \times m), \tag{C-49}$$

where m represents Lebesgue measure.

Lemma C-1. *If*

(i) the strong law of large numbers holds, i.e., $\lim_{C\to\infty} A_C/C = \bar{\pi}$*, P-almost surely, and*

(ii) $\lim_{x\to h}(1-F_T(x))/(1-F_A(x)) := L$ *exists (possibly infinite),*

then

$$\lim_{C\to\infty} \Pi_A(C) = \begin{cases} 0 & L=\infty \\ \frac{\bar{\pi}}{\bar{\pi}+L(1-\bar{\pi})} & L<\infty \end{cases}.$$

Proof. Because the integrand in equation (C-49) is positive and bounded and both measures are finite, Fubini's Theorem (Theorem 2.6.4 Ash, 1972) implies

$$\Pi_A(C) = \int_0^1 \left(\int_\Omega (u(s^{1/A_C(\omega)}))^{C-A_C(\omega)}\, dP(\omega) \right) ds. \tag{C-50}$$

Analogous to the definition of U_n in equation (4), define $U_{C,\omega}$ as

$$U_{C,\omega}(s) = (u(s^{1/A_C(\omega)}))^{C-A_C(\omega)}. \tag{C-51}$$

The first step in the proof is the following result:

Result C-1. If the strong law of large numbers holds, i.e., $\lim_{C\to\infty} A_C/C = \bar{\pi}$, P-almost surely, then for $s \in (0,1)$,

$$\lim_{C\to\infty} U_{C,\omega}(s) = \begin{cases} 0 & L=\infty \\ s^{\frac{1-\bar{\pi}}{\bar{\pi}}L} & L<\infty \end{cases}.$$

Proof of Result C.1. Following the steps in the proof of Theorem 1, first note that the conjugate function to $U_{C,\omega}$ satisfies

$$\hat{U}_{C,\omega}(y) = (C-A_C(\omega)) \log\circ u\left((e^y)^{1/A_C(\omega)}\right) = \left(\frac{C-A_C(\omega)}{A_C(\omega)}\right)\left(A_C(\omega)\,\hat{u}\left(\frac{y}{A_C(\omega)}\right)\right), \quad y \le 0, \tag{C-52}$$

where $\hat{u}$ is the conjugate function to u.[25] The strong law of large numbers implies that, as $C\to\infty$, $A_C\to\infty$, P-almost surely. Thus, by exactly the same argument as used in the proof of Theorem 1,

$$\lim_{C\to\infty} A_C(\omega)\,\hat{u}\left(\frac{y}{A_C(\omega)}\right) = yL, \quad P\text{-a.s. for } y<0.$$

The strong law of large numbers implies

$$\lim_{C\to\infty} \frac{C-A_C(\omega)}{A_C(\omega)} = \frac{1-\bar{\pi}}{\bar{\pi}}, \quad P\text{-a.s.}$$

[25] See Definition A-1 for the definition of the conjugate function.

Thus,

$$\lim_{C\to\infty}\left(\frac{C-A_C(\omega)}{A_C(\omega)}\right)\left(A_C(\omega)\,\hat{u}\left(\frac{y}{A_C(\omega)}\right)\right)=\frac{1-\bar{\pi}}{\bar{\pi}}yL,\quad P\text{-a.s. for } y<0. \tag{C-53}$$

Because $y<0$, equations (C-52) and (C-53) imply that, for $L=\infty$,

$$\lim_{C\to\infty}(C-A_C(\omega))\,\hat{u}\left(\frac{y}{A_C(\omega)}\right)=-\infty \Longrightarrow \lim_{C\to\infty}(u(s^{1/A_C(\omega)}))^{C-A_C(\omega)}=0,\quad P\text{-a.s. for } s\in(0,1),$$

and that, for $L<\infty$,

$$\lim_{C\to\infty}(C-A_C(\omega))\,\hat{u}\left(\frac{y}{A_C(\omega)}\right)=\frac{1-\bar{\pi}}{\bar{\pi}}yL$$
$$\Longrightarrow \lim_{C\to\infty}(u(s^{1/A_C(\omega)}))^{C-A_C(\omega)}=s^{\frac{1-\bar{\pi}}{\bar{\pi}}L},\quad P\text{-a.s. for } s\in(0,1).$$

The result then follows immediately from equation (C-51). □

Equation (C-50), two applications of LDCT, Result C.1, and simple integration then imply that for $L<\infty$,

$$\lim_{C\to\infty}\Pi_A(C)=\int_0^1\left(\lim_{C\to\infty}\int_\Omega(u(s^{1/A_C(\omega)}))^{C-A_C(\omega)}\,dP(\omega)\right)ds=$$
$$\int_0^1\int_\Omega\left(\lim_{C\to\infty}(u(s^{1/A_C(\omega)}))^{C-A_C(\omega)}\right)dP(\omega)\,ds=\int_0^1\left(s^{\frac{1-\bar{\pi}}{\bar{\pi}}L}\right)ds=\frac{\bar{\pi}}{\bar{\pi}+L(1-\bar{\pi})}.$$

Using an identical argument, for $L=\infty$, we see that $\lim_{C\to\infty}\Pi_A(C)=0$. □

Remark 1. The strong law of large numbers (SLLN), i.e.,

$$\lim_{C\to\infty}\frac{A_C}{C}=\bar{\pi},\quad P\text{-a.s.},$$

will always be satisfied if $\{I^i\}$ are independent or negative quadrant dependent (NQD) (Matuła, 1992). In the case of indicator functions, the NQD condition reduces to

$$\mathbb{P}[I^i=1\ \&\ I^j=1]\le\bar{\pi}^2,\quad \text{for all } i\neq j.$$

The SLLN will also hold if $\{I^i\}$ are positive quadrant dependent (PQD), i.e.,

$$\mathbb{P}[I^i=1\ \&\ I^j=1]\ge\bar{\pi}^2,\quad \text{for all } i\neq j,$$

provided that, for each $C\in\mathbb{N}$, the covariance between the ability assignment to the C competitors is not too strongly correlated with A_C, the number of average competitors.[26]

C.2 Effect of aggregate uncertainty on Π_A

Define

$$u^o(a,s) = a\,u(s)^{C-a}\,s^{a-1}, \quad a \in (0,C), \text{ and } s \in (0,1),$$

where u is defined by equation (2). We first show that

Result C-2. $a \hookrightarrow u^o(a,s)$ is strictly convex.

Proof. Explicit differentiation shows that

$$\frac{\partial^2 u^o(a,s)}{\partial a^2} = s^{a-1} \log\left[\frac{s}{u(s)}\right] \left(2 + a \log\left[\frac{s}{u(s)}\right]\right) u(s)^{C-a}. \tag{C-54}$$

Because F_T strictly first-order stochastically dominates F_A, $u(s) < s$ for $s \in (0,1)$. Thus, $\log[s/u(s)] > 0$. This implies that the right-hand side of equation (C-54) is positive and thus, $s \hookrightarrow u^o(s,a)$ is strictly convex. □

Next, note that

$$\lim_{a\to 0} u^o(a,s) = 0 \quad \text{and} \lim_{a\to C} u^o(a,s) = C\,s^{C-1}, \quad s \in (0,1). \tag{C-55}$$

Now, for $a \in (0,C)$, define the function, U^o, by

$$U^o(a) = \int_0^1 u^o(a,s)\,ds.$$

Result C-2 and the definition of U^o imply that U^o is a mixture of strictly convex functions, and thus U^o is strictly convex over $(0,C)$. Extend the definition of U^o to the endpoints of its interval of definition using equation (C-55) and, with a slight abuse of notation, also represent this extension with U^o. Thus,

$$U^o(a) = \begin{cases} 0 & a = 0 \\ \int_0^1 u^o(a,s)\,ds & a \in (0,C) \\ 1 & a = C \end{cases}. \tag{C-56}$$

After this extension, we have that

Result C-3. $a \hookrightarrow U^o(a)$, given by equation (C-56), is a strictly convex function defined over $[0,C]$.

[26] See Birkel (1988) for a precise statement of the condition.

Next, note that a change of variables shows that, for $a \in \{1,2,\dots C-1\}$,

$$\int_0^1 \left(u(s^{1/a})\right)^{C-a} ds = \int_0^1 a\,u(s)^{C-a} s^{a-1}\,ds = U^o(a). \tag{C-57}$$

For $a \in \{1,2,\dots C-1\}$, the probability that an average competitor submits the best performance given that there are a average competitors and $C-a$ talented competitors is represented by the left-hand side of equation (C-57). If there are no average competitors, i.e. $a = 0$, then the probability that an average competitor will submit the best performance is zero; while if all competitors are average, $a = C$, the probability that an average competitor will submit the best performance is 1. By definition, $U^o(0) = 0$ and $U^o(C) = 1$. Thus, if $\Pi_A(a)$ represents the probability that the best performer is average when there are a average competitors,

$$\Pi_A(a) = U^o(a), \quad a = \{0,1,\dots,C\}. \tag{C-58}$$

Suppose that $\tilde{a}$ is any non-degenerate random assignment of types to the C competitors and that $\mathbb{E}[\tilde{a}] = a^o$. The probability that the best performer is average under this random assignment is given by $\mathbb{E}[\Pi_A(\tilde{a})]$. Equation (C-58), Result C-3, and Jensen's inequality thus imply that

$$\mathbb{E}[\Pi_A(\tilde{a})] = \mathbb{E}[U^o(\tilde{a})] > U^o(a^o) = \Pi_A(a^o).$$

In summary, introducing aggregate uncertainty increases the probability that the best performer is average when the number of competitors is finite.

C.3 Effect of aggregate uncertainty on direction of approach

Proposition 1 identifies direction from which $\lim_{n\to\infty} \Pi_A(n)$ is approached. This proposition is derived under the assumption of no aggregate uncertainty, i.e., of a fixed ratio between the number of talented competitors and the number of average competitors. Relaxing the no aggregate uncertainty assumption requires considering talent assignments in which the number of average competitors is only fixed in expectation.

Relaxing the assumption of no aggregate uncertainty can change the conclusion of Proposition 1. A complete analysis of the effects of aggregate uncertainty is beyond the scope of this paper. Our focus is on large competitions and, as we shall see, the effects of aggregate uncertainty are strongest in very small contests. Moreover, the effects are complex and depend in subtle ways on the specific talent assignment mechanism posited. Thus, we only formally model an alternative scenario in which competitor ability is identically and independently distributed. Under this assumption, we show that, when F_T MLR dominates F_A, increasing competition always reduces Π_A, whereas when F_T first-order stochastically dominates F_A but does not MLR dominate F_A, the effect of increased competition, i.e. increasing C, on Π_A can be either negative or positive. We also consider, through a numerical example, the intermediate degrees of aggregate uncertainty which lie between independent random assignment and the fixed-ratio assignment used in the baseline model.

F_T MLR dominates F_A and talent assignments are independent. As the next result shows, when the talented and average competitors' performance distributions are MLR ordered, increasing the number of competitors always reduces the probability that the best performer is average.

Result C-4. Suppose that whether a competitor is talented or average is determined by an independent draw from a Bernoulli distribution that assigns probability $\theta \in (0,1)$ to A and probability $1-\theta$ to T. Then increasing the number of competitors always strictly reduces the probability that the best performer is average when F_T MLR dominates F_A.

Proof. Let $X_1,\ldots,X_C$ be i.i.d. random performance such that, for each $i=1,\ldots,C$, $X_i \overset{\text{dist}}{=} (1-\theta)X_T+\theta X_A$, where $X_T \overset{d}{\sim} F_T$ and $X_A \overset{d}{\sim} F_A$. Let t_i be the type of agent i. Let $b(C) = \arg\max_{1\le i\le C} X_i$ be the index of the best performer. To establish the result, it suffices to show that $\mathbb{P}[t_{b(C+1)}=A] < \mathbb{P}[t_{b(C)}=A]$.

For any fixed vector $(x_1,\ldots,x_C)\in(\ell,h)^C$, by Bayes rule,

$$\mathbb{P}[t_{b(C)}=A|X_1=x_1,\ldots,X_C=x_C] = \frac{\theta}{\theta+(1-\theta)\,(f_T(x_{b(C)})/f_A(x_{b(C)}))}. \tag{C-59}$$

When $X_{C+1}\le x_{b(C)}$, $b(C+1)=b(C)$. Thus,

$$\mathbb{P}[t_{b(C+1)}=A|X_1=x_1,\ldots,X_C=x_C,X_{C+1}\le x_{b(C)}] = \mathbb{P}[t_{b(C)}=A|X_1=x_1,\ldots,X_C=x_C]. \tag{C-60}$$

When $X_{C+1}>x_{b(C)}$, $b(C+1)=C+1$. Thus,

$$\begin{aligned}
&\mathbb{P}[t_{b(C+1)}=A|X_1=x_1,\ldots,X_C=x_C,X_{C+1}>x_{b(C)}]\\
&\quad=\mathbb{P}[t_{b(C+1)}=A|X_{C+1}>x_{b(C)}] = \mathbb{E}\left[\frac{\theta}{\theta+(1-\theta)\,(f_T(X_{C+1})/f_A(X_{C+1}))}|X_{C+1}>x_{b(C)}\right]\\
&\qquad< \frac{\theta}{\theta+(1-\theta)\,(f_T(x_{b(C)})/f_A(x_{b(C)}))} = \mathbb{P}[t_{b(C)}=A|X_1=x_1,\ldots,X_C=x_C],
\end{aligned} \tag{C-61}$$

where the inequality in the last line follows from the hypothesis that F_T MLR dominates F_A and the fact that $x\hookrightarrow \theta/(\theta+(1-\theta)x)$, $x>0$, is decreasing, and the last equality follows from equation (C-59). Equations (C-60) and (C-61), combined with the fact that $X_{C+1}>x_{b(C)}$ with a strictly positive probability, imply that $\mathbb{P}[t_{b(C+1)}=A|X_1=x_1,\ldots,X_C=x_C] < \mathbb{P}[t_{b(C)}=A|X_1=$

$x_1,\ldots,X_C = x_C]$. Integrating over $(x_1,\ldots,x_C)$ with the corresponding probability density shows that $\mathbb{P}[t_{b(C+1)} = A] < \mathbb{P}[t_{b(C)} = A]$. □

F_T MLR dominates F_A but talent assignments are imperfectly correlated. The effect of introducing aggregate uncertainty with respect to type (average/talented) assignment on the relationship between the probability that the best performer is average, Π_A, and the number of competitors, C, is twofold. First, aggregate uncertainty makes it possible that the best performer is average simply because there are no talented competitors in the competitor pool. This *talentless-pool effect* increases Π_A and is strongest when the number of competitors is small. This effect rapidly diminishes as the number of competitors increases. Thus, the talentless-pool effect militates in favor of a decreasing relationship between Π_A and C.

Second, aggregate uncertainty tends to neutralize the differential-threshold effect. Recall that, in Section 2, we argued that the reason why, absent aggregate uncertainty, increasing competition, by pushing up the threshold for best performance, need not favor the talented (even under MLR dominance) is because the threshold is random and the distribution of the random threshold is different for the two types of competitors; consequently, the extent to which the random threshold is pushed up through increased competition is different for the two types. Introducing aggregate uncertainty reduces the asymmetry in the distribution of the random threshold between the two types. When assignments are independent, for any given competitor, the ability distribution of her rivals will be independent of her ability type. In this case, the distribution of the random threshold will be the same for each competitor and, consequently, under MLR dominance, the differential-threshold effect will be completely neutralized. Thus, as shown by Result C-4, when assignments are independent and F_T MLR dominates F_A, increasing competition always favors the talented.

However, when assignments are dependent, for any given competitor, the ability distribution of her rivals will depend on her ability type, in which case introducing aggregate uncertainty can never completely neutralize the differential-threshold effect. Consequently, MLR dominance is no longer sufficient to ensure a decreasing relation between Π_A and C. If the differential-threshold effect favors an increasing relationship between Π_A and C, it can counter the talentless-pool effect produced by aggregate uncertainty over a range of competition sizes and thus produce a non-monotone relationship between Π_A and C.

To illustrate these possibilities, consider the distributions analyzed in Example 2, whose distribution functions are given by equation (7). As pointed out in the discussion of the example, and proven in Appendix D.1, F_T MLR dominates F_A. As we showed in the example, the two distribution functions given by equation (7) satisfy the decreasing reversed hazard-rate ratio condition, in which case the differential-threshold effect favors an increasing relationship between Π_A and C. Using equation (7), we can compute the u function, defined by equation (2) as follows:

$$u(s) = \exp\left(-\frac{e\,(1-s)}{s}\right), \quad s \in (0,\,1). \tag{C-62}$$

When all competitors are average, the probability that the best performer is average is 1; when all competitors are talented, the probability that the best performer is average is 0. When there are a positive number of both talented and average competitors, the probability that the best performer is average given that the total number of competitors is C and the number of average competitors is a is given by equation (3). Thus, if $\pi_A(C,a)$ represents the probability that the best performer is average when there are C total competitors and a average competitors, we have

$$\pi_A(C,a) = \begin{cases} 0 & a = 0 \\ \int_0^1 \left(u(s^{1/a})\right)^{C-a} ds & a \in \{1,2,\ldots C-1\} \\ 1 & a = C \end{cases}. \tag{C-63}$$

Integration shows that, for $a \in \{1,2,\ldots,C-1\}$, when u is given by equation (C-62),

$$\int_0^1 u(s^{1/a})^{C-a} ds = a\,e^{e\,(C-a)}\,(C-a)^a\,e^a\,\Gamma\left(-a,(C-a)\,e\right),$$

where Γ represents the upper incomplete Gamma function. Thus, by equation (C-63),

$$\pi_A(C,a) = \begin{cases} 0 & a = 0 \\ a\,e^{e\,(C-a)}\,(C-a)^a\,e^a\,\Gamma\left(-a,(C-a)\,e\right) & a \in \{1,2,\ldots,C-1\} \\ 1 & a = C \end{cases}. \tag{C-64}$$

The probability that the best performer is average, Π_A, is thus given by

$$\Pi_A(C) = \sum_{a=0}^{C} \pi_A(C,a)\,\mathbb{P}[\tilde{a} = a],$$

where π_A is given by equation (C-64). We model imperfect dependence by assuming that the probability law $a \hookrightarrow \mathbb{P}[\tilde{a} = a]$ is Hypergeometric$[C, C+2\,\kappa, C/2+\kappa]$, C even. This assumption is equivalent to assuming that the competitors are drawn without replacement from a population of potential competitors numbering $C+2\,\kappa$ consisting of $C/2+\kappa$ average competitors and $C/2+\kappa$ talented competitors. If $\kappa = 0$, this talent assignment mechanism is a no-aggregate uncertainty mechanism that assigns talent to exactly half of the competitors. As $\kappa \to \infty$, the Hypergeometric assignment mechanism converges to an independent random assignment. Table 2 presents the probability that the best performer is average, Π_A, for $C = 2,4,6,\ldots 50$ competitors, under no aggregate uncertainty, independent random assignment, and the Hypergeometric assignment (with various choices of the κ parameter).

As can be seen by inspecting Table 2, under the Hypergeometric imperfect dependence assignments, Π_A is initially higher than its limiting value $1/(1+e) \approx 0.269$ (computed in Example 2), falls below its limiting value, and then ultimately increases toward its limiting value. Thus, consistent with the no-aggregate uncertainty case analyzed in Example 2, under the Hypergeometric assignments, Π_A converges to its limiting value from below. However, in contrast to the no-aggregate uncertainty case analyzed in Example 2, convergence is not monotone.

Because the six assignments considered are ordered by second-order stochastic dominance (SSD), i.e., for fixed C, each assignment SSD dominates all assignments to its right, and because, when $a \in \{0, 1, \ldots C\}$, U^o, defined in Result C-3, equals the probability that the best performer is average for any fixed integer $C \geq 2$, Result C-3 implies that Π_A is increasing as one moves from left to right across the table. Observation shows that the results in Table 2 are consistent with Result C-3, i.e., increased aggregate uncertainty favors average competitors. The fact that, under all assignments that feature aggregate uncertainty, Π_A drops very sharply when C increases from 2 to 4, illustrates that the force of the talentless-pool effect produced by aggregate uncertainty diminishes rapidly as the number of competitors increases.

Because the limiting value of Π_A is invariant to assignment mechanisms that satisfy the law of large numbers (Appendix C.1), the differential-threshold effect, which, in this example, favors average competitors, also attenuates as the number of competitors increases, but this effect attenuates more slowly. Thus, under the imperfect dependence Hypergeometric assignment mechanisms, eventually, sooner when dependence is high, $\kappa = 1$, and later when dependence is weaker, $\kappa = 7$, the differential-threshold effect wins out, leading to a reversal in the monotonicity of the Π_A and C relationship.

	Π_A					
		Hypergeometric				
C	NAU	$\kappa = 1$	$\kappa = 3$	$\kappa = 5$	$\kappa = 7$	I
2	0.2283	0.3189	0.3448	0.3518	0.3551	0.3642
4	0.2461	0.2646	0.2802	0.2865	0.2898	0.3009
6	0.2530	**0.2616**	0.2702	0.2745	0.2770	0.2867
8	0.2567	**0.2617**	0.2674	0.2705	0.2725	0.2812
10	0.2590	0.2623	0.2663	0.2688	0.2704	0.2782
12	0.2606	0.2629	0.2659	0.2679	0.2692	0.2764
14	0.2617	0.2634	0.2658	0.2674	0.2685	0.2752
16	0.2626	0.2639	**0.2658**	0.2672	0.2681	0.2743
18	0.2633	0.2643	**0.2659**	0.2670	0.2679	0.2737
20	0.2638	0.2647	0.2660	0.2670	0.2677	0.2731
22	0.2643	0.2650	0.2661	0.2669	0.2676	0.2727
24	0.2646	0.2653	0.2662	**0.2669**	0.2675	0.2724
26	0.2650	0.2655	0.2663	**0.2670**	0.2675	0.2721
28	0.2652	0.2657	0.266	0.2670	0.2675	0.2719
30	0.2655	0.2659	0.2665	0.2670	0.2675	0.2716
32	0.2657	0.2660	0.2666	0.2671	0.2675	0.2715
34	0.2659	0.2662	0.2667	0.2671	0.2675	0.2713
36	0.2661	0.2663	0.2668	0.2672	0.2675	0.2712
38	0.2662	0.2664	0.2669	0.2672	0.2675	0.2710
40	0.2663	0.2666	0.2669	0.2673	**0.2675**	0.2709
42	0.2665	0.2667	0.2670	0.2673	**0.2676**	0.2708
44	0.2666	0.2668	0.2671	0.2674	0.2676	0.2707
46	0.2667	0.2668	0.2671	0.2674	0.2676	0.2707
48	0.2668	0.2669	0.2672	0.2674	0.2676	0.2706
50	0.2669	0.2670	0.2672	0.2675	0.2677	0.2705

Table 2: In all cases considered in the table, one-half of the total number of competitors, C, are expected to be talented. "NAU" denotes a no-aggregate uncertainty assignment in which exactly half of the competitors are talented. "I" denotes the talent assignment in which each competitor's talent results from an independent Bernoulli draw which assigns equal probability to the competitor being talented or average. In the "Hypergeometric" columns, competitors are randomly sampled without replacement from a pool of $C+2\kappa$ potential competitors which contains, $C/2+\kappa$ talented competitors and $C/2+\kappa$ average competitors. Numbers in the table represent the probability that the best performer is average, Π_A. The probabilities at which the direction of monotonicity of $C \hookrightarrow \Pi_A(C)$ changes are highlighted by bold-face type.

F_T first-order stochastically dominates F_A but does not MLR dominate F_A. Absent the assumption that the talented and average competitors' performance distributions are MLR ordered, the relationship between the number of competitors and the probability that the best performer is average need not be monotone even under independent random assignment. Intuitively, absent MLR dominance, even if the threshold for best performance were fixed, increasing the threshold need not favor the talented. Thus, the game-elevation effect alone might favor the average and dominate the talentless-pool effect produced by independent random assignment. For example, consider the following performance distributions:

$$F_T(x) = \begin{cases} 0 & x \le 0 \\ \exp\left(-\frac{1}{x^2}\right) & x > 0 \end{cases}, \quad F_A(x) = \begin{cases} 0 & x \le 0 \\ \frac{x}{\sqrt{1+x^2}} & x > 0 \end{cases}.$$

The associated quantile functions are given by

$$F_T^{-1}(s) = \frac{1}{\sqrt{-\log(s)}}, \quad F_A^{-1}(s) = \frac{s}{\sqrt{1-s^2}}, \quad s \in (0,1).$$

Using the definition of u provided by equation (2), we see that

$$\begin{aligned} u(s) = F_T \circ F_A^{-1}(s) = \exp\left(-\frac{1-s^2}{s^2}\right), \\ u^{-1}(s) = \frac{1}{\sqrt{1-\log(s)}}. \end{aligned} \tag{C-65}$$

First, note that

Result C-5. F_T strictly first-order stochastically dominates F_A.

Proof. Strict stochastic dominance is equivalent to $u(s) < s$ for all $s \in (0,1)$, which is equivalent to $u^{-1}(s) > s$, for all $s \in (0,1)$. Note that the definition of u^{-1} provided by equation (C-65) implies

$$u^{-1}(s) > s \Longleftrightarrow \frac{1}{s\sqrt{1-\log(s)}} > 1, \quad s \in (0,1).$$

Because $s \hookrightarrow s\sqrt{1-\log(s)}$ is concave, $s \hookrightarrow (s\sqrt{1-\log(s)})^{-1}$ is quasi convex. Thus, if $s \hookrightarrow ((s\sqrt{1-\log(s)})^{-1})' < 0$ as $s \to 1$, $s \hookrightarrow (s\sqrt{1-\log(s)})^{-1}$ is decreasing. $((s\sqrt{1-\log(s)})^{-1})' \to -1/2$ as $s \to 1$. Thus, $s \hookrightarrow (s\sqrt{1-\log(s)})^{-1}$ is decreasing. $\lim_{s\to 1}(s\sqrt{1-\log(s)})^{-1} = 1$, so $(s\sqrt{1-\log(s)})^{-1} > 1$, $s \in (0,1)$. □

Next, note that

Result C-6. F_T does not MLR dominate F_A.

Proof. MLR dominance is equivalent to u being convex. Using equation (C-65), we see that

$$\lim_{s\to 1} u''(s) = \lim_{s\to 1} \frac{2\,\exp\left(-\frac{1-s^2}{s^2}\right)\left(2-3\,s^2\right)}{s^6} = -2.$$

Thus, u is not convex. □

Integration shows that for $a \in \{1,2,\ldots,C-1\}$, when u is given by equation (C-65),

$$\int_0^1 \left(u(s^{1/a})\right)^{C-a} ds = \tfrac{1}{2}\,a\,e^{c-a}\,(c-a)^{a/2}\,\Gamma\left(-\frac{a}{2}, c-a\right),$$

where Γ represents the upper incomplete Gamma function. Thus, by equation (C-63), the probability that an average competitor will submit the best performance when there are C total competitors and a average competitors, $\pi_A(C,a)$, is given by

$$\pi_A(C,a) = \begin{cases} 0 & a = 0 \\ \frac{1}{2}\,a\,e^{c-a}\,(c-a)^{a/2}\,\Gamma\left(-\frac{a}{2}, c-a\right) & a \in \{1,2,\ldots C-1\} \\ 1 & a = C \end{cases}. \tag{C-66}$$

Now compare the probability that the best performer is average under type assignments that fix the number of average competitors at half the total number of competitors with a type assignment mechanism under which each competitor's type is determined by an independent draw from a Bernoulli distribution that assigns talent to a competitor with probability $1/2$. Let $\Pi_A(C)$ represent the probability that the best performer is average under fixed assignment when the total number of competitors is C, and let $\Pi_A^o(C)$ represent the probability that the best performer is average under independent random assignment when the total number of competitors is C. Using equation (C-66), we can compute these probabilities as follows:

$$\Pi_A(C) = \pi_A(C, C/2),$$
$$\Pi_A^o(C) = 2^{-C} \sum_{a=0}^{C} \binom{C}{a} \pi_A(C,a).$$

Figure 3 presents the probability that the best performer is average under fixed assignment, Π_A, and independent random assignment, Π_A^o, for $C = 2,4,\ldots 50$. We see that when the number of competitors is very small (i.e., when $C = 2$), Π_A^o is much larger than Π_A due to the talentless-pool effect produced by independent random assignment. When C increases to 4, and then to 6, the probability that there are no talented competitors in the competitor pool under independent random assignment reduces rapidly, and Π_A^o decreases. Once C reaches 6, the talentless-pool effect of independent random assignment is not sufficiently powerful to be the dominant effect.

Because, in this example, the likelihood ratio, f_T/f_A, is decreasing when x is sufficiently large, the game-elevation effect favors average competitors when the number of competitors is sufficiently large. Thus, once C reaches 6, Π_A^o starts to slowly increase. In the limit, Π_A^o and Π_A converge to the same limiting value, provided by Theorem 1.

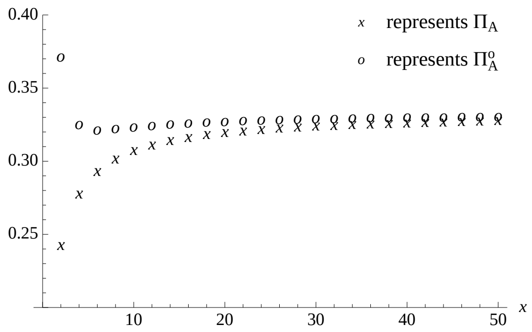


Figure 3: In the figure, the "x" markers represent the probability that the best performer is average under the no-aggregate uncertainty specification used in the baseline model, Π_A. The "o" markers represent the probability that the best performer is average under independent random assignment, Π_A^o. Under both specifications, the probability that a competitor is average equals 1/2.

D Miscellaneous Results

D.1 Proof of MLR dominance for Example 2

Recall the definitions of the talented and average competitors' performance distributions provided in Example 2. Define $\psi : [0,1] \to [0,1]$ and its inverse function as follows:

$$\psi(s) = \begin{cases} \frac{e}{e-\log(s)} & s \in (0,1] \\ 0 & s = 0 \end{cases}, \quad \psi^{-1}(s) = \begin{cases} \exp\left[-\frac{e(1-s)}{s}\right] & s \in (0,1] \\ 0 & s = 0 \end{cases}.$$

Next, note that $F_A = \psi \circ F_T$, and thus $F_A^{-1} = F_T^{-1} \circ \psi^{-1}$. Hence, $u(s) \equiv F_T \circ F_A^{-1}(s) = \psi^{-1}(s)$. The inverse function theorem and the definition of u imply that

$$u'(s) = \frac{f_T(F_A^{-1}(s))}{f_A(F_A^{-1}(s))}.$$

Thus, because $F_A^{-1}(s)$ is increasing in s, to establish MLR dominance (i.e., f_T/f_A increasing), it suffices to show that $u'' > 0$. As shown above, $u = \psi^{-1}$. Differentiation shows that

$$u''(s) = (\phi^{-1})''(s) = \frac{1}{s^4}(e-2s)\exp\left[1 - \frac{e(1-s)}{s}\right].$$

Because $s \leq 1$, $e - 2s > 0$ and hence $u'' > 0$.

D.2 When the target-hitting distribution is log-concave, the condition in Proposition 3.iii is never satisfied unless the hazard rate function is a constant (and *a fortiori* geometrically linear)

Proposition 3.iii shows that in target-hitting contests, $n \hookrightarrow \Pi_A(n)$ is nonincreasing (decreasing) if the hazard rate function, $h = f/(1-F)$, of the target-hitting distribution, F, is (strictly) geometrically concave over $(0, \infty)$. In what follows, we show that, if F is log-concave, then over $(0, \infty)$, h cannot be geometrically concave without being a constant (and *a fortiori* geometrically linear).

First, consider the case in which there exists $x_o > 0$ such that $h'(x_o) > 0$. Because the target-hitting distribution is unimodal and symmetric about 0, it must be that $F(0) = 1/2$, $f(0) > 0$, and $f'(0) \leq 0$. Thus, given that $h = f/(1-F)$ and $h' = (f'(1-F) + f^2)/(1-F)^2$, it must be that $\lim_{x \to 0^+} h(x) > 0$ and $\lim_{x \to 0^+} h'(x) < \infty$, which implies that $\lim_{x \to 0^+} xh'(x)/h(x) = 0$. Hence, given that $h'(x_o) > 0$, $h(x_o) > 0$, and $x_o > 0$, it must be that $\lim_{x \to 0^+} xh'(x)/h(x) = 0 < x_o h'(x_o)/h(x_o)$. Thus, $xh'(x)/h(x)$ cannot be nonincreasing over the whole positive real line, i.e., h cannot be geometrically concave over $(0, \infty)$.

Next, consider the case in which $h'(x) \leq 0$ for all $x > 0$ and $h'(x) < 0$ for some $x > 0$ (otherwise, if $h'(x) = 0$ for all $x > 0$, h would be a constant over $(0, \infty)$). Because the target-hitting distribution, F, is symmetric about 0, $h(x) = f(x)/(1-F(x)) = f(-x)/F(-x) = r(-x)$, where $r = f/F$. Thus, $h'(x) = -r'(-x)$. Hence, given that $h'(x) < 0$ for some $x > 0$, it must be that $r'(x) > 0$ for some $x < 0$. Thus, given that $r' = (\log \circ F)''$, $(\log \circ F(x))'' > 0$ for some $x < 0$, in which case F cannot be log-concave.