IMF Working Paper

The Limits of Meritocracy

by John Morgan, Justin Tumlinson, and Felix Várdy

INTERNATIONAL MONETARY FUND

IMF Working Paper

Monetary and Capital Markets Department

The Limits of Meritocracy[1]

Prepared by John Morgan, Justin Tumlinson, and Felix Vardy

Authorized for distribution by Nigel Jenkinson

November 2018

Abstract

We show that too much meritocracy, modeled as accuracy of performance ranking in contests, can be a bad thing: in contests with homogeneous agents, it reduces output and is Pareto inefficient. In contests with sufficiently heterogeneous agents, discouragement and complacency effects further reduce the benefits of meritocracy. Perfect meritocracy may be optimal only for intermediate levels of heterogeneity.

Authors' E-Mail Address: morgan@berkeley.edu, J.Tumlinson@lboro.ac.uk, fvardy@imf.org

Contents

Figures

The race is not to the swift, nor the battle to the strong, neither yet bread to the wise, nor yet riches to men of understanding, nor yet favor to men of skill: but time and chance happeneth to them all.

Solomon, Ecclesiastes 9:11

I.   Introduction

If secular society has a creed, it is the creed of meritocracy—the unwavering belief that, in the rat race of life, the rewards should go to the best performers. Indeed, while many question how meritocratic our society really is, few challenge the ideal of meritocracy itself. Rather, the merits of meritocracy are thought to be contingent on overcoming the inequality of opportunities and creating a level playing field.

In every day life, the 'rat race' may take many different forms: employees vie for promotions and bonuses, students compete for 'Firsts,' scientists contend for top journals' scarce pages, and so on. What unifies these disparate competitions is the importance of relative—rather than absolute—performance and the often starkly different rewards for winners and losers. In the parlance of economics, they are all *contests*.

The case for meritocracy in contests may seem self-evident: Consider, instead, an entirely *a-meritocratic* contest, where winners are determined at random. Not only would most people consider such a contest unfair, it would also elicit little effort from contestants, since exertion would not improve their chances of winning. Hence, (the perception of) meritocracy, in the sense of accurate performance ranking, is essential, both for moral and for practical reasons.

Imperfections in the fairness of the playing field, measurement error, and the cost of accurately assessing performance still hamper achieving our ideal of perfect meritocracy. Yet, the rise in Big Data and the power of the internet are bringing the goal closer. Nowadays, managers can collect micro-activity data on employees, facilitating more accurate performance measurement (Cowgill, 2015, Lohr, 2014). Educational assessments are becoming more comparable, e.g., via standardized testing and the Bologna Process. And crowd-based peer review is beginning to reduce the randomness of traditional peer review in some academic fields (List, 2017). Meritocracy may finally be within reach.

This raises an important question, however: if meritocracy is good, is more meritocracy necessarily better? Notwithstanding today's "Cult of Meritocracy" (Scaggs, 2018, Frank, 2016, Hayes, 2012, Friedersdorf, 2012), in this paper we show that the answer is *no*: too much meritocracy, in the sense of accuracy of performance ranking, can reduce aggregate output and be Pareto inefficient.

Before presenting our argument, let us fix some terminology. While an
a-meritocratic contest selects winners at random, in a *perfectly* meritocratic contest
the best performers win with certainty. We refer to anything in between as *partially*
meritocratic. One could, of course, move even further away from meritocracy than
'a-meritocratic,' by introducing nepotism and the like. We do not consider such
un-meritocratic contests in this paper.

Arguably, "it is not meritocracy until everyone starts with the same opportunities"
(Wade, 2010). So, to give meritocracy its best shot, we first consider a setting where
inequality of opportunities plays no role. That is, perhaps due to preselection or
remedial action, contestants are homogeneous and, thus, face a perfectly level
playing field. For increasing marginal costs, we show that perfect meritocracy is
always harmful: it hurts the contest organizer by reducing aggregate output,
without benefiting contestants. This holds even though, in our model, perfectly
accurate performance ranking is costless.

The intuition is as follows. When performance ranking is pure noise, nobody does
anything. Hence, injecting some meritocracy into a contest unambiguously raises
output. However, at a critical point, competition becomes so intense that
contestants start 'dropping out,' i.e., they put in zero effort. Crucially, this always
happens before noise is fully eliminated. Without loss of generality, we may assume
that drop-outs earn zero payoff. Homogeneity implies that those who continue to
participate also earn zero in expectation. Thus, from the critical level onward, the
contest organizer captures all the rents and should focus on making the pie as large
as possible.

For strictly increasing marginal costs, Jensen's inequality implies that the output
loss from contestants dropping out exceeds the output gain from the remaining
contestants working harder. Hence, the optimal level of meritocracy equals the
critical level, while perfect meritocracy reduces aggregate output and is Pareto
inefficient. For strictly *de*creasing marginal costs, the opposite result obtains: now
the output gain exceeds the output loss, such that perfect meritocracy is uniquely
output maximizing, as well as Pareto efficient.[2] Finally, for constant marginal costs,
output losses and output gains exactly offset each other. This gives rise to a
threshold result: output is maximized—and production is Pareto efficient—for any
level of meritocracy greater than the critical level.

Focusing on the common case of non-decreasing marginal costs, next we study the
joint optimization problem across meritocracy and prize schedules when the contest
organizer disposes of a fixed prize budget. We first establish that, with
homogeneous contestants, the organizer can do no better than to offer prizes of
equal value. Then we show that within this class, *any* prize schedule can be
optimal, provided the level of meritocracy is judiciously chosen. Subject to a
technical condition, the reverse result also obtains: *any* level of meritocracy can be
optimal, provided we get to choose the number of prizes.

---

[2]For technical reasons, the decreasing marginal cost case is relegated to the Appendix.

The optimal level of meritocracy and the optimal number of prizes turn out to be positively related: the more prizes are on offer (and, thus, the smaller their value) the more meritocratic should be the contest. In fact, perfect meritocracy is optimal if and only if the contest has a "participation-prize" structure (i.e., almost everybody wins a prize). Intuitively, the only way to motivate contestants in an environment where almost all succeed is to make the performance ranking extremely accurate. At the other extreme, a winner-take-all structure is optimal if and only if the contest is almost entirely a-meritocratic.[3] In this case, the only way to motivate agents not to drop out, even though almost nobody succeeds, is to assign the prize virtually at random.[4]

Turning to (homogeneous-prize) contests with heterogeneous-ability agents, first we show that the homogeneous contest model is not a singularity: its results carry over to contests that are merely close to homogeneous. Second, for general levels of heterogeneity, we show that the effect of a rise in meritocracy on output can be partitioned into three, additively separable parts: an attrition effect, a competition effect, and a heterogeneity effect. The attrition effect refers to the output loss from agents dropping out. Therefore, it is either zero or negative. The competition effect refers to the output gain from the remaining contestants working harder on average. Therefore, it is strictly positive. Finally, the heterogeneity effect corrects for the fact that a rise in meritocracy affects contestants *differentially*, depending on their ability: while contestants "on the bubble" (i.e., highly uncertain about winning or losing) are spurred on by greater meritocracy, it discourages the weak and makes the strong complacent. The former are discouraged, because they know that they now stand even less of a chance than before. The latter become complacent, because they realize that outperforming their weaker brothers by just a tiny bit suffices to virtually guarantee themselves a prize.

We find that, in the aggregate, discouragement of the weak and complacency of the strong dominate the spurring-on of the middle, such that the heterogeneity effect is negative. The (rather intuitive) implication is that, as long as there is no attrition, heterogeneity *reduces* the benefits of meritocracy. Heterogeneity also gives meritocracy its best shot, however, because once it kicks in, attrition proceeds from the bottom of the ability distribution. This reduces the cost of agents dropping out when compared to the homogeneous case. As we show, it also makes perfect meritocracy potentially optimal—but only for intermediate levels of heterogeneity. For very high levels of heterogeneity we find that perfect meritocracy is once again strictly suboptimal, even for linear costs. This is due to the high loss of output that meritocracy-induced complacency engenders among the highest ability types.

Our analysis clarifies that whether meritocracy sharpens or softens competition

---

[3]To be precise, the "if" implication requires marginal costs to be *strictly* increasing. For constant marginal costs, the threshold result (together with the technical condition referred to earlier) imply that the winner-take-all contest is optimal for *all* levels of meritocracy.

[4]It is interesting to ponder what this might suggest about the meritocracy of rat races to become CEO of a large corporation.

depends on the heterogeneity of contestants: Under sufficient homogeneity, a rise in meritocracy pits contestants more harshly against each other and, thus, sharpens competition. Under sufficient heterogeneity, a rise in meritocracy discourages the weak and makes the strong complacent and, thus, softens competition. Nevertheless, perfect meritocracy reduces output in *both* cases: because of excessive competition in the sufficiently homogeneous case and deficient competition in the sufficiently heterogeneous case.

Historically, managers monitored performance by random inspection—so-called 'management-by-walking-around.' In many industries today, employee performance can be monitored electronically—most notably in the technology sector. Since promotions, bonuses, and other rewards tend to be awarded according to relative performance, one might expect firms to use this newly and readily available micro-data to improve the accuracy of performance rankings. However, while some firms indeed do so, others explicitly commit to not using the data in this way.[5] Our analysis suggests that it may depend, at least in part, on the heterogeneity of a firm's workforce. A firm that runs a rigorous hiring process will have a workforce that, at a given level, is relatively homogeneous. Since too much meritocracy is output-reducing in such an environment, the firm may be better off (committing to) ignoring the micro-data for purposes of performance ranking. By contrast, a firm whose workforce is more heterogeneous may want to make use of it.[6]

This argument complements other potential explanations from the principal-agent literature. For example, Holmstrom and Milgrom (1991) show that firms may not want make remuneration (too) dependent on measured performance, if doing so makes employees neglect other, less measurable aspects of their jobs. Similarly, Goodhart's Law says that "when a measure becomes a target, it ceases to be a good measure" (Strathern, 1997). For example, employees may start "juking the stats" and thus compromise the integrity of the data. Finally, Lazear (1989) argues that when cooperation among workers is important, we should also expect to see lower-powered incentives.

Our model also puts the Dutch university admissions system in a new light. When the number of applicants exceeds the number of available slots, students in the Netherlands are selected by means of a weighted lottery: the chance of admission is an increasing function of an applicant's grade-point average on the centralized high school exam. The stated motive for the scheme is one of fairness, in the sense that it provides all qualifying students with at least some chance of admission.[7] However, our model suggests that the system may also increase studiousness in high school,

---

[5]See Cowgill (2015) for a description of employee monitoring at a typical, large tech company, as well as for evidence that, sometimes, readily available electronic performance measures are not used. See Lohr (2014) for evidence that, at other times, they are used.

[6]See Morgan *et al.* (2018) for an analysis how meritocracy affects selection into competing contests.

[7]For a short history of the weighted lottery system, which has been in place in one form or another since 1975, as well as for discussions about reforming or abolishing it, see, e.g., KNAW (2013).

through less discouragement of, and less dropping out by, weaker students, and less complacency among stronger students.

The remainder of the paper is organized as follows. Section II. discusses the extant literature. Section III. introduces and analyzes a simple baseline model with a continuum of homogeneous contestants. In Section IV. we study the same model with heterogeneous contestants. Section V. considers finite-player contests à la Lazear and Rosen (1981) and Tullock (1980). Section VI. concludes. Most proofs have been relegated to an online Appendix**??**online Appendix .

II.   Literature Review

The modern literature on contests originated with the seminal papers of Lazear and Rosen (LR, 1981) and Tullock (1980). For a broad overview see, e.g., Konrad (2009) and Corchon (2007).

Within the branch that focuses on optimal contests, our paper is perhaps most closely related to Moldovanu and Sela (2001), Clark and Riis (1998), Siegel and Olszewski (2018, 2016), O'Keeffe *et al.* (1984), and Fu *et al.* (2015). In each of these papers the contest designer varies different levers to achieve his aims.

Moldovanu and Sela (2001) focus on the number of prizes. They show that in a perfectly meritocratic multi-prize contest with heterogeneous players, a winner-take-all structure maximizes output if the contestants have concave or linear costs, but when contestants have convex costs, multiple prizes may be optimal.

Clark and Riis (1998) analyze contests with multiple identical prizes and compare simultaneous versus sequential contest designs. They show that a sequential design produces higher output when contestants are relatively homogeneous, while a simultaneous design dominates when contestants are relatively heterogenous.

Siegel and Olszewski (2018, 2016) show how a standard mechanism-design approach with a continuum of types can be used to approximate contests with many players. They then apply their finding to derive output-maximizing prize schedules—an analysis that is related to the classic papers by Myerson (1981) and Guesnerie and Laffont (1984).

While the papers by Siegel and Olszewski's (2018, 2016) are exceptionally general, like Moldovanu and Sela (2001) and Clark and Riis (1998), they restrict attention to perfectly meritocratic contests. Hence, the particular consideration that motivates our work, namely, the optimal accuracy of performance ranking, is absent.

Our focus on imperfect meritocracy is in line with LR and Tullock (1980). While these seminal papers spawned a vast literature, there has been relatively little focus on *variations* in meritocracy. Some exceptions are Baye *et al.* (1994), Che and Gale

(2000), O'Keeffe *et al.* (1984), and Fu *et al.* (2015).

Baye *et al.* (1994) study highly meritocratic Tullock contests for which pure-strategy equilibria do not exist. They show that these contests do allow for symmetric mixed-strategy equilibria and they derive upper and lower bounds for expected rent dissipation. Che and Gale go through a similar exercise for linear "difference-form" contests. (See also Hirshleifer, 1989, and Baik, 1998.)

Most closely related to our paper are O'Keeffe *et al.* (1984) and Fu *et al.* (2015). O'Keeffe *et al.* study contest design and meritocracy in a 'competitive market' for homogeneous two-player contests. The assumption of perfect competition among contest organizers is taken to imply that all rents accrue to the contestants and that the contest organizer must pursue the *socially efficient* level of output. The main result of the paper is to show that efficiency can be implemented in pure strategy equilibrium by a range of price spreads and less than perfect meritocracy. This follows from equating contestants' FOCs to the social value of output and checking that, at the efficient level, the participation constraint is satisfied.[8]

Our paper is in the spirit of O'Keeffe *et al.* However, rather than focusing on perfect competition and efficiency, we study a monopolistic contest organizer who truly gets to maximize aggregate output. To do so, we first characterize contestants' equilibrium behavior for all levels of meritocracy. Then we show that, in the homogeneous case, output is maximized at the boundary between pure and mixed strategy equilibria. Hence, the organizer raises meritocracy up to the point where contestants start dropping out. Unlike efficiency, output maximization can thus be pursued even when the monetary value of output is unknown or cannot be agreed upon. This matters because a major justification for organizing a contest rather than paying a piece rate is precisely the inability to assign monetary value to a unit of output.

Finally, Fu *et al.* (2015) study a model of endogenous entry in nested lottery contests. They derive output-maximizing combinations of entry fees (or subsidies), prizes, and discriminatoriness. They also identify conditions under which the optimum can be achieved by solely adjusting the discriminatory power of a Tullock contest with a single prize. Focusing on monomial cost functions and symmetric equilibria, Fu *et al.* anticipate some of our findings for Tullock contests. In particular, they also find that the contest organizer may prefer a noisier winner selection mechanism. They relate this to a trade-off between, on the one hand, encouraging entry and, on the other, more competitive bidding for a given number of entrants. The former pushes towards more, while the latter pushes towards less

---

[8]Essentially, the argument of O'Keeffe *et al.* (1984) is as follows. Efficiency of equilibrium requires that the social marginal benefit of output equals marginal cost equals private marginal benefit. Since the social marginal benefit is finite, so must be the private marginal benefit. To see that the latter is inconsistent with perfect meritocracy, notice that, in that case, an infinitesimal unilateral deviation from the symmetric equilibrium would cause the probability of winning to jump discontinuously from $1/2$ to 1.

noise.

Our analysis shows that, even in the absence of entry costs, the contest organizer may prefer less than perfect discriminatoriness—and it clarifies that this solely depends on the convexity of the cost function. We also relax the symmetric equilibrium assumption.

III. Baseline Model

A. Set-Up

A unit mass of homogeneous, risk-neutral agents $i \in I \equiv [0,1]$ participate in a contest. Agent i's measured performance $Y_i \in [0,\infty)$ depends on his output $X_i \in [0,\infty)$ and i.i.d. noise $E_i \in (0,\infty)$. Specifically,

$$Y_i = X_i \cdot E_i$$

The cost of producing X_i is $C(X_i) \in [0,\infty)$. We assume that: 1) $C(\cdot)$ is analytic; 2) $C(0) = 0$; 3) $C'(\cdot) > 0$ for $X > 0$; and 4) $C''(\cdot) \geq 0$.

Let $\eta_{C''/C'}$ denote the elasticity of the cost function's Arrow-Pratt curvature, C''/C', with respect to X. Similarly, η_C and $\eta_{C'}$ denote the output elasticities of costs and marginal costs, respectively. For analytical convenience, we assume that $\eta_{C''/C'}$ does not fall "too fast." That is,

$$\eta_{C''/C'} \equiv \frac{\mathrm{d}\ln(C''/C')}{\mathrm{d}\ln X} \geq -1 \ . \tag{1}$$

(If $C''(X) = 0$, then $\eta_{C''/C'} \equiv 0$.) As we shall see, condition (1) implies that there exists at most one interior, payoff-maximizing output. In turn, this guarantees uniqueness of equilibrium. While convenient, in Appendix 2. we show that condition (1) is not essential for our argument.[9]

Usually, it will be more practical to work in logs, which we denote by lower case letters. Hence, the log of measured performance, y_i, is

$$y_i = x_i + \varepsilon_i \ ,$$

where $x_i \in [-\infty, \infty)$ and $\varepsilon_i \in \mathbb{R}$. We also define $c(x) \equiv C(e^x) = C(X)$. In our leading example, the elasticity of marginal costs is constant; i.e., the cost function is of the form $C(X) = kX^\alpha = ke^{\alpha x} = c(x)$, $k > 0$, $\alpha \geq 1$.

The following lemma translates our assumptions on $C(\cdot)$ into properties of $c(\cdot)$.

---

[9] All cost functions of the form $C(X) = kX^\alpha$, $k > 0, \alpha \geq 1$, satisfy (1)—as do all otherwise admissible polynomials of order 3 or less. On the other hand, even though it is strictly convex, the 4th-order polynomial $C(X) = 7X^2 - 3X^3 + 1/2 \cdot X^4$ violates (1) for $X \in \mathbb{R}_+ \setminus \left[\frac{1}{9}(28 - \sqrt{217}, \frac{1}{9}(28 + \sqrt{217}\right] \approx \mathbb{R}_+ \setminus [1.47, 4.74]$.

Lemma 1 *1)* $c(-\infty) = c'(-\infty) = 0$. *For* $x > -\infty$: *2)* $c'(x)$, $c''(x) > 0$; *3)* $c(x)$, $c'(x)$ *are weakly log-convex; 4)* $\eta_{C'} + 1 = c''(x)/c'(x) \geq c'(x)/c(x) = \eta_C \overset{(=)}{>} 1$, *where the last (in)equality depends on whether* $C''(\cdot) \overset{(=)}{>} 0$; *5) If* $C(X) = kX^\alpha$, $k > 0$, $\alpha \geq 1$, *then* $\eta_{C'} + 1 = \eta_C = \alpha$.

We assume that the distribution of ε_i belongs to a location-scale family with location parameter zero and scale parameter $\sigma > 0$. Precision $1/\sigma$ measures the accuracy of performance evaluation, which we interpret as meritocracy. Perfect meritocracy arises when $\sigma \to 0$. Noise ε_i admits a CDF $F\left(\frac{\varepsilon_i}{\sigma}\right)$ with associated PDF $\frac{1}{\sigma}f\left(\frac{\varepsilon_i}{\sigma}\right)$. Density $f(\cdot)$ is assumed to be continuous, differentiable on $\mathbb{R}$, strictly log-concave, and to have full support. Notice that strict log-concavity implies that f is single peaked.

The mass $m \in (0,1)$ of contestants with highest measured performances y receive a prize $v > 0$. Among a continuum of agents, the quantiles of measured performances are perfectly predictable. Hence, to win a prize, an agent must surpass a *deterministic* performance threshold, or 'standard,' $\theta \in \mathbb{R}$. If prizes need to be rationed among a mass of contestants with identical y, prizes are allocated at random. Unless more than m contestants "drop out" (i.e., choose $x_i = -\infty$), this almost surely does not happen, since ε_i is i.i.d. and atomless.

The contest organizer's objective is to maximize expected aggregate output, where aggregate output is given by

$$O \equiv \int_{i \in I} X_i \mathrm{d}i = \int_{i \in I} e^{x_i} \mathrm{d}i .$$

Initially, the organizer's only instrument is the degree of meritocracy, σ. Later we will also allow him to chose the prize structure—i.e., the number and value of prizes, (m, v), for a fixed budget $b = mv$.

B.   Equilibrium

An agent who produces log-output x_i when the standard is θ achieves an expected payoff

$$\pi(x_i, \theta) = v\bar{F}\left(\frac{\theta - x_i}{\sigma}\right) - c(x_i) . \tag{2}$$

Here, $\bar{F} \equiv 1 - F$ denotes the decumulative distribution function (DCDF) of ε_i/σ.

The first-order condition (FOC) for the agent's choice variable, x_i, to be an interior extremum requires marginal benefit to equal marginal cost (see Figure 1):

$$\frac{v}{\sigma}f\left(\frac{\theta - x_i}{\sigma}\right) = c'(x_i) . \tag{3}$$

The SOC for a maximum requires that $c''(x) > -v/\sigma^2 f'[(\theta - x)/\sigma]$—i.e., the marginal cost curve must be steeper than the marginal benefit curve, such that the former intersects the latter from below. By re-applying the FOC, the SOC can be written somewhat more elegantly as

$$soc \equiv - \left[\frac{\frac{1}{\sigma} f'\left(\frac{\theta - x}{\sigma}\right)}{f\left(\frac{\theta - x}{\sigma}\right)} + \frac{c''(x)}{c'(x)} \right] < 0 \; . \tag{4}$$

In the panels of Figure 1, the FOC and SOC are both satisfied at the point x_s. More generally, let x_s denote the largest intersection point between marginal costs and marginal benefits.

The participation constraint requires that, at an interior maximum, $\pi \geq 0$. Otherwise, the agent is better off dropping out and playing $x = -\infty$. This pays zero for all θ.[10] Dropping out is also optimal if the FOC has no solution since, in that case, marginal costs must dominate marginal benefits for all x.

Denote the (pure) best-response correspondence by $\hat{x}(\theta)$, i.e., $\hat{x}(\theta) \equiv \{x : \arg\max_{x \in [-\infty, \infty)} \pi(x; \theta)\}$. Since agents are homogeneous, $\hat{x}(\cdot)$ is the same for all agents. Slightly abusing notation, we treat the correspondence $\hat{x}(\cdot)$ as a function when it is single-valued. The next lemma shows that $\hat{x}(\theta)$ has a particularly simple structure.

Lemma 2 *For all θ, $\hat{x}(\theta) \subset \{-\infty, x_s(\theta)\}$.*

Lemma 2 says that, depending on the standard, agents either produce x_s or they drop out. In other words, Figure 1 is generic. To see why this is so, notice that the FOC in (3) can be rewritten as

$$\frac{\frac{v}{\sigma} f\left(\frac{\theta - x_i}{\sigma}\right)}{c'(x_i)} = 1 \; . \tag{5}$$

Together, strict log-concavity of f and weak log-convexity of $c'(x)$ (Lemma 1) imply that the LHS of (5) is strictly log-concave in x. Since log-concavity begets single-peakedness, we may conclude that marginal costs and marginal benefits have at most two points of intersection, only one of which can be a maximum. Next, observe that $\lim_{x \to \infty} \frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)/c'(x) = 0$. This means that, if the two curves intersect at all, the largest intersection point, x_s, must be a maximum, and since there is at most one interior maximum, it is the only one. Hence, optimal output equals either x_s or $-\infty$—or a mixture of the two. Notice that this result critically

---

[10]If more than m contestants drop out, arguably, the standard is $\theta = -\infty$ and some prizes are assigned at random among the drop-outs. Notice, however, that this can never happen in equilibrium: any $x > -\infty$, no matter how small, would guarantee a prize and, hence, constitute a profitable deviation. Therefore, we can safely restrict attention to $\theta \in \mathbb{R}$.

depends on our technical assumption (1) that $\eta_{C''/C'} \geq -1$, for it implies (in fact, is equivalent to) log-convexity of $c'(x)$. The case where this inequality fails to hold is analyzed in Appendix 2..

Figure 1. The marginal cost curve intersects the marginal benefit curve at most twice. If the curves intersect at all, the largest instersection, $x_s(\theta)$, constitutes the unique interior maximum. However, for θ sufficiently large, the participation constraint will be violated at $x_s(\theta)$, and the best response is to drop out of the contest altogether by playing $-\infty$.
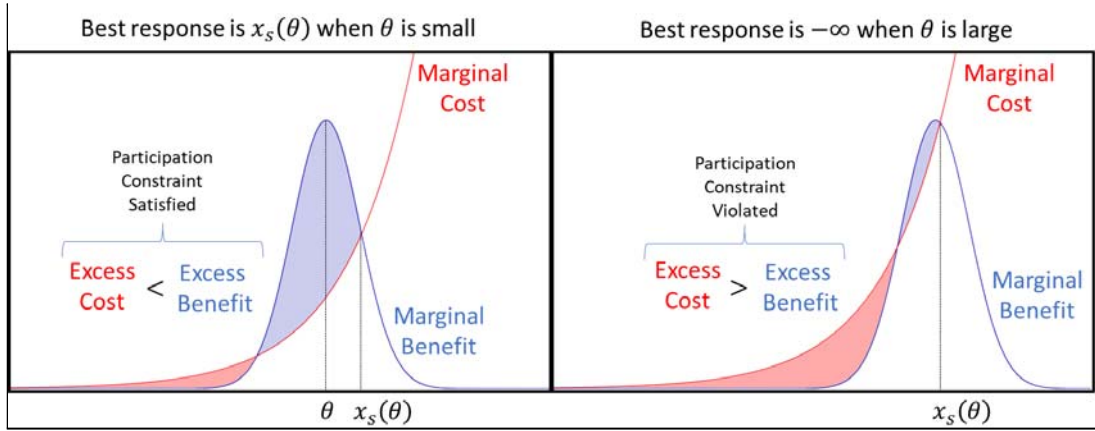
Since $\hat{x}(\theta) \subset \{-\infty, x_s(\theta)\}$, a mixed best response is fully characterized by the probability, $\hat{\gamma}_i(\theta)$, that the agent produces $x_s(\theta)$. (If $x_s(\theta)$ does not exist, $\hat{\gamma}_i(\theta) \equiv 0$.) Let $\{\hat{\gamma}_i\}_{i \in I}$ denote a profile of mixed best responses, one for each agent. Such a profile is symmetric if $\hat{\gamma}_i = \hat{\gamma}_j \equiv \hat{\gamma}$ for all $i, j \in I$.

Best-response profile $\{\hat{\gamma}_i\}_{i \in I}$ generates aggregate output O and mass of winners W, where

$$O\left(\{\hat{\gamma}_i\}_{i \in I}\right) = e^{x_s(\theta)} \int_{i \in I} \hat{\gamma}_i(\theta) \, \mathrm{d}i , \tag{6}$$

$$W\left(\theta, \{\hat{\gamma}_i\}_{i \in I}\right) = \bar{F}\left[\frac{\theta - x_s(\theta)}{\sigma}\right] \int_{i \in I} \hat{\gamma}_i(\theta) \, \mathrm{d}i . \tag{7}$$

A tuple $\{\theta, \{\hat{\gamma}_i\}_{i \in I}\}$ is an equilibrium if $W\left(\theta, \{\hat{\gamma}_i\}_{i \in I}\right) = m$. It is denoted by $\{\theta^*, \{\hat{\gamma}_i^*\}_{i \in I}\}$. An equilibrium is symmetric and denoted by $\{\theta^*, \hat{\gamma}^*\}$, if $\{\hat{\gamma}_i^*\}_{i \in I}$ is symmetric. Finally, a symmetric equilibrium is in pure strategies if $\hat{\gamma}^* = 1$. (Notice that $\hat{\gamma}^* = 0$ is incompatible with equilibrium.) A pure-strategy equilibrium is denoted by $\{\theta^*, x_s(\theta^*)\}$.

Since aggregate output is what we are after, we only need to distinguish between equilibria that generate different levels of O^*. The next lemma implies that, without loss of generality, we may restrict attention to symmetric equilibria.

Lemma 3 *For every equilibrium $\left\{\theta^*, \{\hat{\gamma}_i^*\}_{i \in I}\right\}$ there exists a symmetric equilibrium $\left\{\theta^*, \hat{\gamma}^*\right\}$ that generates the same aggregate output.*

The intuition for Lemma 3—which also constitutes the proof—is as follows. Let $\gamma \equiv \int_{i \in I} \hat{\gamma}_i^* \mathrm{d}i$. From (6) and (7) we know that, due to the atomicity of agents, a population playing $\{\hat{\gamma}_i^*\}_{i \in I}$ generates the same aggregate output and the same mass of winners as when everybody mixes between $x_s(\theta^*)$ and $-\infty$ with probabilities γ and $1 - \gamma$, respectively. Furthermore, since $\{\hat{\gamma}_i^*\}_{i \in I}$ constitutes a best-response profile and agents are homogeneous, γ is also a best response. Together, these observations imply that, if $\left\{\theta^*, \{\hat{\gamma}_i^*\}_{i \in I}\right\}$ is an equilibrium with output O^*, so is $\{\theta^*, \gamma\}$, while the latter is symmetric.

Equilibria form "output equivalence classes," one for each symmetric equilibrium. To see this, consider a strictly-mixed symmetric equilibrium $\{\theta^*, \hat{\gamma}^*\}$ and notice that any tuple $\left(\theta^*, \{\gamma_i\}_{i \in I}\right)$ such that $\int_{i \in I} \gamma_i \mathrm{d}i = \hat{\gamma}^*$ also constitutes an equilibrium and generates the same output.

Lemma 3 allows us to restrict attention to symmetric equilibria, $\{\theta^*, \hat{\gamma}^*\}$. These equilibria are characterized by the following system of FOC, SOC, participation constraint (PC), and market-clearing condition (MCC):

$$\text{FOC} \quad : \quad \frac{v}{\sigma} f\left[\frac{\theta - x_s(\theta)}{\sigma}\right] = c'\left[x_s(\theta)\right] \ ,$$

$$\text{SOC} \quad : \quad \frac{\frac{1}{\sigma} f'\left[\frac{\theta - x_s(\theta)}{\sigma}\right]}{f\left[\frac{\theta - x_s(\theta)}{\sigma}\right]} > \frac{c''\left[x_s(\theta)\right]}{c'\left[x_s(\theta)\right]} \ ,$$

$$\text{PC} \quad : \quad v\bar{F}\left[\frac{\theta - x_s(\theta)}{\sigma}\right] \left\{ \begin{array}{ll} \geq c\left[x_s(\theta)\right] & \text{if} \quad \hat{\gamma} = 1 \\ = c\left[x_s(\theta)\right] & \text{if} \quad \hat{\gamma} \in (0,1) \end{array} \right. \ ,$$

$$\text{MCC} \quad : \quad \hat{\gamma}\bar{F}\left[\frac{\theta - x_s(\theta)}{\sigma}\right] = m \ .$$

The next proposition establishes existence and uniqueness of (symmetric) equilibrium. It also characterizes equilibrium behavior as a function of σ.

Proposition 1 *Equilibrium exists and is unique. For σ large, equilibrium is in pure strategies, and all agents produce $x_s(\theta^*) > -\infty$, which is the unique interior maximizer. For σ small, agents mix between x_s and dropping out (i.e., $x = -\infty$).*

Formally,

$$\hat{x}(\theta_\sigma^*; \sigma) = \left\{ \begin{array}{ll} x_s(\theta_\sigma^*; \sigma) & \text{for} \quad \sigma > \underline{\sigma} \\ \{-\infty, x_s(\theta_\sigma^*; \sigma)\} & \text{for} \quad \sigma \leq \underline{\sigma} \end{array} \right. \ ,$$

where $\underline{\sigma} = vf\left[\bar{F}^{-1}(m)\right]/c'\left[c^{-1}(mv)\right] > 0$. Output $x_s(\theta^*_\sigma; \sigma)$ is played with probability $\hat{\gamma}^*$, where $\hat{\gamma}^* \overset{(=)}{\underset{<}{}} 1$ if $\sigma \overset{(\geq)}{\underset{<}{}} \underline{\sigma}$. Furthermore, $d\hat{\gamma}^*/d\sigma > 0$ and $\lim_{\sigma\to 0}\hat{\gamma}^* = m$.

Proposition 1 reflects that, for large σ, the intensity of competition is low. In this case, everybody participates and earns strictly positive expected payoffs playing the unique interior maximizer $x_s(\theta^*)$.[11] For small σ, the intensity of competition is high, and contestants compete away all rents. In this case, contestants strictly mix between $x_s(\theta^*)$ and dropping out. The two cases are illustrated in Figure 2.

Figure 2. The interior maximizer is $x_s(\theta^*_\sigma)$. Under low levels of meritocracy ($\sigma > \underline{\sigma}$, left panel) agents' benefits at $x_s(\theta^*_\sigma)$ exceed costs, so all particpate in equilibrium. However, under high levels of meritocracy ($\sigma \leq \underline{\sigma}$, right panel), a fraction $1 - \gamma^*_\sigma$ of agents drop out such that, in equilibrium, expected payoffs are exactly zero.
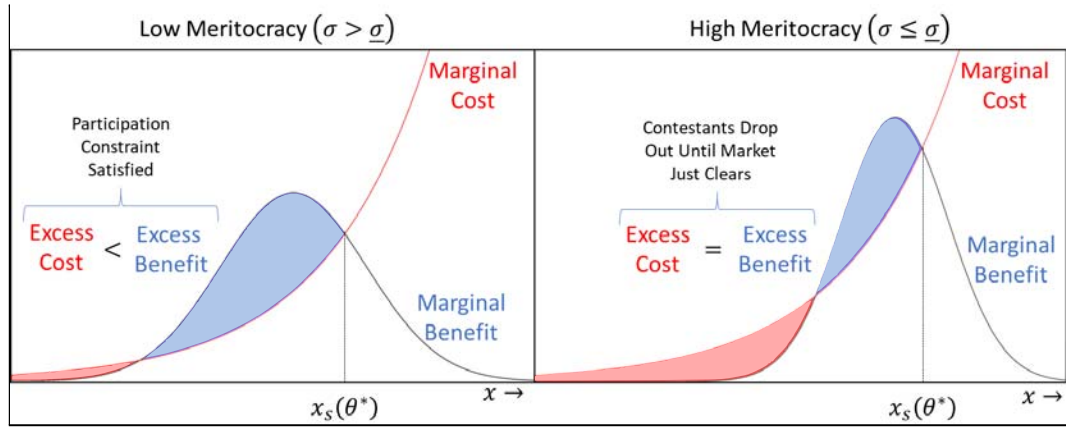
Uniqueness of the equilibrium standard is driven by $c''(\cdot) > 0$. Convexity of c implies that $dx_s(\theta)/d\theta < 1$, which makes the mass of winners $W(\theta)$ strictly decreasing in the standard. Hence, there exists at most one θ that clears the market.

C.   Limits of Meritocracy

We are now ready to prove our main result, namely that, in homogeneous contests, perfect meritocracy is too much of a good thing. It is a corollary of the following lemma, in which we calculate the elasticity of output with respect to σ.

---

[11]To reduce notation, we will, at times, supress the dependence of θ^*, $\hat{x}$, and x_s on σ.

Lemma 4

$$\frac{dO^*/O^*}{d\sigma/\sigma} = \begin{cases} \frac{-1}{\eta_{C'}+1} & for \quad \sigma > \underline{\sigma} \\ \frac{\eta_C - 1}{-soc} & for \quad \sigma < \underline{\sigma} \end{cases} . \tag{8}$$

From Lemma 1 we know that $\eta_{C'} \geq 0$. Since $C(0) = 0$, it is easily verified that $\eta_C \overset{(=)}{>} 1$ if $C''(\cdot) \overset{(=)}{>} 0$. (The reverse implication is not true.) Furthermore, the SOC implies that $-soc > 0$. Hence, $-1/(\eta_{C'} + 1) < 0$, while $(\eta_C - 1)/(-soc) \overset{(=)}{>} 0$. Thus, Lemma 4 has the following implication:

Theorem 1 *Fix a prize structure (m, v). Aggregate output is maximized at $\sigma = \underline{\sigma}$.*

If $C''(\cdot) > 0$, then $\underline{\sigma}$ is the unique maximizer, while perfect meritocracy (i.e., $\sigma \to 0$) is neither output maximizing nor Pareto efficient. If $C''(\cdot) = 0$, then all $\sigma \in (0, \underline{\sigma}]$ are output maximizing.

The intuition for Theorem 1 is as follows. When performance measurement is pure noise, nobody does anything. Hence, injecting some meritocracy into a contest unambiguously raises output. However, at a certain point, competition becomes so intense that contestants are pushed down to their reservation payoff of zero. Crucially, the level of meritocracy at which this happens, $\underline{\sigma}$, is reached before noise is fully eliminated. If σ is reduced below $\underline{\sigma}$, contestants start dropping out with positive probability. Drop-outs earn zero. Homogeneity implies that those who continue to participate also earn zero, in expectation. Hence, contestants earn no rents and the contest organizer should focus on making the pie as large as possible. Jensen's inequality implies that, if marginal costs are increasing, then the output loss from drop-outs exceeds the output gain from the remaining contestants working harder. Thus, reducing σ below $\underline{\sigma}$ reduces aggregate output without benefitting the contestants. This is illustrated in Figure 3.

Remaining contestants work harder because, as illustrated in Figure 2, a reduction in σ raises the marginal benefit of output. We refer to this as the "competition effect" of a rise in meritocracy.

Figure 3. Aggregate output O^*, participation $\hat{\gamma}^*$, and output of non-drop-outs x_s as a function of meritocracy, when $C''() > 0$.
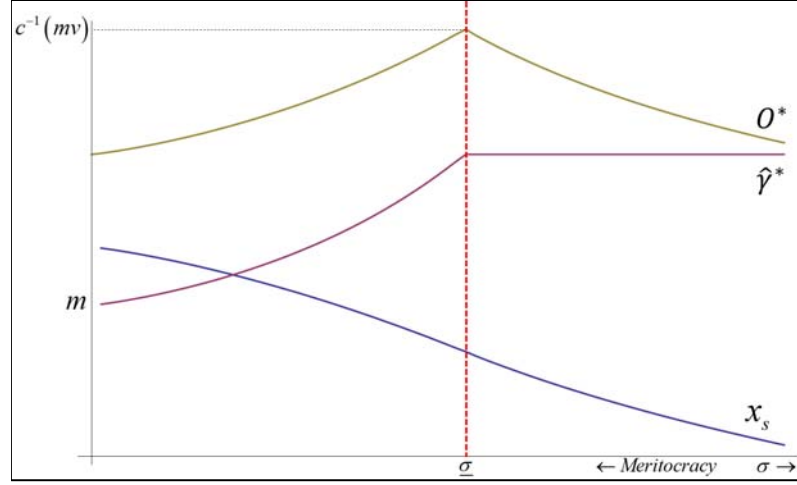
For constant marginal costs, the output loss from drop-outs is exactly off-set by the output gain from the remaining contestants working harder. Thus, reducing σ below $\underline{\sigma}$ keeps aggregate output constant. Finally, as suggested by the intuition and formally proved in the Appendix, for strictly decreasing marginal costs, the output gain from the remainder working harder exceeds the output loss from contestants dropping out. Hence, perfect meritocracy is uniquely output maximizing.

For future reference, we decompose the marginal benefit of meritocracy—expressed as an elasticity—into two terms:

$$\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} = K + A . \tag{9}$$

Here, K denotes the *competition effect* of a change in σ, while A denotes the *attrition effect*. When $\sigma > \underline{\sigma}$, no one drops out in response to a small reduction in σ. Hence, $A \equiv 0$ and only the competition effect is operative. Lemma 4 then implies that $K = -1/(\eta_{C'} + 1) < 0$. When σ falls below $\underline{\sigma}$, the attrition effect kicks in. Its value is calculated—in fact, defined—as a residual: $A \equiv \frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} - K$. Notice that Theorem 1 can be interpreted as saying that, for all $\sigma \leq \underline{\sigma}$ (i.e., from the moment it kicks in), the attrition effect dominates the competition effect. Hence, in a homogeneous contest, the optimal level of meritocracy is $\underline{\sigma}$.

Recall from Proposition 1 that $\underline{\sigma} = vf\left[\bar{F}^{-1}(m)\right]/c'\left[c^{-1}(mv)\right]$. Using the tautology $v = c\left[c^{-1}(vm)\right]/m$, we can rewrite $\underline{\sigma}$ as

$$\underline{\sigma} = \frac{f\left[\bar{F}^{-1}(m)\right]}{\bar{F}\left[\bar{F}^{-1}(m)\right]} \bigg/ \frac{c'\left[c^{-1}(mv)\right]}{c\left[c^{-1}(mv)\right]} . \tag{10}$$

We may conclude that $d\underline{\sigma}/dm < 0$ and $d\underline{\sigma}/dv \leq 0$—i.e., optimal meritocracy is strictly increasing in m and weakly increasing in v. To see this, notice that: 1) the numerator in (10) is strictly decreasing in m, because $\bar{F}^{-1}(\cdot)$ is strictly decreasing while the hazard rate $f/\bar{F}$ is strictly increasing; 2) the denominator is weakly increasing in mv, because $c^{-1}(\cdot)$ is strictly increasing while c'/c is weakly increasing. (The last claim follows from log-convexity of $c(\cdot)$, which was proved in Lemma 1.[12])

At $\sigma = \underline{\sigma}$, contestants earn zero expected profit and win a prize v with probability m. Agents' equilibrium output is therefore $x_s\left(\theta_{\underline{\sigma}}^*\right) = c^{-1}(mv)$, and aggregate output equals

$$O_{\underline{\sigma}}^* = C^{-1}(mv) \ .$$

Notice that $O_{\underline{\sigma}}^*$ does not depend on m and v separately, but only on the prize budget $b \equiv mv$. This means that for a given budget, *any* prize structure (m, v), $m \in (0, 1)$, $v = b/m$, can be optimal, provided we get to choose σ and set it equal to $\underline{\sigma}(m, v)$. As we shall see, under certain conditions, the reverse is also true: for a given budget, *any* $\sigma > 0$ can be optimal, provided we get to choose the appropriate prize structure.

Let $\underline{\sigma}_b(m)$ denote $\underline{\sigma}$ as a function of m when the prize budget is b, i.e.,

$$\underline{\sigma}_b(m) = \frac{f\left[\bar{F}^{-1}(m)\right]}{\bar{F}\left[\bar{F}^{-1}(m)\right]} \left/ \frac{c'\left[c^{-1}(b)\right]}{c\left[c^{-1}(b)\right]} \right. \ . \tag{11}$$

The following proposition follows immediately from (11) and the fact that the hazard rate $f/\bar{F}$ is strictly increasing.

Proposition 2 *Fix a prize budget $b > 0$. The fewer and more valuable are the prizes, the less meritocratic is the optimal contest. Perfect meritocracy is output maximizing iff the contest provides "participation prizes" (i.e., $m \to 1$).*

Formally: 1) $d\underline{\sigma}_b(m)/dm < 0$; 2) $\lim_{m \to 1} \underline{\sigma}_b(m) = 0$; and 3) $\lim_{m \to 0} \underline{\sigma}_b(m) = \lim_{z \to \infty} \kappa f(z)/\bar{F}(z)$, where $\kappa = c\left[c^{-1}(b)\right]/c'\left[c^{-1}(b)\right]$.

From $d\underline{\sigma}_b(m)/dm < 0$ we may conclude that the $\underline{\sigma}$-reducing effect of a rise in m always dominates the $\underline{\sigma}$-raising effect of the associated fall in v. Hence, offering more prizes of lesser value requires a rise in meritocracy. At the extreme where almost everybody receives a prize (i.e., $m \to 1$), the only way to keep contestants on their toes is to remove all room for error. This is achieved by making "participation-prize contests" perfectly meritocratic (i.e., $\sigma \to 0$).[13]

At the other extreme, the contest organizer offers ever fewer but ever larger prizes (i.e., $m \to 0$ and $v = b/m \to \infty$). Despite their high value, the extreme scarcity of

---

[12] Hence, whether $\underline{\sigma}$ is strictly or weakly decreasing in v does not depend on strict or weak convexity of $C(\cdot)$, but rather on strict or weak log-convexity of $c(\cdot)$.

[13] Technically, a participation prize contest with a continuum of agents corresponds to the limit as $n \to \infty$ of a perfectly meritocratic n player contest with $n-1$ equal prizes.

prizes in these "winner-take-all" contests makes agents' efforts fall ever farther below the performance standard (i.e., $\theta^* - x_s \to \infty$). To prevent contestants from getting discouraged and dropping out, optimal performance ranking becomes ever more noisy. Proposition 2 shows that whether the optimal contest remains somewhat meritocratic in the limit or turns into a crapshoot depends on whether the hazard rate $f(z)/\bar{F}(z)$ remains bounded when $z \to \infty$. Bounded $f/\bar{F}$ corresponds to (relatively) "fat tails," which makes outliers relatively common. Therefore, a finite value of the scale parameter suffices to dissuade agents from dropping out and keep them motivated. When $f/\bar{F}$ is unbounded, tails are thin and outliers exceedingly rare. For agents not to drop out, σ must go to infinity. The two cases are illustrated in Figure 4, which depicts $\underline{\sigma}_b(m)$ for Logistic and Normal noise. While $\lim_{m \to 0} \underline{\sigma}_b(m)$ remains bounded in the former case, it goes to ∞ in the latter. Finally, notice that whether any $\sigma > 0$ can be optimal for an appropriately chosen prize structure also turns on whether $\lim_{z \to 0} f(z)/\bar{F}(z) = \infty$.

Figure 4. Optimal meritocracy $\underline{\sigma}_b$ as a function of m, when $C(X) = kX^a$, $k > 0$, $a = 1$. For this class of cost functions, $\underline{\sigma}_b$ does not depend on b (or v), nor on k. (See Example 1).
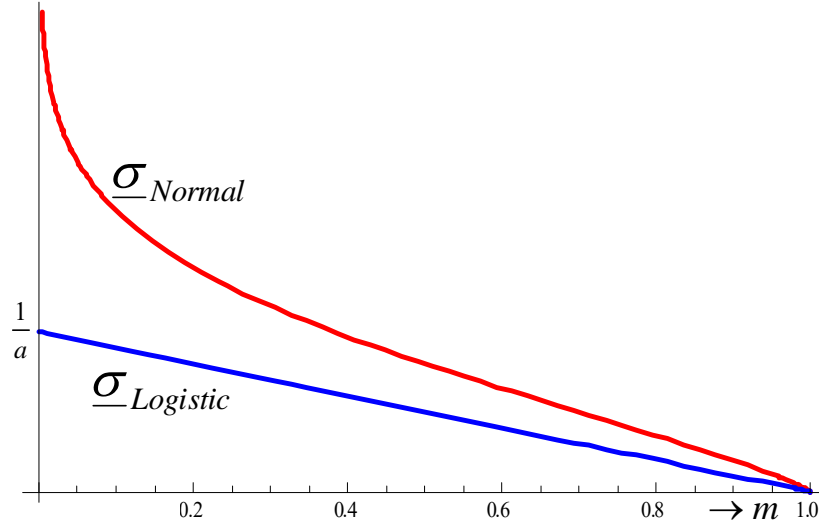
Example 1 Let $C(X) = kX^\alpha$, $k > 0, \alpha \geq 1$. Then

$$\underline{\sigma}_b(m) = \frac{1}{\alpha} \frac{f\left[\bar{F}^{-1}(m)\right]}{\bar{F}\left[\bar{F}^{-1}(m)\right]}.$$

For this cost function, optimal meritocracy depends only on m and α—and not on b (or v), nor on k.

For Normal and Logistic noise, $\underline{\sigma}_b(m)$ is[14]

$$\underline{\sigma}_b^{Normal}(m) = \frac{1}{\alpha}\frac{e^{-\left\{Erfc^{-1}[2(1-m)]\right\}^2}}{\sqrt{2\pi m}},$$

$$\underline{\sigma}_b^{Logistic}(m) = \frac{1}{\alpha}(1-m).$$

These curves are depicted in Figure 4.

So far we have assumed that the prize budget is fixed. Determining the optimal budget requires that we assign some strictly increasing monetary value $p(O)$ to aggregate output. The optimal budget b^* is then

$$b^* \in \arg\sup_{b \geq 0} p\left(c^{-1}(b)\right) - b. \tag{12}$$

Notice that a social planner would replace the second term in (12) by $c(b)$.

Finally, suppose that the contest organizer only cares about the output of winners. For example, this would be the case in an architectural or business-plan competition, where only the best ideas are executed. The organizer's objective function is then

$$\sup_{\sigma > 0} \int_{i \in I} \bar{F}\left(\frac{\theta - x_i}{\sigma}\right) e^{x_i} di.$$

In equilibrium, this reduces to

$$\sup_{\sigma > 0} m e^{x_s(\theta_\sigma^*)}.$$

Implicitly differentiating the equilibrium conditions and solving for $dx_s/(d\sigma/\sigma)$ reveals that

$$\frac{dx_s}{d\sigma/\sigma} = \begin{cases} \frac{-1}{c''/c'} & \text{for} \quad \sigma > \underline{\sigma} \\ \frac{1}{soc} & \text{for} \quad \sigma < \underline{\sigma} \end{cases} < 0.$$

Hence, a rise in meritocracy is unambiguously beneficial, and the organizer achieves his maximum payoff,

$$\lim_{\sigma \to 0} m e^{x_s(\theta_\sigma^*)} = m e^{c^{-1}(v)},$$

under perfect meritocracy. Intuitively, a contest organizer who solely cares about the output of winners can safely ignore drop-outs. Since the output of non-drop-outs continues to increase as σ falls, perfect meritocracy is always optimal.

D.   Multiple Prize Levels

In closing, we study the optimal prize structure when the contest organizer can offer multiple prize levels. Specifically, suppose that there are $\mathbf{m} = \{m_1, ..., m_\lambda\}$ prizes

---

[14]$Erfc^{-1}(z)$ denotes the inverse of the complementary error function $2/\sqrt{\pi}\int_z^\infty \exp\left(-z^2\right) dz$.

with values $\mathbf{v} = \{v_1, ..., v_\lambda\}$, $\lambda \in \mathbb{N}$. Here, $v_1 \geq ... \geq v_\lambda$ and $\sum_{j=1}^{\lambda} m_j < 1$. How does this affect the optimal contest? The following proposition shows that, with homogeneous contestants, multiple prize levels serve no purpose.

Proposition 3 *A contest organizer offering multiple prize levels* $(\mathbf{m}, \mathbf{v})$ *can do no better than to set* $v_1 = v_2 = ... = v_\lambda$.

The intuition for Proposition 3 is straight forward: In an optimal single-prize-level contest, the organizer not only appropriates all the surplus, but also maximizes productive efficiency. Hence, contestants' behavior is first-best. Offering multiple prize levels can only make things worse. Indeed, for strictly increasing marginal costs, the organizer is left strictly worse off when $v_i \neq v_j$ induces contestants to produce heterogeneous levels of output. This again follows from Jensen's inequality. The intuition also indicates why, at the cost of some technical complications, Proposition 3 can be extended to allow for considerably more general prize functions, that assign prize $v(q)$ to contestants at the q-th performance quantile.

When the contest organizer only benefits from the output of winners, the number of winners tends not be a free parameter. Hence, we do not analyze that case here.[15]

IV.   Heterogeneous Contestants

We now study the effects of meritocracy in contests with heterogeneous contestants. First we show that the homogeneous baseline model is not a singularity: its results carry over to contests with sufficiently (rather than perfectly) homogeneous contestants. Second, for general heterogeneous contests, we identify and isolate the three forces that, together, determine the effect on aggregate output of a rise in meritocracy, namely: 1) competition; 2) attrition; and 3) heterogeneity.

---

[15]For completeness: A contest organizer who only benefits from the output of winners and gets to choose m for fixed budget b lets $\sigma \to 0$ and chooses m such that

$$\sup_{m \in (0,1)} m \cdot c^{-1}\left(\frac{b}{m}\right) \ .$$

The FOC for an interior maximum is

$$-\frac{1}{c'\left[c^{-1}\left(\frac{b}{m}\right)\right]}\frac{b}{m} + c^{-1}\left(\frac{b}{m}\right) = 0 \ .$$

Furthermore, the SOC is always satisfied. Finally, the FOC can be rewritten as

$$c(x) = xc'(x) \ .$$

where $x = c^{-1}(v)$.

Except for heterogeneity in ability, the model is as in Section III.. An agent's index $i \in I$ now also labels his ability, which is strictly increasing in i. Let $G(\cdot)$ denote the CDF of i over $I = [0, 1]$. The associated density, which is assumed to exist and to be strictly positive everywhere, is denoted by $g(\cdot)$. As before, the contest organizer tries to maximize aggregate output, which is now given by

$$O_{G(\cdot)} = \int_0^1 e^{x(i)} \mathrm{d}G(i) \ .$$

The cost for contestant $i \in I$ of producing output $X_i \in [0, \infty)$ is given by $C(X_i, i) \in [0, \infty)$. We assume that, for all $i \in I$, $C(\cdot, i)$ has the same properties as $C(\cdot)$ in the homogeneous model of Section III.. For example, this means that $C_X(\cdot, i) > 0$ for $X > 0$, while $C_{XX}(\cdot, i) \geq 0$, where subscripts denote partial derivatives. In addition, we assume that $C_X(\cdot, \cdot)$ is differentiable in i and that $C_{Xi}(\cdot, \cdot) < 0$ for all $X > 0$, $i \in I$. Hence, higher ability means lower (marginal) cost—or, technically speaking, $C(\cdot, \cdot)$ is strictly submodular in its arguments. Let $\eta_{C_X, X}$ denote the elasticity of marginal costs with respect to output, while $\eta_{C_X, i} \equiv -C_{Xi}/C_X$ denotes (minus) the *semi*-elasticity of marginal costs with respect to ability. Also, define $c(x, i) \equiv C(e^x, i) = C(X, i)$.

Let $\hat{x}(\theta, i)$ denote the set of pure best-responses for a contestant i facing standard θ. A (type-symmetric) best-response profile $\hat{\chi}(\theta)$ is a function that assigns a pure best-response $\hat{x} \in \hat{x}(\theta, i)$ to each ability type $i \in I$. Equilibrium is defined as a tuple $\{\theta^*, \hat{\chi}(\theta^*)\}$ of a performance standard and a best-response profile that clears the market for prizes, i.e.,

$$W(\hat{\chi}(\theta^*), \theta^*) = \int_{i \in I} \bar{F} \left[\frac{\theta^* - \hat{x}(\theta^*, i)}{\sigma} \right] \mathrm{d}G(i) = m \ .$$

Notice that we restrict attention to type-symmetric, pure-strategy equilibria. This is essentially without loss of generality because, for any θ, the measure of contestants willing to mix is zero. Indeed, for all θ, almost all agents $i \in I$ strictly prefer to either produce $x_s(\theta, i)$ or drop out (See Corollary 1 in the Appendix).

In the following proposition, "single-crossing in ability" means that, if boundary type $\underline{i} \in I$ is indifferent between participating and dropping out, then all $i \overset{(<)}{>} \underline{i}$ strictly prefer participating (dropping out). "Unique" means unique up to a zero measure of contestants.

Proposition 4 *In the contest with heterogeneous contestants, equilibrium exists and is unique.*

Participation satisfies single-crossing in ability. That is, if $\underline{i}$ exists, then

$$\hat{x}(\theta^*, i) = \begin{cases} x_s(\theta^*, i) & for \quad i > \underline{i} \\ \{-\infty, x_s(\theta^*, i)\} & for \quad i = \underline{i} \\ -\infty & for \quad i < \underline{i} \end{cases} \ .$$

Otherwise, $\hat{x}(\theta^*, i) = x_s(\theta^*, i)$ for all $i \in I$.

For $\sigma > 0$ sufficiently small, $\underline{i}$ exists. Furthermore, $d\underline{i}/d\sigma < 0$ and $\lim_{\sigma \to 0} \underline{i} = 1 - m$.

We now show that the baseline model of Section III. is not a singularity. To do so, we extend the set of allowable σ to $[0, \infty)$ (i.e., we close the set by including zero) and define output at $\sigma = 0$ as $O^*(0) \equiv \lim_{\sigma \to 0} O^*(\sigma)$. That this limit exists is shown in the proof of Theorem 2. Next, let $\hat{\sigma}_{G(\cdot)} \in [0, \infty)$ denote an output maximizing level of σ in the heterogeneous contest with abilities G. Since the optimal σ is clearly bounded and $O^*(\sigma)$ is continuous on $[0, \infty)$, $\hat{\sigma}_{G(\cdot)}$ must exist. Finally, let $\mathbf{1}_{\{i=i_0\}}$ denote the degenerate distribution that puts mass 1 on $i = i_0$, $i_0 \in I$, while $\underline{\sigma}_{\mathbf{1}_{\{i=i_0\}}}$ denotes $\underline{\sigma}$ in the homogeneous contest with cost function $C(\cdot) \equiv C(\cdot, i_0)$.

Proposition 5 *Suppose $C_{XX}(\cdot, i) > 0$. In a sufficiently homogeneous contest, perfect meritocracy is strictly suboptimal. Formally,*

$$\hat{\sigma}_{G(\cdot)} \xrightarrow{i \xrightarrow{\text{Pr}} i_0} \underline{\sigma}_{\mathbf{1}_{\{i=i_0\}}} > 0 \ .$$

For strictly increasing marginal costs, Proposition 5 extends Theorem 1 to sufficiently, rather than perfectly, homogeneous contests. Obviously, this is essential for the relevance of the original result.

To see why the proposition restricts attention to $C_{XX}(\cdot, i) > 0$, recall that in a homogenous contest with constant marginal costs, any $\sigma \in (0, \underline{\sigma}]$ maximizes aggregate output. Similarly, in an *approximately* homogeneous contest, any $\sigma \in \left[0, \underline{\sigma}_{\mathbf{1}_{\{i=i_0\}}}\right]$ *approximately* maximizes aggregate output. However, since $O^*_{G(\cdot)}(\sigma)$ is almost flat in σ on this interval, a sequence of exact maximizers may jump around a lot as $i \xrightarrow{\text{Pr}} i_0$. Indeed, for $C_{XX} = 0$, $\hat{\sigma}_{G(\cdot)}$ may converge to any $\sigma \in \left[0, \underline{\sigma}_{\mathbf{1}_{\{i=i_0\}}}\right]$, or not converge at all, depending on the details of the convergence process of i to i_0.

We now turn to the other extreme and study contests that are very heterogeneous. Notice that (the degree of) heterogeneity is a joint property of CDF G and the cost function C. This is illustrated by the fact that a population of contestants can be "homogenized" in two different but equivalent ways: either by 1) fixing $C(\cdot, \cdot)$ and letting $i \xrightarrow{\text{Pr}} i_0$, as was done in Proposition 5; or by 2) fixing $G(\cdot)$ and uniformly letting $\eta_{C_X, i} \to 0$ for all $(X, i) \in [0, \infty) \times I$. In both cases, (almost) all contestants end up facing the same (marginal) cost function, making the contest homogeneous.

Without loss of generality, we may assume that abilities i are uniformly distributed on $I \in [0, 1]$. To see this, notice that we can always relabel abilities according to $i' \equiv G(i)$. The original contest with abilities $i \sim G(\cdot)$ and cost function $C(\cdot, \cdot)$ is then isomorphic to a contest with uniform abilities i' and cost function $C(\cdot, G^{-1}(\cdot))$. It only remains to observe that $C(\cdot, G^{-1}(\cdot))$ continues to satisfy our assumptions for cost functions.

Once we have normalized abilities i and turned them into population quantiles, notice that heterogeneity is solely a property of the cost function. This is the motivation behind the following definition. We say that a statement X is true for "sufficiently heterogeneous contests," if the statement holds for all (admissible) cost functions whose marginal costs fall sufficiently fast in ability i. Formally, there needs to exist a $\beta > 0$ such that X holds if $\eta_{C_X,i} \geq \beta$ for all $(X, i) \in [0, \infty) \times I$.

In terms of productive efficiency, heterogeneity clearly favors transferring output from low to high ability contestants, while convexity of C in X disfavors it. This suggests that for sufficiently heterogeneous contests, or close to linear costs, attrition raises output and, hence, perfect meritocracy should be optimal. Our next result shows that this intuition is wrong. The theorem does rely on the assumption that we can make $\eta_{C_X,i}$ large without $\eta_{C_X,X}$ blowing up as well. For that purpose we assume that $\eta_{C_X,X}$ is uniformly bounded in (X, i).[16]

Theorem 2 *Fix a prize structure (m, v). For sufficiently heterogeneous contests, perfect meritocracy is strictly suboptimal.*

Theorem 2 is driven by complacency of high ability types in a world of perfectly accurate performance ranking. When measurement is noisy, high types build in a "margin of error" when choosing how much to produce, i.e., they set $\hat{x} > \theta^*$. This way they insure against an unlucky draw of ε. The more meritocratic the contest becomes, the smaller the chance of an unlucky draw and, hence, the less need to exceed the standard. Indeed, when $\sigma \to 0$, the output of all non-drop-outs converges to θ^*. Theorem 2 shows that, for small σ and sufficiently heterogeneous contests, the output loss from complacency dominates the efficiency gain from attrition. This result does rely on the assumption that the elasticity of marginal costs with respect to output remains bounded. To see why, notice that if $\eta_{C_X,X}$ explodes when $\eta_{C_X,i}$ becomes large, then the complacency-induced drop in output as meritocracy rises may remain small even among very-high-ability types, simply because the "margin of error" is small to begin with. As a result, the efficiency effect could dominate the complacency effect.

We now consider the benefits of meritocracy away from extreme homo- and heterogeneity. (The CDF G may or may not be uniform.) Using Leibniz' rule, we find for the marginal benefit of meritocracy expressed in terms of (semi)elasticities:

$$\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} = \underbrace{\int_{\max\{0,\underline{i}\}}^{1} \frac{e^{x_s(\theta^*,i;\sigma)}}{O^*} \frac{\mathrm{d}x_s\left(\theta^*,i;\sigma\right)}{\mathrm{d}\sigma/\sigma} g\left(i\right) \mathrm{d}i}_{=\tilde{K}+\tilde{H}+\tilde{A}_{ind}} + \underbrace{1_{\{\underline{i}\}} \frac{e^{x_s(\theta^*,\underline{i};\sigma)}}{O^*} g\left(\underline{i}\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma}}_{=\tilde{A}_{dir}} . \quad (13)$$

Here, $\max\{0, \underline{i}\} \equiv \underline{i}$ if $\underline{i}$ exists and 0 otherwise, while $1_{\{\underline{i}\}}$ denotes the indicator function that takes on the value 1 if $\underline{i}$ exists and 0 otherwise.

---

[16] Alternatively, we may assume that $C\left(\cdot, \cdot\right)$ is multiplicatively separable in X and i. This makes $\eta_{C_X,X}$ and $\eta_{C_X,i}$ fully independent of each other.

In order to further decompose $\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma}$, we focus on cost functions with constant elasticity of marginal costs, $\eta_{C_X,X}$.

Remark 1 *Elasticity $\eta_{C_X,X}$ is constant iff $C(X,i) = kX^\alpha \cdot l(i)$, where $k > 0$ and $\alpha = \eta_{C,X} = \eta_{C_X,X} + 1$. Our earlier assumptions on $C(\cdot,\cdot)$ imply that $l'(\cdot)$ exists and is < 0.*

When $\eta_{C_X,X}$ is constant, we can decompose $\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma}$ into: 1) competition effect $\tilde{K}$; 2) heterogeneity effect $\tilde{H}$; and 3) attrition effect $\tilde{A}$, which in turn consists of a direct effect $\tilde{A}_{dir}$ and an indirect effect $\tilde{A}_{ind}$. That is,

$$\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} = \tilde{K} + \tilde{H} + \underbrace{\left(\tilde{A}_{dir} + \tilde{A}_{ind}\right)}_{=\tilde{A}} .$$

Recall that, in a homogeneous contest, the competition effect is given by $K \equiv -1/(\eta_{C'} + 1)$. When $\eta_{C_X,X}$ is constant, the obvious analogue for heterogeneous contests is

$$\tilde{K} \equiv -\frac{1}{\eta_{C_X,X} + 1} = -\frac{1}{\alpha} .$$

Clearly, the attrition effect $\tilde{A}$ is operative only when boundary type $\underline{i}$ exists; otherwise, $\tilde{A} \equiv 0$. Also notice that attrition has both a direct and an indirect effect on output. The direct effect, $\tilde{A}_{dir}$, refers to the output lost from the contestants who actually drop out. It is given by the second term on the RHS of (13) and, provided it is operative, is always strictly positive.

Dropping out also has an indirect effect on output, namely, through a (relative) fall in standards. This indirect effect, $\tilde{A}_{ind}$, is incorporated in the first term of (13). To see this, notice that

$$\frac{\mathrm{d}x_s(\theta^*,i;\sigma)}{\mathrm{d}\sigma/\sigma} = \frac{\partial x_s(\theta^*,i;\sigma)}{\partial\sigma/\sigma} + \frac{\partial x_s(\theta^*,i;\sigma)}{\partial\theta^*}\frac{\mathrm{d}\theta^*}{\mathrm{d}\sigma/\sigma} . \tag{14}$$

In Lemma 30 in the Appendix, it is shown that $\frac{\partial\theta^*}{\partial\sigma/\sigma}$ is of the form

$$\frac{\mathrm{d}\theta^*}{\mathrm{d}\sigma/\sigma} = B - 1_{\{\underline{i}\}}Z\frac{\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} , \tag{15}$$

where, $B, Z > 0$ and $\mathrm{d}\underline{i}/(\mathrm{d}\sigma/\sigma) < 0$. The second term in (15) corresponds to the indirect effect of attrition, and its sign confirms that attrition leads to a (relative) fall in standards.

Together, (14) and (15) imply that the indirect effect on $x_s\left(\theta^*_\sigma, i\right)$ is $1_{\{\underline{i}\}}Z\left(-\mathrm{d}\underline{i}\right)/\left(\mathrm{d}\sigma/\sigma\right)\cdot\partial x_s\left(\theta^*_\sigma, i\right)/\partial\theta^*$. Hence, the indirect effect on output is

$$\tilde{A}_{ind} \equiv 1_{\{\underline{i}\}}Z\frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma}\int_{\underline{i}}^1 e^{x_s(\theta^*,i)}\frac{\partial x_s\left(\theta^*, i\right)}{\partial\theta^*}g\left(i\right)\mathrm{d}i \ . \tag{16}$$

Notice that the sign of $\tilde{A}_{ind}$ is ambiguous, because a fall in standards raises the output of contestants who need a "lucky break" (i.e., $x_s\left(i\right) < \theta^*$), while it lowers the output of contestants who merely need to avoid an "unlucky break" (i.e., $x_s\left(i\right) > \theta^*$). This follows directly from the FOC.

Finally, we define $\tilde{H}$ as the residual:

$$\begin{aligned}
\tilde{H} &\equiv \frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} - \tilde{K} - \tilde{A}\\
&= \int_{\max\{0,\underline{i}\}}^1 \frac{e^{x_s(\theta^*,i;\sigma)}}{O^*}\left[\frac{\partial x_s\left(\theta^*, i; \sigma\right)}{\partial\sigma/\sigma} + \frac{\partial x_s\left(\theta^*, i; \sigma\right)}{\partial\theta^*}B\right]g\left(i\right)\mathrm{d}i + \frac{1}{\alpha} \ .
\end{aligned}$$

When all contestants (strictly prefer to) participate, $\tilde{A} = 0$. Hence, in the interior, $\tilde{H}$ corresponds to the differential effect of meritocracy on output in the heterogeneous contest as compared to the homogeneous contest with the same elasticity of marginal cost. That is why we refer to $\tilde{H}$ as the heterogeneity effect.

The following proposition shows that $\tilde{K}$, $\tilde{H}$, $\tilde{A}$ can be unambiguously signed. Furthermore, $\tilde{H} \xrightarrow{i \overset{\mathrm{Pr}}{\to} i_0} 0$, which is consistent with our interpretation of $\tilde{H}$ as the effect of heterogeneity.

Proposition 6 *Suppose $\eta_{C_X,X}$ is constant. The marginal benefit of meritocracy, $\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma}$, can be decomposed into additively separable competition, attrition, and heterogeneity effects. The competition effect $\tilde{K}$ is < 0. Provided the marginal type, $\underline{i}$, exists, the attrition effect $\tilde{A}$ is > 0, while $\tilde{A} \xrightarrow{i \overset{\mathrm{Pr}}{\to} i_0} A_{\mathbf{1}_{\{i=i_0\}}}$. Otherwise, $\tilde{A} = 0$. The heterogeneity effect $\tilde{H}$ is > 0, while $\tilde{H} \xrightarrow{i \overset{\mathrm{Pr}}{\to} i_0} 0$.*

By analogy with the homogeneous contest, we have called $\tilde{K} = -1/\alpha$ the competitive effect of a rise in meritocracy. It is always negative because, ceteris paribus, a fall in σ raises competition which, in turn, increases aggregate output. The terms $\tilde{A}$ and $\tilde{H}$ capture the effects of attrition and heterogeneity, respectively. Proposition 6 establishes that the latter is strictly positive and, provided it is operative, so is the former. This means that both attrition and heterogeneity *reduce* the marginal benefit of meritocracy.

Starting from a sufficiently high value, suppose we reduce σ. Unlike in the homogeneous case, aggregate output may continue to rise beyond the point where

the attrition effect kicks in. That is, in a heterogeneous contest, the output loss from drop-outs may be more than compensated for by the remaining contestants working harder. The reason is that attrition proceeds from the bottom of the ability distribution, which implies that drop-outs are of lower ability than those who remain. As a result, while in the homogeneous case $\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} > 0$ iff $A > 0$, in the heterogeneous case $\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma}$ may very well be < 0 even though $\tilde{A} > 0$.

That heterogeneity reduces the benefits of meritocracy—i.e., $\tilde{H} > 0$—is not obvious. A fall in σ lifts the peak of f and thins the tails. This raises the marginal benefit of output for agents operating close to the standard and reduces it for those operating farther away. Optimal output follows suit. Put differently, a rise in meritocracy discourages the weak, encourages the middle, and makes the strong complacent. Proposition 6 shows that discouragement and complacency of the extremes dominate encouragement of the middle. The assumption that $\eta_{C_X,X}$ is constant is not innocuous in this regard: Suppose we selectively lower $\eta_{C_X,X}$ for agents operating close to the standard. This amplifies the encouragement effect of a rise in meritocracy and thus could make $\tilde{H}$ negative. (Alternatively, we can raise $\eta_{C_X,X}$ for high ability agents and/or decrease $\eta_{C_X,X}$ for low ability agents. This reduces complacency and/or discouragement effects.)[17]

$\tilde{H} > 0$ implies that, up to the point where contestants start dropping out, a rise in meritocracy is less valuable in heterogeneous than in homogeneous contests. However, beyond this point, the marginal benefit ordering can be reversed. To see this, recall from Proposition 5 that in sufficiently homogeneous contests attrition is always something to be avoided. By contrast, in heterogeneous contests, some attrition may be warranted if the drop-outs are of low ability. Indeed, as the following example illustrates, even perfect meritocracy may be optimal.

---

[17]Notice, however, that when η_{C_X} is variable, the definition of $\tilde{K}$ and the interpretation of $\tilde{H}$ are no longer obvious.

Figure 5. Output, attrition, and the decomposition of $(dO^*/O^*)/(d\sigma/\sigma)$ into competition, attrition, and heterogeneity effects, for different levels of heterogeneity (Example 2).
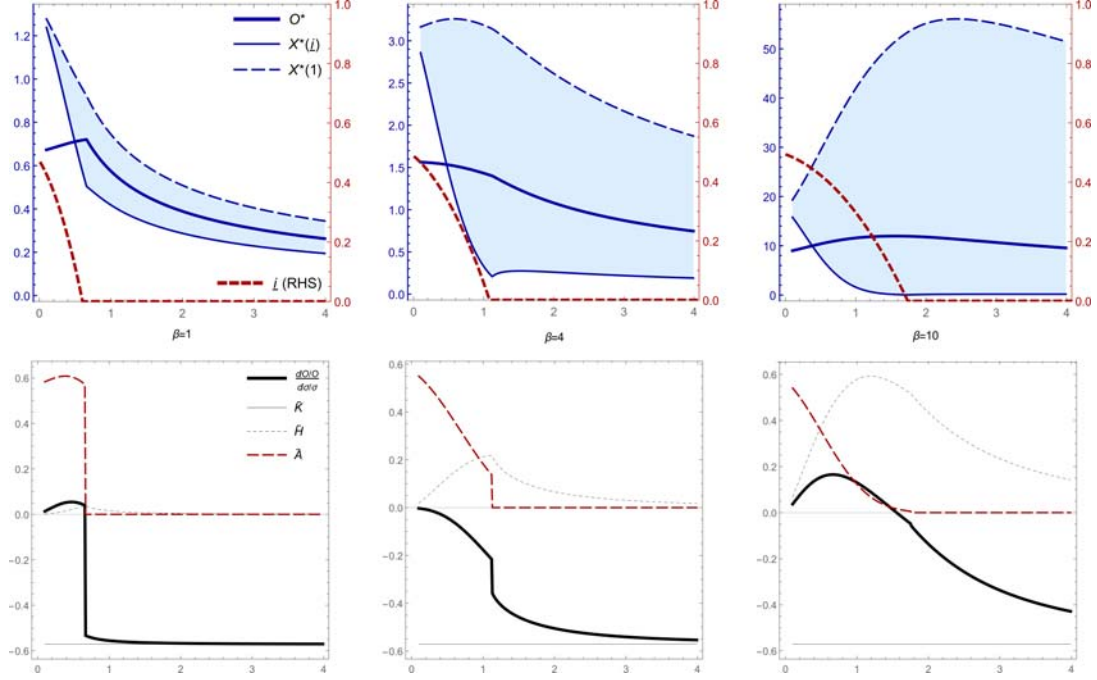
Example 2 *Let $i \sim U[0,1]$, $\varepsilon \sim N(0, \sigma^2)$, and $C(X, i) = X^{7/4}/e^{\beta i}$. Furthermore, $m = 1/2$ and $v = 1$.*

- **Low heterogeneity:** *Let $\beta = 1$. As in the homogeneous case, aggregate output is maximized at the value of σ where contestants start dropping out. As soon as the attrition effect, $\tilde{A}$, kicks in, it dominates the competition effect $\tilde{K}$. The heterogeneity effect, $\tilde{H}$, is tiny. (Figure 5, left panels.)*

- **Medium heterogeneity:** *Next, let $\beta = 4$. Now the competition effect dominates the sum of the (diminished) attrition effect and the (enlarged) heterogeneity effect. Hence, aggregate output is maximized under perfect meritocracy. (Center panels.)*

- **High heterogeneity:** *When $\beta = 10$, the heterogeneity effect $\tilde{H}$ dominates, such that $\hat{\sigma}$ reverts to being strictly positive. Imperfect meritocracy keeps high-ability contestants on their toes and (most) low ability types from dropping out. (Right panels.) If we increase β even further, $\hat{\sigma}$ becomes so large that even the lowest ability type strictly prefers to participate. Even though low types contribute essentially nothing to aggregate output, such a large σ is optimal because it keeps high types on their toes.*

V. Discrete Number of Contestants

So far, we have used the power of the continuum to facilitate the analysis. We now show that the results of the baseline model carry over, in essence, to single-prize contests with a discrete number of homogeneous agents.

With atomistic agents, contestants produce maximum output, and the organizer collects all the rents, at the highest level of meritocracy for which a pure-strategy equilibrium (PSE) exists. At this point, contestants are just indifferent between participating and dropping out. As we shall see, the same is true for discrete agents. However, rather than dropping out with some probability when meritocracy rises beyond the threshold, discrete agents mix across multiple output levels. With increasing marginal costs, mixing is inefficient.

The key to adapting the baseline model to discrete contestants lies in ensuring that the probability of winning remains location-scale and log-concave. To win in the continuum, recall that a contestant's output measure, y, had to surpass a *deterministic* performance threshold, θ. The associated probability of winning was $\bar{F}\left[(\theta - x)/\sigma\right]$. In a two-player contest, player 1's output measure y_1 must surpass his rival's *stochastic* measure y_2. The probability of winning is now $\Pr\left[\varepsilon_1 - \varepsilon_2 > x_2 - x_1\right]$. For our earlier analysis to carry through, the distribution of the difference $\varepsilon_1 - \varepsilon_2$—rather than the distribution of ε_i itself—needs to be location-scale and log-concave. For more than two players, it is the difference between y_i and the maximum of all other players' output measures that must have these properties.

Below, we first study the homogeneous 2-player contest pioneered by Lazear and Rosen (LR, 1981). LR focus on PSE. To guarantee existence thereof, they restrict attention to performance measurement that is "sufficiently" noisy. We allow for arbitrary noisiness and, as for the baseline model, derive the output-maximizing level of meritocracy. Then we study a particular instance for which the analysis extends to $n > 2$ players; namely, when errors are drawn from an extreme value type-I maximum distribution (EVTIM). Since the LR contest with EVTIM noise is isomorphic to an n-player Tullock contest, practitioners may find this case particularly useful.

A. 2-Player Lazear-Rosen Contests

Consider a contest with two identical, risk neutral agents competing for a single prize of value v. Agents' cost functions and measurement errors are as in the baseline model of Section III., except that now the distribution of the *difference* of log errors belongs to a log-concave, location-scale family, rather than the error itself. The distribution is assumed to be symmetric around its location parameter, zero. The scale parameter is $\rho > 0$, the CDF is Φ, the DCDF is $\bar{\Phi}$, and the PDF is φ. If

the two agents have the same measured performance, which almost surely does not happen, the winner is determined by the toss of a fair coin. We refer to this setup as the 2-player LR contest.

For two commonly-used distributions, the following example shows how σ translates into ρ.

Example 3 *(i) If $\varepsilon_1, \varepsilon_2 \sim Normal\,(0, \sigma)$, then $(\varepsilon_1 - \varepsilon_2) \sim Normal\left(0, \sqrt{2}\sigma\right)$. Hence, $\rho = \sqrt{2}\sigma$.*

(ii) If $\varepsilon_1, \varepsilon_2 \sim EVTIM\,(\mu, \sigma)$, then $(\varepsilon_1 - \varepsilon_2) \sim Logistic\,(0, \sigma)$. Hence, in this case, $\rho = \sigma$.

For all noise levels ρ, existence of a symmetric equilibrium follows from Theorem 6 of Dasgupta and Maskin (1986). Equilibrium may be in mixed strategies, however.

Suppose agent $j \in \{1, 2\}$ plays $x_j \in \mathbb{R}$. Then for agent i, $i \neq j$, the expected payoff from playing $x_i \in \mathbb{R}$ is

$$\pi_i\,(x_i, x_j) = v \Pr\,(x_i + \varepsilon_i \geq x_j + \varepsilon_j) - c\,(x_i) = v\bar{\Phi}\left(\frac{x_j - x_i}{\rho}\right) - c\,(x_i) \ .$$

The FOC for an interior extremum is

$$\frac{\mathrm{d}\pi_i\,(x_i, x_j)}{\mathrm{d}x_i} = \frac{v}{\rho}\varphi\left(\frac{x_j - x_i}{\rho}\right) - c'\,(x_i) = 0 \ . \tag{17}$$

As in the baseline, strict log-concavity of φ and weak log-convexity of c' guarantee that, for each x_j, at most one solution to the FOC satisfies the SOC for a maximum.

For a profile of mutually consistent solutions (x_1, x_2) to the FOCs and SOCs to constitute a PSE—denoted by (x_1^*, x_2^*)—the players' participation constraints need to be satisfied. In Lemma 5 (below) we show that all PSEs are symmetric. Hence, the participation constraint reduces to $v/2 \geq c\,(x^*)$, where $x^* = x_1^* = x_2^*$.

Agents' behavior in a PSE mirrors that of the full-participation equilibrium in the baseline model. When performance measurement is pure noise, neither player exerts any effort, each wins with probability $1/2$ and enjoys an expected payoff $v/2$. Initially, raising meritocracy increases output and reduces agents' payoffs until, eventually, the participation constraint becomes binding. Increasing meritocracy even further causes the interior best replies to yield negative expected payoffs. At this point, a PSE ceases to exist.

Let $\bar{x} \equiv c^{-1}\,(v/2)$, i.e., $\bar{x}$ denotes the symmetric log-output level that pushes contestants' expected payoffs down to zero. Similarly, let $\bar{X} \equiv C^{-1}\,(v/2)$. Next, define $\underline{\rho} \equiv v\varphi\,(0)\,/c'\,(\bar{x})$; that is, $\underline{\rho}$ corresponds to the (unique) noise level for which $\bar{x}$ satisfies the players' FOCs.

At $\underline{\rho}$, agents' participation constraints hold with equality. Therefore,

$$\underline{\rho} = \frac{\varphi(0)}{\bar{\Phi}(0)} \Bigg/ \frac{c'\left[c^{-1}(v/2)\right]}{c\left[c^{-1}(v/2)\right]} . \tag{18}$$

(Cf., the expression for $\underline{\sigma}$ in (10).) When ρ falls below $\underline{\rho}$, rather than dropping out, discrete agents start mixing across multiple output levels. This claim is proved in the next proposition, which also formalizes a number of claims made earlier.

Lemma 5 *(i) For all $\rho > 0$, there exists a symmetric equilibrium. A PSE exists iff $\rho \geq \underline{\rho}$. The PSE is unique and symmetric, where $x^* = c'^{-1}\left[\frac{v}{\rho}\varphi(0)\right]$.*

(ii) In the PSE, individual effort strictly increases with meritocracy. If $\rho = \underline{\rho}$, then $x^* = \bar{x}$ and $O^* = 2\bar{X}$, while contestants' equilibrium payoffs are zero.

(iii) For $\rho < \underline{\rho}$, equilibrium payoffs are zero in all symmetric equilibria (which are in mixed strategies).

(iv) If $C''(\cdot) \overset{(=)}{>} 0$, then in a MSE, $O^* \overset{(\leq)}{<} 2\bar{X}$. If $C''(\cdot) = 0$ and $\rho < \underline{\rho}$, then in a symmetric MSE, $O^* = 2\bar{X}$.

Part *(i)* of Lemma 5 implies that only mixed-strategy equilibria exist for $\rho < \underline{\rho}$. Among these MSEs, the symmetric equilibria yield zero expected payoffs to contestants (part *(iii)*). Notice that MSEs may also exist for $\rho > \underline{\rho}$, and these equilibria may yield strictly positive payoffs. In fact, non-symmetric MSEs may yield strictly positive payoffs even if $\rho \leq \underline{\rho}$.[18] According to part *(iv)* of the lemma, MSEs fail to maximize output when marginal costs are strictly increasing. The reason is that mixing over actions with unequal marginal costs is wasteful.

Lemma 5 effectively implies the next proposition, which is the 2-player analogue of Theorem 1. It shows that the optimal level of meritocracy is equal to $\underline{\rho}$, $\underline{\rho} > 0$. We may conclude that the central result of the baseline model carries over, in essence, to LR contests.

Theorem 1.LR *In a 2-player LR contest, aggregate output is maximized in the PSE with noise level $\underline{\rho} > 0$.*

If $C''(\cdot) > 0$, then $\underline{\rho}$ is the unique maximizer, while perfect meritocracy (i.e., $\rho \to 0$) is neither output maximizing nor Pareto efficient. If $C''(\cdot) = 0$, then all $\rho \in (0, \underline{\rho}]$ are both output maximizing and Pareto efficient in a symmetric MSE.

Finally, like $\underline{\sigma}$, threshold $\underline{\rho}$ is weakly decreasing in v. This follows from (18) and weak log-convexity of $c(\cdot)$.

---

[18]Whether they exist remains an open question.

B. n-Player Tullock Contests

Now consider a contest with an arbitrary but finite number of symmetric players competing for a single prize. It is readily verified that Lemma 5 generalizes to n players, provided that the probability of winning remains log-concave and location-scale. An n-player extension of Theorem 1 LR is then easily proved.[19]

Unfortunately, this result is not readily applicable. In order to win in an n-player contest, a player's output measure y_i must exceed the maximum of all other players' output measures y_j, $j \neq i$. The difference between a location-scale random variable and the maximum of multiple other such random variables is a complicated object, whose distribution is not generally known nor likely to be location-scale. In the remainder of this section, we explicate an important exception.

Lemma 6 *An n-player LR contest with noise $\varepsilon_i \sim EVTIM(0, \sigma)$ gives rise to a Tullock contest with discriminatoriness $r = 1/\sigma$. That is, player i's probability of winning is logistic and given by*

$$\Pr\left\{\max_{j \neq i}\{x_j + \varepsilon_j\} \leq x_i + \varepsilon_i\right\} = \frac{e^{\frac{x_i}{\sigma}}}{\sum_{j=1}^{n} e^{\frac{x_j}{\sigma}}} = \frac{X_i^r}{\sum_{j=1}^{n} X_j^r} \ .$$

The proof of Lemma 6 follows McFadden (1973, 1974) and Fu and Lu (2012). It relies on the fact that: 1) the maximum of EVTIM random variables with scale σ is again distributed EVTIM with scale σ; and 2) the difference of two EVTIMs with scale σ is logistically distributed with the same scale.

The logistic distribution is log-concave and location-scale. Hence, in the special case of a Tullock contest, the n-player extension of Theorem 1.LR does in fact apply.

Let $\bar{r}$ denote the highest level of discriminatoriness for which there exists a PSE in a Tullock contest. Then:

Theorem 1.T *In an n-player Tullock contest, aggregate output is maximized in the PSE with discriminatoriness $\bar{r}$. Here,*

$$\bar{r} = \frac{c'\left[c^{-1}(v/n)\right]}{c\left[c^{-1}(v/n)\right]} \bigg/ \left(1 - \frac{1}{n}\right) \ .$$

If $C''(\cdot) > 0$, then $\bar{r}$ is the unique maximizer, while a perfectly discriminating contest (i.e., $r \to \infty$) is neither output maximizing nor Pareto efficient. If $C''(\cdot) = 0$, then all $r \in [\bar{r}, \infty)$ are output maximizing.

---

[19]To be precise, Lemma 5 generalizes to n players, except for uniqueness of the PSE. Notice, however, that Proposition 1 and its n-player extension do not crucially depend on uniqueness of the PSE. Furthermore, for the special case of the Tullock contest, we do in fact prove uniqueness. See Lemma 36 in the Appendix.

Let $\bar{x}_n$ denote the n-player analogue of $\bar{x}$—i.e., $\bar{x}_n = c^{-1}(v/n)$. Theorem 1.T says that discriminatoriness in a Tullock contest is optimal when r equals the curvature of the cost function at $\bar{x}_n$ divided by the probability of losing. Observe that a rise in n increases the probability of losing which, in turn, reduces $\bar{x}_n$. Furthermore, log-convexity of $c(\cdot)$ implies that as $\bar{x}_n$ falls, c'/c decreases. Hence:

Remark 2 *In a Tullock contest, optimal discriminatoriness, $\bar{r}$, is strictly decreasing in the number of contestants, n.*

Intuitively, the larger the number of contestants, the less volatile the stochastic performance threshold a contestant needs to surpass in order to win. To compensate for this reduction in randomness, noise in performance measurement must increases—i.e., $\bar{r}$ falls.

Finally, notice that optimal meritocracy (discriminatoriness) continues to be weakly increasing in v.

These properties are illustrated in the following example.

Example 4 *If $C(X) = kX^{\alpha}$, $k > 0, \alpha \geq 1$, then $\bar{r} = \alpha \frac{n}{n-1}$. Hence, $\bar{r} \downarrow \alpha$ when $n \to \infty$.*

Notice that Example **??** corresponds to the endogenous entry model of Fu *et al.* (2015) for the case where entry costs are zero. In that case, the expression for r^*, defined in Theorem 5 of their paper, indeed reduces to $\alpha n/(n-1)$.

To summarize, our finding that meritocracy can be too much of a good thing is not an artifact of the continuum, but also obtains in contests with discrete players. The basic intuition also carries through: when meritocracy exceeds the threshold where rents are completely dissipated, contestants start to mix. When marginal costs are increasing, this is inefficient. Hence, aggregate output must fall.

VI.   Conclusion

We have shown that in contests, too much meritocracy can be a bad thing: with sufficiently homogeneous contestants, perfectly accurate performance ranking is both Pareto inefficient and output reducing. Hence, it hurts the contest organizer without benefiting contestants.

For heterogeneous contestants, we have decomposed the effect of a rise in meritocracy into competition, attrition, and heterogeneity effects. While the competition effect raises aggregate output, both attrition and heterogeneity effects

reduce it. The optimal level of meritocracy is then determined by the interplay of these forces. Whether maximum meritocracy is something to strive for depends on the degree of heterogeneity of contestants or, equivalently, on the degree of supermodularity of marginal costs in effort and ability. For sufficiently heterogeneous contests we have shown that perfect meritocracy is once more suboptimal, since it discourages the weak and makes the strong complacent.

While contests are a pervasive feature of modern life, our finding that too much meritocracy can be counter-productive may be particularly relevant for education policy. In this regard, the Dutch and French education systems constitute interesting contrasts. While admission to elite "grandes écoles" in France is notoriously meritocratic, scarce university slots in the Netherlands are allocated by means of a weighted lottery.

In the context of US college admissions, it is interesting to observe how the Scholastic Aptitude Test (SAT) has changed over the years. Initially, test takers were not expected to complete the test, let alone achieve a perfect score (Jacobsen, 2017). Hence, unlike today, there was no censoring from above.[20] Later, an essay was added to the test, which was "not as immediately quantifiable as the other elements" (JROTC, 2008). In 2016, the penalty for wrong answers was abolished, while the number of multiple-choice options was reduced from four to three. Arguably, all these changes increased the noisiness of the test results. Whether this was in any way optimal is an open question.

---

[20]The original SAT from 1926 featured 315 questions with a time limit of 97 minutes.

Since the "recentering" of 1995, students can miss as many as four questions and still get a perfect score. As a result, perfect scores have risen 36 fold, from about 1 in 50,000 to 1 in 1,400. See Jacobsen (2017).

References

Bagnoli, M., and T. Bergstrom, 2004, "Log-concave probability and its applications," *Economic Theory* 26: 445-469.

Baik, Kyung Hwan, 1998, "Difference-form Contest Success Functions and Effort Levels in Contests," *European Journal of Political Economy* 14: 685-701.

Baye, Michael R., Dan Kovenock, and Casper G. de Vries, 1994, "The Solution to the Tullock Rent-seeking Game When $R > 2$: Mixed-strategy Equilibria and Mean Dissipation Rates," *Public Choice* 81: 363-380.

Che, Yeon-Koo, and Ian Gale, 2000, "Difference-form Contests and the Robustness of All-pay Auctions," *Games and Economic Behavior* 30: 22-43.

Clark, D. J., and C. Riis, 1998, "Competition Over More Than One Prize," *American Economic Review* 88(1), 276–289.

Corchón, Luis C., 2007, "The Theory of Contests: A Survey," *Review of Economic Design* 11: 69-100.

Cornes, R. and R. Hartley, 2005, "Asymmetric Contests With General Technologies," *Economic Theory* 26: 923–946.

Cowgill, Bo. 2015, "Competition and Productivity in Employee Promotion Contests," *Working Paper.*

Dasgupta, P. and E. Maskin, 1986, "The existence of equilibrium in discontinuous economic games, I: Theory," *The Review of Economic Studies* 53: 1-26.

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

Friedersdorf, 2012, "The Cult of Smartness: How Meritocracy Is Failing America," *Atlantic Monthly*, June 14, 2012, available at: *www.theatlantic.com/politics/archive/2012/06/the-cult-of-smartness-how-meritocracy-is-failing-america/258492/*

Fu, Qiang, and Jingfeng Lu, 2012, "Micro Foundations of Multi-Prize Lottery Contests: A Perspective of Noisy Performance Ranking," *Social Choice and Welfare* 38: 497-517.

Fu, Qiang, Qian Jiao, and Jingfeng Lu, 2015, "Contests with endogenous entry," *International Journal of Game Theory* 44: 387–424.

Guesnerie, R. and J.J. Laffont, 1984, "A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm," *Journal of public Economics* 25: 329-369.

Hayes, C., 2012, *Twilight of the Elites: America After Meritocracy*, New York: Crown Publishing.

Holmstrom, Bengt, and P. Milgrom, 1991, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, & Organization* 7: 24-52.

Hirshleifer, J., 1989, "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice* 63: 101-112.

Jacobsen, Erik, 2017, "A (Mostly) Brief History Of The SAT And ACT Tests," available at: *www.erikthered.com/tutor/sat-act-history.html*.

Jensen, M. K., 2016, "Existence, uniqueness, and comparative statics in contests." In *Equilibrium Theory for Cournot Oligopolies and Related Games* (pp. 233-244), Cham: Springer.

JROTC, 2008, "What You Need to Know About the Revised Scholastic Aptitude Test," available at: *www.jrotccollegeprep.com/parents.php?content=incl_content_gen_adv_14*.

KNAW, 2013, "Is Loting (On)rechtvaardig?" available at: *www.knaw.nl/nl/actueel/agenda/is-loting-on-rechtvaardig*.

Konrad, Kai, 2009, *Strategy and Dynamics in Contests*, Oxford: Oxford University Press.

Lazear, Edward, 1989, "Pay Equality and Industrial Politics," *Journal of Political Economy* 97: 561-80.

Lazear, Edward and Sherwin Rosen, 1981, "Rank-Order Tournaments as Optimal Labor Contracts." *Journal of Political Economy* 89: 841-864.

List, Benjamin, 2017. "Crowd-based peer review can be good and fast." *Nature News*, 546 (7656): 9.

Lohr, Steve, 2012. "The age of big data." *New York Times*, February 11, 2012, available at: *www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html*

Lohr, Steve, 2014, "Unblinking Eyes Track Employees: Workplace Surveillance Sees Good and Bad," *New York Times*, June 21, 2014, available at: *www.nytimes.com/2014/06/22/technology/workplace-surveillance-sees-good-and-bad.html?_r=0*

May, Theresa, 2017, Speech in Nottinghamshire on April 6 to Launch the Conservative Campaign Ahead of Local Elections in May 2017. available at: *www.politicshome.com/news/uk/political-parties/conservative-party/theresa-may/news/84903/full-theresa-mays-speech*

Moldovanu, B. and A. Sela, 2001, "The Optimal Allocation of Prizes in Contests,"*American Economic Review* 91: 542-558.

Morgan, John, Dana Sisak, and Felix Várdy, 2018, "The Ponds Dilemma," *The Economic Journal* 128: 1634-1682.

Myerson, R. B., 1981, "Optimal Auction Design," *Mathematics of Operations Research,* 6(1), 58-73.

O'Keeffe, W. Kip Viscusi, and Richard J. Zeckhauser, 1984, "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics* 2: 27-56.

PISA, 2016, "PISA 2015: Results in Focus," available at: *http://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf.*

Olszewski, W., and Siegel, R., 2016, "Large contests," *Econometrica* 84(2): 835-854.

Olszewski, W., and Siegel, R., 2018, "Performance-Maximizing Contests," Working Paper.

Scaggs, Alexandra, 2018, "How to Win a Debate in the Cult of Meritocracy," *Financial Times*, March 18, 2018, available at: *ftalphaville.ft.com/2018/03/28/1522201766000/How-to-win-a-debate-in-the-cult-of-meritocracy/*

Solomon."Ecclesiastes" 9:11.

Strathern, Marilyn, 1997, "'Improving ratings': audit in the British University system," *European Review* 5: 305-321.

Tullock, Gordon, 1980, *Towards a Theory of the Rent-seeking Society*, Edward Elgar Publishing.

Wade, Laura, 2010, "Interview: Laura Wade, playwright," *The Scotsman*, 8 April.

Online Appendix:

Proofs

Baseline Model

Setup

  Proof of Lemma 1:

Part 1: Notice that $c(-\infty) = C(0) = 0$. The claim that $c'(-\infty) = 0$ follows from the fact that, for $x \in [-\infty, \infty)$, $c(x) = \int_{-\infty}^{x} c'(\xi)\,\mathrm{d}\xi$ is finite.

Part 2: Observe that

$$c'(x) = \frac{\mathrm{d}c(x)}{\mathrm{d}x} = \frac{\mathrm{d}C(e^x)}{\mathrm{d}e^x}\frac{\mathrm{d}e^x}{\mathrm{d}x} = \frac{\mathrm{d}C(X)}{\mathrm{d}X}X = XC'(X) \ .$$

For $x = \ln X > -\infty$, we have $X > 0$ and $C'(X) > 0$. Hence, $c'(x) > 0$.

Next,

$$
\begin{aligned}
c''(x) &= \frac{\mathrm{d}^2 c(x)}{(\mathrm{d}x)^2} = \frac{\mathrm{d}^2 C(e^x)}{(\mathrm{d}x)^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}C(e^x)}{\mathrm{d}x}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}C(e^x)}{\mathrm{d}e^x}\frac{\mathrm{d}e^x}{\mathrm{d}x}\right) \\
&= \frac{\mathrm{d}C(e^x)}{\mathrm{d}e^x}e^x + e^x\frac{\mathrm{d}^2 C(e^x)}{\mathrm{d}x\mathrm{d}e^x} = \frac{\mathrm{d}C(e^x)}{\mathrm{d}e^x}e^x + e^x\frac{\mathrm{d}^2 C(e^x)}{(\mathrm{d}e^x)^2}\frac{\mathrm{d}e^x}{\mathrm{d}x} \\
&= \frac{\mathrm{d}C(X)}{\mathrm{d}X}X + X^2\frac{\mathrm{d}^2 C(X)}{(\mathrm{d}X)^2} = XC'(X) + X^2 C''(X) \ .
\end{aligned}
$$

We already know that $XC'(X) > 0$, while $C''(X) \geq 0$ by assumption. Hence $c''(x) > 0$.

Part 3: Notice that

$$
\begin{aligned}
\frac{\mathrm{d}\left[c''(x)/c'(x)\right]}{\mathrm{d}x} &= \frac{\mathrm{d}\left[1 + X \cdot C''(X)/C'(X)\right]}{\mathrm{d}X}\frac{\mathrm{d}X}{\mathrm{d}x} \\
&= \left[X\frac{\mathrm{d}\left[C''(X)/C'(X)\right]}{\mathrm{d}X} + \frac{C''(X)}{C'(X)}\right]e^x \ .
\end{aligned}
$$

Therefore, $\mathrm{d}\left[c''(x)/c'(x)\right]/\mathrm{d}x \geq 0$, such that $c'(x)$ is weakly log-convex, iff

$$X\frac{\mathrm{d}\left[C''(X)/C'(X)\right]}{\mathrm{d}X} \geq -\frac{C''(X)}{C'(X)} \ .$$

This condition is equivalent to:

$$C''(X) = 0 \ \text{ or } \ \frac{\mathrm{d}\left[C''(X)/C'(X)\right]/\left[C''(X)/C'(X)\right]}{\mathrm{d}X/X} = \eta_{C''/C'} \geq -1 \ .$$

Now notice that the inequality holds by assumption. Hence, $c'(x)$ is weakly log-convex.

Finally, log-convexity of $c'(\cdot)$ implies log-convexity of $c(\cdot)$. This follows from Lemma 3 in Bagnoli and Bergstrom (2004).

Part 4: First, observe that

$$c''(x)/c'(x) = 1 + XC''(X)/C'(X) = 1 + \eta_{C'} ,$$

and

$$c'(x)/c(x) = XC'(X)/C(X) = \eta_C .$$

Because $c(\cdot)$ is log-convex, it follows that

$$1 + \eta_{C'} = c''(x)/c'(x) \geq c'(x)/c(x) = \eta_C . \tag{19}$$

Finally, because $C(0) = 0$,

$$C(X) = \int_0^X C'(X)\,\mathrm{d}Z \overset{(<)}{\leq} XC'(X) ,$$

depending on whether $C''(\cdot) \overset{(=)}{>} 0$. Hence,

$$\eta_C = c'(x)/c(x) = XC'(X)/C(X) \overset{(=)}{>} XC'(X)/XC'(X) = 1 . \tag{20}$$

Together, (19) and (20) imply the result.

Part 5: Follows from straight-forward calculation.

This completes the proof of Lemma 1.   ■

Equilibrium   This section contains a proof of Proposition 1. It consists of two parts. First, we prove existence and uniqueness of equilibrium. Then we show that, as a function of σ, output $\hat{x}(\theta_\sigma^*; \sigma)$ takes on the shape claimed in the proposition.

Existence and Uniqueness
We first establish some useful properties of the pure best-response correspondence $\hat{x}(\theta)$.

Lemma 7 *For all θ, $\hat{x}(\theta)$ is non-empty, compact-valued, and upper hemicontinuous (UHC).*

Moreover, (i) there exists θ' such that for all $\theta > \theta'$, $\hat{x}(\theta) = -\infty$, (ii) $\lim_{\theta \to \infty} \theta - \hat{x}(\theta) = \infty$, (iii) $\lim_{\theta \to -\infty} \theta - \hat{x}(\theta) = -\infty$, and (iv) $\lim_{\theta \to -\infty} \hat{x}(\theta) = -\infty$.

Proof. Notice that $\pi(x, \theta)$ is jointly continuous in x and θ. Moreover, for every θ, clearly, $\hat{x}(\theta)$ is bounded by some finite bound $d(\theta)$. Hence, we may limit our search for best responses $\hat{x}(\theta)$ to the compact interval $D(\theta) \equiv [-\infty, d(\theta)]$. Applying Berge's Maximum theorem to $\pi(x, \theta)$ on $D(\theta)$ implies that $\hat{x}(\theta)$ is non-empty, compact-valued, and UHC for all θ.

Part (i): Consider an arbitrary $x > -\infty$ and notice that

$$\lim_{\theta \to \infty} \pi(x, \theta) = \lim_{\theta \to \infty} v\bar{F}\left(\frac{\theta - x}{\sigma}\right) - c(x) < 0 .$$

The inequality holds because the chance of winning, $\bar{F}\left(\frac{\theta - x}{\sigma}\right)$, goes to zero when $\theta \to \infty$, while $c(x) > 0$. By continuity, there exists θ' such that for all $\theta > \theta'$, any output $x > -\infty$ violates the participation constraint. Therefore, it must be that $\hat{x}(\theta) = -\infty$ for all $\theta > \theta'$.

Part (ii): $\lim_{\theta \to \infty} \theta - \hat{x}(\theta) = \infty$ follows immediately from (i).

Part (iii): To prove that $\lim_{\theta \to -\infty} \theta - \hat{x}(\theta) = -\infty$, suppose to the contrary that $\lim_{\theta \to -\infty} \theta - \hat{x}(\theta)$ remains bounded from below by some $d > -\infty$. Then

$$\lim_{\theta \to -\infty} \pi[\hat{x}(\theta), \theta] = \lim_{\theta \to -\infty} v\bar{F}\left[\frac{\theta - \hat{x}(\theta)}{\sigma}\right] - c[\hat{x}(\theta)]$$

$$\leq \lim_{\theta \to -\infty} v\bar{F}\left[\frac{\theta - \hat{x}(\theta)}{\sigma}\right] \leq v\bar{F}\left(\frac{d}{\sigma}\right) < v .$$

Now consider the alternative output schedule $\hat{x}(\theta) = \frac{1}{2}\theta$. For this $\hat{x}(\theta)$,

$$\lim_{\theta \to -\infty} \pi[\hat{x}(\theta), \theta] = \lim_{\theta \to -\infty} v\bar{F}\left(\frac{\frac{1}{2}\theta}{\sigma}\right) - c\left(\frac{1}{2}\theta\right) = v .$$

Hence, $\hat{x}(\theta) = \frac{1}{2}\theta$ is a profitable deviation. Contradiction.

Part (iv): Finally, to prove that $\lim_{\theta \to -\infty} \hat{x}(\theta) = -\infty$, suppose to the contrary that $\hat{x}(\theta)$ remains bounded from below by $d > -\infty$. In that case, the marginal cost is bounded from below by $c'(x)|_{x=d} > 0$, while the marginal benefit $\frac{v}{\sigma}f\left[\frac{\theta - \hat{x}(\theta)}{\sigma}\right] \to 0$ for $\theta \to -\infty$, because $\lim_{\theta \to -\infty} \theta - \hat{x}(\theta) = -\infty$ (see part (iii)). Therefore, $\hat{x}(\theta)$ is not optimal. Contradiction. ∎

Lemma 8 *If $\hat{x}(\cdot)$ is single-valued at θ', then it is locally differentiable and $\hat{x}'(\theta') < 1$. In particular,*

$$\hat{x}'(\theta') \in \left\{ \begin{array}{ll} (-\infty, 0) & \text{if } \hat{x}(\theta') < \theta' \\ (0, 1) & \text{if } \hat{x}(\theta') \geq \theta' \end{array} \right. .$$

Proof. If $\hat{x}(\theta') = -\infty$ and single-valued, then $x = -\infty$ is strictly better than any other value of x. By Berge's maximum theorem, $\pi\left[\hat{x}(\theta), \theta\right]$ is continuous in θ. It follows that, in a neighborhood of θ', $x = -\infty$ remains the unique best response. Hence, $\hat{x}(\theta)$ is differentiable around θ' (with zero derivative).

If $\hat{x}(\theta') > -\infty$ and single-valued, then differentiability of $\hat{x}(\theta)$ in a neighborhood of θ' follows from applying the implicit function theorem (IFT) to the FOC at the point $(\theta', \hat{x}(\theta'))$. Furthermore, implicitly differentiating the FOC with respect to θ yields

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}\theta} = \frac{\frac{v}{\sigma^2} f'}{\frac{v}{\sigma^2} f' + c''} \ .$$

The SOC guarantees that the denominator of this expression is strictly positive. Hence, if $f' \leq 0$ (equivalently $\hat{x}(\theta') \geq \theta'$), then $\hat{x}'(\theta') \leq 0 < 1$. Alternatively, if $f' > 0$ (equivalently $\hat{x}(\theta') < \theta'$) then $0 < \hat{x}'(\theta') < 1$, because $c'' > 0$ (see Lemma 1, part 2). ∎

The following lemma implies that marginal cost and marginal benefit curves intersect at most twice.

Lemma 9 *The ratio $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x)$ is either single-peaked or strictly decreasing in x. Furthermore, $\lim_{x \to \infty} \frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x) = 0$.*

Proof. Since $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ is log-concave in x by assumption and $1/c'(x)$ is weakly log-concave by Lemma 1, the ratio is strictly log-concave, because the product of log-concave functions is log-concave. Log-concavity implies quasiconvexity which, in turn, implies unimodality. For x large enough, the ratio strictly decreases since the numerator is a density with finite variance and $c''(x) \geq 0$. To see that $\lim_{x \to \infty} \frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x) = 0$, notice that $\lim_{x \to \infty} \frac{1}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) \to 0$, and that $c'(x)$ remains bounded away from zero when $x \to \infty$, which is implied by parts 2 and 3 of Lemma 1. ∎

Recall that the largest crossing point between $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ and $c'(x)$ is denoted by x_s. We use Lemma 9 to prove that $\hat{x}(\theta) \subset \{-\infty, x_s(\theta)\}$, which is Lemma 2 in the main text.

Proof of Lemma 2: Fix a θ. Notice that Lemma 9 implies that marginal cost $c'(x)$ crosses marginal benefit $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ at most twice. Furthermore, since $\lim_{x \to \infty} \frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x) = 0$, if the curves do not cross then $c'(x) \geq \frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ for all $x \in \mathbb{R}$. If they cross once then $c'(x)$ crosses $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ from below. If they cross twice then $c'(x)$ crosses $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ first from above at $x_t < \theta$ and then from below at $x_s > \theta$. Together, these observations imply that only $-\infty$ and $x_s(\theta)$ are potential best responses. ∎

The following lemma further refines the characterization of $\hat{x}(\theta)$ given in Lemma 2.

Lemma 10 *Either $\hat{x}(\theta) = x_s(\theta)$ for all θ, or there exists a unique $\bar{\theta} \in \mathbb{R}$ such that*

$$\hat{x}(\theta) = \begin{cases} x_s(\theta) & for \quad \theta < \bar{\theta} \\ \{-\infty, x_s(\theta)\} & for \quad \theta = \bar{\theta} \\ -\infty & for \quad \theta > \bar{\theta} \end{cases} .$$

Proof. From Lemma 2 we know that $\hat{x}(\theta) \subset \{-\infty, x_s(\theta)\}$. The envelope theorem implies that $\frac{\partial}{\partial x}\pi(x,\theta)\big|_{x=x_s(\theta)} = 0$. Hence,

$$\frac{d}{d\theta}\pi[x_s(\theta),\theta] = \frac{\partial}{\partial\theta}\pi[x_s(\theta),\theta] = -\frac{v}{\sigma}f\left[\frac{\theta - x_s(\theta)}{\sigma}\right] < 0 .$$

Therefore, there can exist at most one $\bar{\theta}$ such that $\pi\left(-\infty, \bar{\theta}\right) = \pi\left[x_s\left(\bar{\theta}\right), \bar{\theta}\right]$. If it exists, $\hat{x}(\theta)$ must be of the form

$$\hat{x}(\theta) = \begin{cases} x_s(\theta) & if \quad \theta < \bar{\theta} \\ \{-\infty, x_s(\theta)\} & if \quad \theta = \bar{\theta} \\ -\infty & if \quad \theta > \bar{\theta} \end{cases} .$$

If it does not exist, then either $\hat{x}(\theta) = -\infty$ for all θ, or $\hat{x}(\theta) = x_s(\theta)$ for all θ. Clearly, $\hat{x}(\theta) = -\infty$ for all θ is inconsistent with profit maximization. Hence, in this case, it must be that $\hat{x}(\theta) = x_s(\theta)$ for all θ. ∎

Let $\Omega(\theta)$ denote the set of masses of winners, w, that can arise when agents symmetrically best respond to standard θ. The inverse, $\Omega^{-1}(w)$, denotes the set of standards that give rise to a mass w of winners. Formally, for θ and $w \in (0,1)$,

$$\Omega(\theta) \equiv \left\{w : \exists \hat{\Gamma}(\theta) \text{ st. } W[\theta, \hat{\gamma}(\theta)] = w\right\}, \text{ while}$$

$$\Omega^{-1}(w) \equiv \left\{\theta \in \mathbb{R} : \exists \hat{\Gamma}(\theta) \text{ st. } W[\theta, \hat{\gamma}(\theta)] = w\right\} .$$

As with $\hat{x}(\cdot)$, we treat the correspondences $\Omega(\cdot)$ and $\Omega^{-1}(\cdot)$ as functions when they are single-valued. Finally, recall that $\bar{\theta}$ solves that $\pi\left(-\infty, \bar{\theta}\right) = \pi\left[x_s\left(\bar{\theta}\right), \bar{\theta}\right]$.

Lemma 11   1. *If $\bar{\theta}$ does not exist, then $\Omega(\theta)$ is single-valued, differentiable, and strictly decreasing for all $\theta \in \mathbb{R}$.*

   2. *If $\bar{\theta}$ exists then: i) $\Omega(\theta)$ is single-valued, differentiable, and strictly decreasing for all $\theta < \bar{\theta}$; ii) $\Omega\left(\bar{\theta}\right) = \left[0, \lim_{\theta\uparrow\bar{\theta}}\Omega(\theta)\right]$; and iii) $\Omega(\theta) = 0$ for $\theta > \bar{\theta}$.*

   3. $\lim_{\theta\to-\infty}\Omega(\theta) = 1$ *and* $\lim_{\theta\to\infty}\Omega(\theta) = 0$.

Proof.
Part 1: If $\bar{\theta}$ does not exist, then we know from Lemma 10 that $\hat{x}(\theta) = x_s(\theta)$ for all θ. Hence, $\Omega(\theta)$ is given by

$$\Omega(\theta) = \bar{F}\left[\frac{\theta - x_s(\theta)}{\sigma}\right]$$

and single-valued for all θ.

Next, Lemma 8 and differentiability of $\bar{F}(\cdot)$ imply that $\Omega(\theta)$ is differentiable in θ. Moreover,

$$\frac{\mathrm{d}\Omega(\theta)}{\mathrm{d}\theta} = -\left[1 - \frac{\mathrm{d}x_s(\theta)}{\mathrm{d}\theta}\right]\frac{1}{\sigma}f < 0 \;,$$

where the inequality follows from Lemma 8. Hence, $\Omega(\theta)$ is strictly decreasing in θ.

Part 2: If $\bar{\theta}$ does exist, then Lemma 10 implies

$$\hat{x}(\theta) = \begin{cases} x_s(\theta) & \text{if} \quad \theta < \bar{\theta} \\ \{-\infty, x_s(\theta)\} & \text{if} \quad \theta = \bar{\theta} \\ -\infty & \text{if} \quad \theta > \bar{\theta} \end{cases}$$

. For $\theta < \bar{\theta}$, the proof that $\Omega(\theta)$ is single-valued, differentiable, and strictly decreasing is the same as when $\bar{\theta}$ does not exist. For $\theta > \bar{\theta}$, $\Omega(\theta) = 0$ follows immediately from $\hat{x}(\theta) = -\infty$.

Finally, for $\theta = \bar{\theta}$, let $\gamma \in [0, 1]$ denote the probability that a contestant plays $x_s(\theta)$, and notice that the mass of winners is $W(\bar{\theta}, \gamma) = \gamma\bar{F}\left[\frac{\bar{\theta} - x_s(\bar{\theta})}{\sigma}\right]$. Therefore, the set of masses of winners is a closed interval:

$$\Omega(\bar{\theta}) = \left[0, \bar{F}\left[\frac{\bar{\theta} - x_s(\bar{\theta})}{\sigma}\right]\right] = \left[0, \lim_{\theta\uparrow\bar{\theta}}\Omega(\theta)\right] \;.$$

Part 3. From the second part of Lemma 7 we know that $\lim_{\theta\to\infty}\theta - \hat{x}(\theta) = \infty$ and $\lim_{\theta\to-\infty}\theta - \hat{x}(\theta) = -\infty$. Hence, $\lim_{\theta\to\infty}\Omega(\theta) = \lim_{\theta\to\infty}\bar{F}\left[\frac{\theta-\hat{x}(\theta)}{\sigma}\right] = 0$, while $\lim_{\theta\to-\infty}\Omega(\theta) = \lim_{\theta\to-\infty}\bar{F}\left[\frac{\theta-\hat{x}(\theta)}{\sigma}\right] = 1$. ∎

Lemma 12 *There exists a unique symmetric equilibrium. The equilibrium standard, θ^*, is unique.*

Proof. Lemma 11 implies that, for all $m \in (0, 1)$, $\Omega^{-1}(m)$ exists and is single-valued. Hence, there exists a unique potential equilibrium standard $\theta^* = \Omega^{-1}(m)$.

If $\theta^* < \bar{\theta}$ or $\bar{\theta}$ does not exist, then Lemmas 10 and 11 imply that $\hat{x}(\theta^*) = x_s(\theta^*)$ and $\Omega(\theta^*) = m$. Hence, $(\theta^*, x_s(\theta^*))$ constitutes the unique (symmetric) equilibrium.

If $\theta^* = \bar{\theta}$, then $\hat{x}(\theta^*) = \{-\infty, x_s(\bar{\theta})\}$ and $\Omega(\theta^*) = [0, \lim_{\theta \uparrow \bar{\theta}} \Omega(\theta)]$. Denote by $\gamma \in [0,1]$ the probability of playing $x_s(\bar{\theta})$. Market clearing requires that, in a symmetric equilibrium,

$$\gamma \bar{F} \left[\frac{\bar{\theta} - x_s(\bar{\theta})}{\sigma} \right] = m \ .$$

On the RHS, $0 < m \le \bar{F} \left[\frac{\bar{\theta} - x_s(\bar{\theta})}{\sigma} \right]$, where the last inequality follows from the fact that, by assumption, $\Omega^{-1}(m) = \theta^* = \bar{\theta}$. The LHS is strictly increasing in γ, running from 0 to $\bar{F} \left[\frac{\bar{\theta} - x_s(\bar{\theta})}{\sigma} \right]$. Together with the intermediate value theorem, this implies that there exists a unique $\hat{\gamma}^* \in (0,1]$ that clears the market. Therefore, in the unique symmetric equilibrium, $\theta^* = \bar{\theta}$ and contestants mix between $x_s(\bar{\theta})$ and $-\infty$ with probabilities $\hat{\gamma}^*$ and $1 - \hat{\gamma}^*$, respectively. ∎

$\hat{x}(\theta^*_\sigma; \sigma)$ as a function of σ.

It remains to prove that, as a function of σ, $\hat{x}(\theta^*_\sigma; \sigma)$ takes on the shape claimed in Proposition 1. First we show that the distance between x_s and the standard vanishes when $\sigma \to 0$.

Lemma 13 *Let θ_σ be some function of σ. Then*

$$\lim_{\sigma \to 0} \theta_\sigma - x_s(\theta_\sigma; \sigma) = 0 \ .$$

Proof. When $\sigma \to 0$, notice that density $\frac{1}{\sigma} f\left(\frac{\cdot}{\sigma}\right)$ converges to the Dirac measure with mass 1 at zero. As x_s is an intersection point between $c'(x)$ and and $\frac{v}{\sigma} f\left(\frac{\theta_\sigma - x}{\sigma}\right)$, it immediately follows that x_s converges to θ_σ when $\sigma \to 0$. ∎

The following lemma shows that the equilibrium standard, θ^*_σ, remains bounded as $\sigma \to 0$.

Lemma 14 *As $\sigma \to 0$, the equilibrium standard θ^*_σ remains bounded.*

Proof.
Claim 1: As $\sigma \to 0$, θ^*_σ remains bounded from above.

Proof of Claim 1: Suppose to the contrary that there exists a converging (sub)sequence $\{\sigma_n\}_{n \in \mathbb{N}} \xrightarrow{n \to \infty} 0$ such that $\{\theta^*_{\sigma_n}\}_{n \in \mathbb{N}} \xrightarrow{n \to \infty} \infty$. As $n \to \infty$, the payoff

from producing $x_s\left(\theta^*_{\sigma_n}; \sigma_n\right)$ approaches

$$\lim_{n\to\infty} v\bar{F}\left[\frac{\theta^*_{\sigma_n} - x_s\left(\theta^*_{\sigma_n}; \sigma_n\right)}{\sigma_n}\right] - c\left[x_s\left(\theta^*_{\sigma_n}; \sigma_n\right)\right] = -\lim_{n\to 0} c\left[x_s\left(\theta^*_{\sigma_n}; \sigma_n\right)\right] < 0\ ,$$

which violates the participation constraint. Lemma 2 then implies that all contestants choose $x = -\infty$. Hence, $\lim_{n\to 0} W\left(\theta^*_{\sigma_n}\right) = 0 < m$, which contradicts that θ^*_σ is an equilibrium standard.

Claim 2: As $\sigma \to 0$, θ^*_σ remains bounded from below.

Proof of Claim 2: Suppose to the contrary that there exists a converging (sub)sequence $\{\sigma_n\}_{n\in\mathbb{N}} \overset{n\to\infty}{\longrightarrow} 0$ such that $\{\theta^*_{\sigma_n}\}_{n\in\mathbb{N}} \overset{n\to\infty}{\longrightarrow} -\infty$. Then, in the limit, contestants can achieve the highest possible payoff, v, by producing $x_n = \theta^*_{\sigma_n} + \delta\sqrt{\sigma_n}$, $\delta > 0$. To see this, notice that

$$v\bar{F}\left(\frac{\theta^*_{\sigma_n} - x_n}{\sigma_n}\right) - c\left(x_n\right) = v\bar{F}\left(\frac{-\delta}{\sqrt{\sigma_n}}\right) - c\left(\theta^*_{\sigma_n} + \delta\sqrt{\sigma_n}\right) \overset{n\to\infty}{\longrightarrow} v\ ,$$

where convergence to v follows from $\{\sigma_n\} \overset{n\to\infty}{\longrightarrow} 0$ and $\{\theta^*_{\sigma_n}\} \overset{n\to\infty}{\longrightarrow} -\infty$. Because $\bar{F}\left(\frac{-\delta}{\sqrt{\sigma_n}}\right) \overset{n\to\infty}{\longrightarrow} 1$, all contestants are guaranteed a prize. Therefore, $\lim_{n\to 0} W\left(\theta^*_{\sigma_n}\right) = 1 > m$. This contradicts that $\theta^*_{\sigma_n}$ is an equilibrium standard. ∎

Lemma 14 allows us to invoke the Bolzano-Weierstrass theorem. It says that every $\{\sigma_n\} \overset{n\to\infty}{\longrightarrow} 0$ has an associated convergent (sub)sequence of equilibrium standards, $\{\theta^*_{\sigma_n}\}_{n\in\mathbb{N}}$. We denote the limit point by $\theta^*_{\sigma=0}$.

To simplify notation, from hereon, we go back to suppressing the dependence of θ, $\hat{x}$ and x_s on σ. For example, we may write $x_s\left(\theta^*_\sigma\right)$ or $x_s\left(\theta^*\right)$ for $x_s\left(\theta^*_\sigma; \sigma\right)$.

Lemma 15 *When $\sigma \to 0$, the probability a non-drop-out wins goes to 1. That is,*

$$\lim_{\sigma\to 0} \bar{F}\left[\frac{\theta^* - x_s\left(\theta^*\right)}{\sigma}\right] = 1\ .$$

Proof. Suppose, to the contrary, that there exists a sequence $\{\sigma_n\}_{n\in\mathbb{N}} \to 0$ with associated converging (sub)sequence $\{\theta^*_{\sigma_n}\}_{n\in\mathbb{N}} \to \theta^*_{\sigma=0}$ of standards such that the probability of winning, $\bar{F}\left[\frac{\theta^*_{\sigma_n} - x_s\left(\theta^*_{\sigma_n}\right)}{\sigma_n}\right]$, remains bounded away from 1. That is, $\lim_{n\to 0} \bar{F}\left[\frac{\theta^*_{\sigma_n} - x_s\left(\theta^*_{\sigma_n}\right)}{\sigma_n}\right] \leq d < 1$, for some constant $0 < d < 1$.

Now consider an output strategy $x_n = \theta^*_{\sigma_n} + \delta\sqrt{\sigma_n}$, $\delta > 0$. The limit payoff satisfies

$$
\begin{aligned}
\lim_{n\to 0} v\bar{F}\left(\frac{\theta^*_{\sigma_n} - x_n}{\sigma_n}\right) - c\left(x_n\right) &= \lim_{n\to 0} v\bar{F}\left(\frac{-\delta}{\sqrt{\sigma_n}}\right) - c\left(\theta^*_{\sigma_n} + \delta\sqrt{\sigma_n}\right) \\
&= v - c\left(\theta^*_{\sigma=0}\right) \\
&> vd - c\left(\theta^*_{\sigma=0}\right) \\
&\geq \lim_{n\to 0} v\bar{F}\left[\frac{\theta^*_{\sigma_n} - x_s\left(\theta^*_{\sigma_n}\right)}{\sigma_n}\right] - c\left[x_s\left(\theta^*_{\sigma_n}\right)\right] ,
\end{aligned}
$$

where we have used that $\left\{\theta^*_{\sigma_n}\right\}_{n\in\mathbb{N}} \to \theta^*_{\sigma=0}$ and $\lim_{n\to 0}\theta^*_{\sigma_n} - x_s\left(\theta^*_{\sigma_n}\right) = 0$ (Lemma 13). Hence, x_n is a profitable deviation. This contradicts Lemma 15, according to which $\hat{x}\left(\theta\right) \subset \left\{-\infty, x_s\left(\theta\right)\right\}$. Therefore, $\lim_{\sigma\to 0}\bar{F}\left[\frac{\theta^*_\sigma - x_s\left(\theta^*_\sigma\right)}{\sigma}\right] = 1$. ∎

Lemma 16 *For σ sufficiently large, $\hat{x}\left(\theta^*\right) = x_s\left(\theta^*\right)$.*

Proof. From Lemma 15, recall that $\hat{x}\left(\theta^*\right) \subset \left\{-\infty, x_s\left(\theta^*\right)\right\}$. Clearly, all contestants playing $x = -\infty$ is inconsistent with equilibrium, because $0 < m < 1$. Hence, in equilibrium, some contestants must be playing $x_s\left(\theta^*\right)$.

Next, observe that the equilibrium payoff of playing $x_s\left(\theta^*\right)$ is at least

$$
\pi^*\left[x_s\left(\theta^*\right), \theta^*\right] \geq vm - c\left[x_s\left(\theta^*\right)\right] \geq 0 , \tag{21}
$$

where the lower bound is reached if all other contestants also play $x_s\left(\theta^*\right)$—rather than $-\infty$.

Recall that $x_s\left(\theta^*\right)$ satisfies the FOC

$$
c'\left[x_s\left(\theta^*\right)\right] = \frac{v}{\sigma}f\left[\frac{\theta^* - x_s\left(\theta^*\right)}{\sigma}\right] .
$$

Clearly, the RHS goes to zero when $\sigma \to \infty$. Because $c'\left(-\infty\right) = 0$ and $c'\left(\cdot\right)$ is strictly increasing, it must be that, on the LHS, $x_s\left(\theta^*\right) \to -\infty$. Therefore, $\pi^*\left[x_s\left(\theta^*\right), \theta^*\right] \to mv > 0$ as $\sigma \to 0$. We may conclude that dropping out is not a best response. Hence, $\hat{x}\left(\theta^*\right) = x_s\left(\theta^*\right)$.

This completes the proof. ∎

Lemma 17 *For σ sufficiently small, $\hat{x}\left(\theta^*\right) = \left\{-\infty, x_s\left(\theta^*\right)\right\}$. Both $x = -\infty$ and $x = x_s\left(\theta^*\right)$ are played with strictly positive probabilities. Equilibrium payoffs are zero.*

Proof. From Lemma 2 we know that $\hat{x}(\theta^*) \subset \{-\infty, x_s(\theta^*)\}$. Notice that all contestants playing $x = -\infty$ with probability 1 is inconsistent with equilibrium, because $0 < m < 1$. For σ sufficiently small, everybody playing $x = x_s(\theta^*)$ is likewise inconsistent with equilibrium. To see this, recall that $\lim_{\sigma \to 0} \bar{F}\left[\frac{\theta^* - x_s(\theta^*)}{\sigma}\right] = 1$ (Lemma 15), which contradicts market clearing, since $m < 1$.

We may conclude that, for small σ, $\hat{x}(\theta^*) = \{-\infty, x_s(\theta^*)\}$ and that both elements are played with strictly positive probability. Since $\theta^* > -\infty$, we have $\pi^*[-\infty, \theta^*] = 0$. This implies that equilibrium payoff must be zero. ∎

Recall that $\hat{\gamma}^*$ denotes the probability that $x_s(\theta^*)$ is played in equilibrium.

Lemma 18 (i) If $\hat{x}(\theta^*) = x_s(\theta^*)$, then

$$\frac{dx_s(\theta^*)}{d\sigma/\sigma} = -\frac{1}{c''/c'} < 0 .$$

(ii) If $\hat{x}(\theta^*) = \{-\infty, x_s(\theta^*)\}$, then

$$\frac{dx_s(\theta^*)}{d\sigma/\sigma} = -\frac{1}{-soc} < 0 ,$$

$$\frac{d\hat{\gamma}^*/\hat{\gamma}^*}{d\sigma/\sigma} = \frac{c'/c}{-soc} > 0 ,$$

$$\lim_{\sigma \to 0} \hat{\gamma}^* = m .$$

Proof.
Part (i): If $\hat{x}(\theta^*) = x_s(\theta^*)$, then equilibrium is characterized by the following FOC and market-clearing condition:

$$\frac{v}{\sigma} f\left(\frac{\theta^* - x_s}{\sigma}\right) = c'(x_s) ,$$

$$\bar{F}\left(\frac{\theta^* - x_s}{\sigma}\right) = m .$$

Implicitly differentiating the FOC with respect to σ, re-applying the FOC, and solving for $dx_s(\theta^*, \sigma)/d\sigma$ yields

$$\frac{dx_s(\theta^*, \sigma)}{d\sigma} = \frac{f'/f \cdot \left(\frac{d\theta^*}{d\sigma} - \frac{\theta^* - x_s}{\sigma}\right) - 1}{f'/f + \sigma c''/c'} . \tag{22}$$

Implicitly differentiating the market-clearing condition with respect to σ yields

$$\frac{d\theta^*}{d\sigma} = \frac{dx_s}{d\sigma} + \frac{\theta^* - x_s}{\sigma} . \tag{23}$$

Substituting (23) back into (22) and solving we find

$$\frac{\mathrm{d}x_s\left(\theta^*,\sigma\right)}{\mathrm{d}\sigma/\sigma} = -\frac{1}{c''/c'} < 0 \ , \tag{24}$$

where the strict inequality follows from Lemma 1.

Part (ii): If $\hat{x}\left(\theta^*\right) = \{-\infty, x_s\left(\theta^*\right)\}$, then equilibrium is characterized by the following FOC, market-clearing, and zero-profit conditions:

$$\frac{\upsilon}{\sigma}f\left(\frac{\theta^*-x_s}{\sigma}\right) = c'\left(x_s\right) \ ,$$

$$\hat{\gamma}^*\bar{F}\left(\frac{\theta^*-x_s}{\sigma}\right) = m \ ,$$

$$\upsilon\bar{F}\left(\frac{\theta^*-x_s}{\sigma}\right) = c\left(x_s\right) \ .$$

Implicitly differentiating the FOC with respect to σ and solving for $\mathrm{d}x_s\left(\theta^*,i\right)/\mathrm{d}\sigma$ once again yields (22) above.

Implicitly differentiating the zero-profit condition and solving for $\mathrm{d}\theta^*/\mathrm{d}\sigma$ yields

$$\frac{\mathrm{d}\theta^*}{\mathrm{d}\sigma} = \frac{\theta^*-x_s}{\sigma} \ . \tag{25}$$

Substituting (25) back into (22), we find

$$\frac{\mathrm{d}x_s}{\mathrm{d}\sigma/\sigma} = \frac{1}{soc} < 0 \ . \tag{26}$$

where soc is defined as in eqn. (4).

Implicitly differentiating the market clearing condition and applying the FOC, (25) and (26), we get:

$$\frac{\mathrm{d}\hat{\gamma}^*/\hat{\gamma}^*}{\mathrm{d}\sigma/\sigma} = \frac{c'}{-soc}\frac{\hat{\gamma}^*}{m\upsilon} \ .$$

Together, market clearing and zero profit conditions imply that
$\hat{\gamma}^* = m/\bar{F}\left(\frac{\theta^*-x_s}{\sigma}\right) = m\upsilon/c\left(x_s\right)$. Hence,

$$\frac{\mathrm{d}\hat{\gamma}^*/\hat{\gamma}^*}{\mathrm{d}\sigma/\sigma} = \frac{c'/c}{-soc} > 0 \ . \tag{27}$$

Finally, $\lim_{\sigma\to 0}\hat{\gamma}^* = m$ follows immediately from Lemma 15 and the market clearing condition. ∎

We use Lemma 18 to establish single-crossing in σ.

Lemma 19 *There exists a unique $\underline{\sigma} > 0$ such that*

$$\hat{x}(\theta^*) = \begin{cases} \{-\infty, x_s(\theta^*)\} & \text{iff} \quad \sigma \leq \underline{\sigma} \\ x_s(\theta^*) & \text{iff} \quad \sigma > \underline{\sigma} \end{cases} .$$

Proof. If $\hat{x}(\theta^*) = x_s(\theta^*)$, then all agents strictly prefer to participate, such that the equilibrium payoff is

$$\pi^*[\hat{x}(\theta^*), \theta^*] = vm - c[x_s(\theta^*)] > 0 . \tag{28}$$

From Lemma 18 we know that $\mathrm{d}x_s(\theta_\sigma^*)/\mathrm{d}\sigma < 0$. Together with (28), this implies that $\mathrm{d}\pi^*/\mathrm{d}\sigma > 0$. Notice, however, that $\mathrm{d}\pi^*/\mathrm{d}\sigma > 0$ precludes profits falling back to zero once they are in strictly positive territory. Together with Lemmas 17 and 16, this proves the lemma. $\blacksquare$

Finally, we determine the value of $\underline{\sigma}$:

Lemma 20

$$\underline{\sigma} = \frac{vf[\bar{F}^{-1}(m)]}{c'[c^{-1}(mv)]} .$$

Proof. Observe that $\underline{\sigma}$ is such that, for $x = x_s(\theta^*)$: 1) the FOC is satisfied; 2) the equilibrium payoff is zero; and 3) the market clears with all contestants participating. That is,

$$\frac{v}{\sigma} f\left(\frac{\theta - x_s}{\sigma}\right) = c'(x_s) ,$$

$$v\bar{F}\left(\frac{\theta - x_s}{\sigma}\right) = c(x_s) ,$$

$$\bar{F}\left(\frac{\theta - x_s}{\sigma}\right) = m .$$

Using the market clearing and zero profit conditions, we find

$$\frac{\theta - x_s}{\sigma} = \bar{F}^{-1}(m) ,$$

$$c^{-1}\left[v\bar{F}\left(\frac{\theta - x_s}{\sigma}\right)\right] = x_s .$$

Substituting these expressions into the FOC, we get

$$\frac{v}{\sigma} f[\bar{F}^{-1}(m)] = c'\left(c^{-1}\left\{v\bar{F}[\bar{F}^{-1}(m)]\right\}\right) .$$

Solving for σ yields the result. $\blacksquare$

This completes the proof of Proposition 1.

Limits of Meritocracy   Proof of Lemma 4: First we consider $\sigma > \underline{\sigma}$. In that case, $\hat{x}(\theta^*) = x_s(\theta^*)$ and all contestants participate. Hence, $O^* = e^{x_s}$. From Lemma 18 we know that $\mathrm{d}x_s(\theta^*, \sigma)/\mathrm{d}\sigma = -1/(\sigma c''/c') < 0$. Therefore,

$$\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} = \frac{e^{x_s}}{O^*}\frac{\mathrm{d}x_s}{\mathrm{d}\sigma/\sigma} = -\frac{1}{c''/c'} = -\frac{1}{\eta_{C'} + 1} \ .$$

Next we consider $\sigma < \underline{\sigma}$. In that case, $\hat{x}(\theta^*) = \{-\infty, x_s(\theta^*)\}$ and output is given by $O^* = \hat{\gamma}^* e^{x_s}$. Here, $\hat{\gamma}^*$ denotes the equilibrium probability that $x_s(\theta^*)$ is played. From Lemma 18 we know that, now, $\mathrm{d}x_s(\theta^*_\sigma)/(\mathrm{d}\sigma/\sigma) = -\frac{1}{-soc} < 0$, and $(\mathrm{d}\hat{\gamma}^*/\hat{\gamma}^*)/\mathrm{d}\sigma = \frac{1}{\sigma}\frac{c'/c}{-soc} > 0$. Hence,

$$\frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} = \frac{\sigma}{\hat{\gamma}^* e^{x_s}}\frac{\mathrm{d}\hat{\gamma}^* e^{x_s}}{\mathrm{d}\sigma} = \frac{\sigma}{\hat{\gamma}^* e^{x_s}}\left(e^{x_s}\frac{\mathrm{d}\hat{\gamma}^*}{\mathrm{d}\sigma} + \hat{\gamma}^* e^{x_s}\frac{\mathrm{d}x_s}{\mathrm{d}\sigma}\right)$$

$$= \frac{\mathrm{d}\hat{\gamma}^*/\hat{\gamma}^*}{\mathrm{d}\sigma/\sigma} + \frac{\mathrm{d}x_s}{\mathrm{d}\sigma/\sigma} = \frac{c'/c - 1}{-soc} = \frac{\eta_C - 1}{-soc} \ .$$

This completes the proof.   ■

Multiple Prize Levels   Proof of Proposition 3: We prove the proposition by showing that the optimal single-prize-level contest is first-best from the organizer's point of view. That is: 1) he appropriates all the rents; and 2) production is efficient.

To see that 1) holds, notice that at $\sigma = \underline{\sigma}$, all agents participate, and they earn zero profits. Therefore, the organizer appropriates all the rents. To see that 2) holds, recall that, at $\underline{\sigma}$, all contestants participate and produce the same output $x_s(\theta^*) > -\infty$. Efficiency of production then follows from the fact that C is strictly convex.

This completes the proof.   ■

Heterogeneous Contestants

Lemma 21 *Fix some $x > -\infty$ and $i \in I$: 1) $c(-\infty, i) = c_x(-\infty, i) = 0$; 2) $c(x, i)$, $c_x(x, i)$, $c_{xx}(x, i) > 0$, while $c_i(x, i)$, $c_{xi}(x, i) < 0$; 3) $c(\cdot, i)$ and $c_x(\cdot, i)$ are weakly log-convex; 4) $\eta_{C_X} + 1 = c_{xx}(x, i)/c_x(x, i) \overset{(=)}{\geq} c_x(x, i)/c(x, i) = \eta_C \overset{(=)}{>} 1$ depending on whether $C_{xx}(\cdot, i) \overset{(=)}{>} 0$; 5) If $C(X) = kX^\alpha \cdot l(i)$, $k > 0$, $\alpha \geq 1$, then $\eta_{C_X} + 1 = \eta_C = \alpha$.*

Proof. The proofs of these claims are either trivial or essentially the same as those in Lemma 1. ■

**Existence and Uniqueness of Equilibrium With Heterogeneous
Contestants**   This section contains a proof of Proposition 4.

Lemma 22 *For all $(\theta, i) \in \mathbb{R} \times I$, $\hat{x}(\theta, i)$ is non-empty, compact-valued, and UHC.*

Moreover, for all $i \in I$, $\lim_{\theta \to \infty} \theta - \hat{x}(\theta, i) = \infty$ and $\lim_{\theta \to -\infty} \theta - \hat{x}(\theta, i) = -\infty$.

Proof. Recall our assumption that, for each $i \in I$, $C(\cdot, i)$ has the same properties as $C(\cdot)$ in the homogeneous model. Therefore, the proof of the lemma is essentially the same as that of Lemma 7. ∎

Lemma 23 *If $\hat{x}(\theta, i)$ is single-valued at $(\theta', i') \in \mathbb{R} \times I$, then it is locally differentiable. Furthermore, $\partial \hat{x}(\theta, i') / \partial \theta|_{\theta = \theta'} < 1$, and $\partial \hat{x}(\theta', i) / \partial i|_{i = i'} \underset{(=)}{>} 0$ if*
$\hat{x}(\theta', i') \underset{(=)}{>} -\infty$.

Proof. The proof of differentiability in θ and $\partial \hat{x}(\theta, i') / \partial \theta|_{\theta = \theta'} < 1$ is identical to the proof of Lemma 8, while the proof of differentiability in i is analogous.

If $\hat{x}(\theta', i') > -\infty$, then we can implicitly differentiate the FOC with respect to i. This yields
$$\left. \frac{\partial \hat{x}(\theta', i)}{\partial i} \right|_{i=i'} = -c_{xi}[\hat{x}(\theta', i'), i'] > 0 \, ,$$
where we have used the envelope theorem and the fact that $c_{xi} < 0$ (Lemma 21). Finally, that $\left. \frac{\partial \hat{x}(\theta', i)}{\partial i} \right|_{i=i'} = 0$ when $\hat{x}(\theta', i') = -\infty$ is trivial. ∎

Lemma 24 *Fix $i' \in I$. Either $\hat{x}(\theta, i') = x_s(\theta, i')$ for all $\theta \in \mathbb{R}$, or there exists a unique $\bar{\theta}_{i'} \in \mathbb{R}$ such that*

$$\hat{x}(\theta, i') = \begin{cases} x_s(\theta, i') & for \quad \theta < \bar{\theta}_{i'} \\ \{-\infty, x_s(\theta, i')\} & for \quad \theta = \bar{\theta}_{i'} \\ -\infty & for \quad \theta > \bar{\theta}_{i'} \end{cases} .$$

Fix $\theta' \in \mathbb{R}$. Either $\hat{x}(\theta', i) = x_s(\theta', i)$ for all $i \in I$, or $\hat{x}(\theta', i) = -\infty$ for all $i \in I$, or there exists a unique $\underline{i}_{\theta'} \in I$ such that

$$\hat{x}(\theta', i) = \begin{cases} x_s(\theta', i) & for \quad i > \underline{i}_{\theta'} \\ \{-\infty, x_s(\theta', i)\} & for \quad i = \underline{i}_{\theta'} \\ -\infty & for \quad i < \underline{i}_{\theta'} \end{cases} .$$

Proof. Notice that, for fixed $i \in I$, Lemmas 2, 9, and 10 carry over unchanged. This proves the claimed structure of $\hat{x}(\theta, i')$ as a function of θ.

That $\hat{x}(\theta', i)$ takes on the form claimed in the lemma follows from Lemma 2 and the fact that

$$\frac{\mathrm{d}}{\mathrm{d}i}\pi\left[\hat{x}(\theta', i), \theta', i\right] = -c_i\left[\hat{x}(\theta', i), i\right] > 0 \ ,$$

where we have used the envelope theorem and the fact that $c_i < 0$ (Lemma 21). ∎

Lemma 24 implies:

Corollary 1 *Fix some $\theta' \in \mathbb{R}$. For almost all $i \in I$, $\hat{x}(\theta', i)$ is single-valued and differentiable in i, as well as differentiable in θ at θ'. Around an ability i' where $\hat{x}(\theta', i')$ is multi-valued, $\hat{x}(\theta', i')$ jumps upward from $-\infty$ to $x_s(\theta', i)$, as i increases.*

Fix some $i' \in I$. For almost all $\theta \in \mathbb{R}$, $\hat{x}(\theta, i')$ is single-valued and differentiable in θ, as well as differentiable in i at i'. Around a standard θ' where $\hat{x}(\theta', i')$ is multi-valued, $\hat{x}(\theta, i')$ jumps downward from $x_s(\theta, i')$ to $-\infty$, as θ increases.

As in the baseline, $\Omega(\theta)$ is defined as the set of masses of winners, w, that can arise when agents best respond to standard θ. The inverse, $\Omega^{-1}(w)$, denotes the set of standards that give rise to a mass w of winners. We treat the correspondences $\Omega(\cdot)$ and $\Omega^{-1}(\cdot)$ as functions when they are single-valued.

We use Corollary 1 to show:

Lemma 25 $\Omega(\theta)$ *is single-valued, differentiable, and strictly decreasing in θ for all $\theta \in \mathbb{R}$. Moreover, $\lim_{\theta \to \infty} \Omega(\theta) = 0$ and $\lim_{\theta \to -\infty} \Omega(\theta) = 1$.*

Proof. Without loss of generality, we may assume that i is uniformly distributed. Hence, $G(i) = i$.

Corollary 1 implies that, for a given θ, the best-response profile $\hat{\chi}(\theta)$ is uniquely determined up to a zero measure of types $i = \underline{i}_\theta$. Hence, for all $\theta > -\infty$, $\Omega(\theta)$ is single-valued and equal to

$$\Omega(\theta) = W[\theta, \hat{\chi}(\theta)] = \int_{i \in I \setminus \{\underline{i}_\theta\}} \bar{F}\left[\frac{\theta - \hat{x}(\theta, i)}{\sigma}\right] \mathrm{d}i \ ,$$

where we have simply ignored the zero measure of contestants of type $\underline{i}_\theta$ for whom $\hat{x}(\theta, i)$ is multi-valued.

Provided the derivative exists,

$$\begin{aligned}
\frac{\mathrm{d}\Omega(\theta)}{\mathrm{d}\theta} &= \frac{\mathrm{d}}{\mathrm{d}\theta}\int_{i \in I \setminus \{\underline{i}_\theta\}} \bar{F}\left[\frac{\theta - \hat{x}(\theta, i)}{\sigma}\right] \mathrm{d}i \\
&= -\int_{i \in I \setminus \{\underline{i}_\theta\}} \frac{1}{\sigma}f\left[\frac{\theta - \hat{x}(\theta, i)}{\sigma}\right]\left[1 - \frac{\mathrm{d}\hat{x}(\theta, i)}{\mathrm{d}\theta}\right] \mathrm{d}i \ .
\end{aligned}$$

From Corollary 1 we know that at all θ, $\hat{x}(\theta, i)$ is differentiable in θ for almost all i. Therefore, $\Omega(\theta)$ is differentiable for all θ. Moreover, because $\mathrm{d}\hat{x}(\theta, i)/\mathrm{d}\theta < 1$ by Lemma 23, it is easily verified that $\mathrm{d}\Omega(\theta)/\mathrm{d}\theta < 0$.

Finally, $\lim_{\theta \to \infty} \Omega(\theta) = 0$ and $\lim_{\theta \to -\infty} \Omega(\theta) = 1$ follow from the second part of Lemma 22, where it was shown that $\lim_{\theta \to \infty} \theta - \hat{x}(\theta, i) = \infty$ and $\lim_{\theta \to -\infty} \theta - \hat{x}(\theta, i) = -\infty$. ∎

Lemma 26 *With heterogeneous contestants, equilibrium exists and is unique.*

Proof. Lemma 25 and the intermediate value theorem imply that $\Omega(\theta) = m$ has a unique solution $\theta^* \in \mathbb{R}$. In turn, Corollary 1 implies that the associated best-response profile $\hat{\chi}(\theta^*)$ is uniquely determined up to a zero measure of contestants $i \in I$. Hence, an equilibrium $(\hat{\chi}(\theta^*), \theta^*)$ exists and is (essentially) unique. ∎

Lemma 27 *Participation satisfies single-crossing in ability.*

That is, if $i' \in I$ is willing to participate (drop out) then $i \overset{(<)}{>} i'$ participates (drops out) for sure.

Proof. The result follows from $c_{xi}(\cdot, \cdot) < 0$, i.e., submodularity of x and i (Lemma 21). Specifically, if i' is willing to participate and $i > i'$, then

$$0 \leq \pi\left[\hat{x}(\theta, i'), \theta, i'\right] \leq \pi\left[\hat{x}(\theta, i'), \theta, i\right] < \pi\left[\hat{x}(\theta, i), \theta, i\right] ,$$

where the second inequality follows from $c_i(\cdot, \cdot) < 0$ and the last inequality follows from $c_{xi}(\cdot, \cdot) < 0$. Hence, $i > i'$ participates for sure.

Similarly, if i' is willing to drop out and $i < i'$, then

$$0 \geq \pi\left[\hat{x}(\theta, i'), \theta, i'\right] > \pi\left[\hat{x}(\theta, i), \theta, i'\right] > \pi\left[\hat{x}(\theta, i), \theta, i\right] .$$

Hence, $i < i'$ drops out for sure. ∎

Recall that $\underline{i} \in I$ denotes a type who is indifferent between participating and dropping out.

Lemma 28 *For $\sigma > 0$ sufficiently small, $\underline{i}$ exists. Moreover, $\lim_{\sigma \to 0} \underline{i} = 1 - m$.*

Proof.
Claim 1: Fix $(\theta, i) \in \mathbb{R} \times I$. For σ sufficiently small, the marginal cost $c_x(\cdot, i)$ intersects marginal revenue $\frac{v}{\sigma} f\left(\frac{\theta-x}{\sigma}\right)$ exactly twice, first from above at $x_t(\theta, i) < \theta$, then from below at $x_s(\theta, i) > \theta$.

Proof of Claim 1: Identical to the proof of Lemma 41.

Claim 2: Fix a $(\theta, i) \in \mathbb{R} \times I$. For σ sufficiently small, $\hat{x}(\theta, i) \subset \{-\infty, x_s(\theta, i)\}$.

Proof of Claim 2: Corollary of Claim 1.

Claim 3: When $\sigma \to 0$, the winning probability of all non-drop-outs goes to 1.

Proof of Claim 3: Observe that, for fixed i, Lemmas 13, 14, and, hence, Lemma 15 carry over unchanged. This proves the claim.

Together with market clearing, Claim 3 implies that the mass of participants converges to $m < 1$ when $\sigma \to 0$. Because π^* is continuous and strictly increasing in i, $\underline{i}$ must indeed exist and $\lim_{\sigma \to 0} \underline{i} = 1 - m$. ∎

More Notation:

Let $\bar{F}(\underline{i}) \equiv \bar{F}\left[\frac{\theta^* - x_s(\theta^*, \underline{i})}{\sigma}\right]$ and $f(\underline{i}) \equiv f\left[\frac{\theta^* - x_s(\theta^*, \underline{i})}{\sigma}\right]$, while $f \equiv f\left[\frac{\theta^* - x_s(\theta^*, i)}{\sigma}\right]$; i.e., for generic i we drop the argument of f, but write $f(\underline{i})$ when $i = \underline{i}$. Along the same lines, $c(\underline{i}) \equiv c[x_s(\theta^*, \underline{i}), \underline{i}]$, $c_x(\underline{i}) \equiv c_x[x_s(\theta^*, \underline{i}), \underline{i}]$, and $c_i(\underline{i}) \equiv c_i[x_s(\theta^*, \underline{i}), \underline{i}]$, while $c_{xx}/c_x \equiv c_{xx}[x_s(\theta^*, i), i]/c_x[x_s(\theta^*, i), i]$. Also, let $soc \equiv -\left[\frac{1}{\sigma} f'/f + c_{xx}/c_x\right] < 0$. Finally, depending on the context, we may drop one or more of the arguments of $x_s(\theta, i; \sigma)$.

Lemma 29 *Suppose $\underline{i}$ exists. For all $\sigma > 0$, $d\underline{i}/d\sigma < 0$.*

Proof. WLOG, we once more assume that i is uniformly distributed. Hence $G(i) = i$.

Analogous to (22), implicitly differentiating the FOC with respect to σ and solving for $dx_s(\theta^*_\sigma)/d\sigma$ yields

$$\frac{dx_s(i)}{d\sigma} = \frac{1}{\sigma} \frac{f'/f \cdot \left[\frac{d\theta^*}{d\sigma} - \frac{\theta^* - x_s(i)}{\sigma}\right] - 1}{-soc}, \tag{29}$$

where we have suppressed the dependence of x_s on θ^* and σ.

By definition, a contestant of ability $\underline{i}$ is indifferent between participating and dropping out, i.e.,

$$\bar{F}(\underline{i}) v - c(\underline{i}) = 0. \tag{30}$$

Implicitly differentiating (30) with respect to σ while using the envelope theorem yields:

$$-\frac{v}{\sigma}f\left(\underline{i}\right)\left[\frac{\mathrm{d}\theta^*}{\mathrm{d}\sigma}-\frac{\theta^*-x_s\left(\underline{i}\right)}{\sigma}\right]=c_i\left(\underline{i}\right)\frac{\mathrm{d}\underline{i}}{\mathrm{d}\sigma}\ .$$

Applying the FOC and solving for $\mathrm{d}\underline{i}/\mathrm{d}\sigma$ we find

$$\frac{\mathrm{d}\underline{i}}{\mathrm{d}\sigma}=\frac{c_x\left(\underline{i}\right)}{-c_i\left(\underline{i}\right)}\left[\frac{\mathrm{d}\theta^*}{\mathrm{d}\sigma}-\frac{\theta^*-x_s\left(\underline{i}\right)}{\sigma}\right]\ . \tag{31}$$

Implicitly differentiating the market clearing condition with respect to σ gives

$$-\int_{\underline{i}}^1\left[\frac{\mathrm{d}\theta^*/\mathrm{d}\sigma-\mathrm{d}x_s\left(i\right)/\mathrm{d}\sigma}{\sigma}-\frac{\theta^*-x_s\left(i\right)}{\sigma^2}\right]f\mathrm{d}i-\bar{F}\left(\underline{i}\right)\frac{\mathrm{d}\underline{i}}{\mathrm{d}\sigma}=0\ . \tag{32}$$

Next, we substitute into (32) the expressions for $\mathrm{d}x_s\left(i\right)/\mathrm{d}\sigma$ and $\mathrm{d}\underline{i}/\mathrm{d}\sigma$ given in (29) and (31), respectively. Solving for $\mathrm{d}\theta^*/\mathrm{d}\sigma$ then yields:

$$\frac{\mathrm{d}\theta^*}{\mathrm{d}\sigma}=\frac{1}{\sigma}\frac{\int_{\underline{i}}^1\frac{1}{\sigma}\frac{\left[\theta^*-x_s(i)\right]\frac{c_{xx}}{c_x}-1}{-soc}f\mathrm{d}i+\left[\theta^*-x\left(\underline{i}\right)\right]\frac{c_x(\underline{i})}{-c_i(\underline{i})}\bar{F}\left(\underline{i}\right)}{\int_{\underline{i}}^1\frac{1}{\sigma}\frac{1}{-soc}\frac{c_{xx}}{c_x}f\mathrm{d}i+\frac{c_x(\underline{i})}{-c_i(\underline{i})}\bar{F}\left(\underline{i}\right)}\ . \tag{33}$$

Substituting (33) back into (31) and solving for $\mathrm{d}\underline{i}/\mathrm{d}\sigma$ we find

$$\frac{\mathrm{d}\underline{i}}{\mathrm{d}\sigma}=\frac{1}{\sigma}\frac{\int_{\underline{i}}^1\frac{\left[x_s(\underline{i})-x_s(i)\right]\frac{c_{xx}}{c_x}-1}{-soc}f\mathrm{d}i}{\frac{-c_i(\underline{i})}{c_x(\underline{i})}\int_{\underline{i}}^1\frac{1}{-soc}\frac{c_{xx}}{c_x}f\mathrm{d}i+\sigma\bar{F}\left(\underline{i}\right)}\ . \tag{34}$$

Notice that the numerator of (34) is strictly negative, because: 1) $x_s\left(i\right)$ is strictly increasing in ability i; 2) $c_{xx}/c_x>0$; and 3) $-soc>0$. Furthermore, the denominator is strictly positive, because $-c_i\left(\underline{i}\right)/c_x\left(\underline{i}\right)>0$. Hence, $\mathrm{d}\underline{i}/\mathrm{d}\sigma<0$. $\blacksquare$

This completes proof of Proposition 4.

Non-singularity of Baseline Model   Proof of Proposition 5: Let $\{i_n\}_{n\in\mathbb{N}}$ denote a sequence of random abilities $i_n\in I$ that converge in probability to $i_0\in I$. The CDF and PDF of i_n are denoted by $G_n\left(\cdot\right)$ and $g_n\left(\cdot\right)$, respectively. The degenerate distribution that puts mass 1 on $i=i_0$ is denoted by $\mathbf{1}_{\{i=i_0\}}$

Claim 1: For all $\theta\in\mathbb{R}$, $\Omega_{G_n}\left(\theta\right)\overset{i_n\overset{\mathrm{Pr}}{\to}i_0}{\longrightarrow}\Omega_{\mathbf{1}_{\{i=i_0\}}}\left(\theta\right)$.

Proof of Claim 1: Fix some $\theta\in\mathbb{R}$. From Lemma 24 we know that $\hat{x}\left(\theta,i\right)$ is single-valued for all $i\in I\setminus\{\underline{i}_\theta\}$. First, suppose that $\underline{i}_\theta\neq i_0$ (or, equivalently, that

$\theta \neq \bar{\theta}_{i_0}$). Then $\hat{x}(\theta, i_0) \in \{-\infty, x_s(\theta, i_0)\}$ is single-valued. Once more applying Lemma 24, we may conclude that, in a neighborhood of i_0, $\hat{x}(\theta, i)$ is single-valued and continuous in i. Furthermore: if $\hat{x}(\theta, i_0) = x_s(\theta, i_0)$ then $\hat{x}(\theta, i) = x_s(\theta, i)$; and if $\hat{x}(\theta, i_0) = -\infty$ then $\hat{x}(\theta, i) = -\infty$.

Convergence in probability to i_0 means that $\lim_{n \to \infty} \Pr(|i_0 - i_n| > \varepsilon) = 0$. I.e., ever more probability mass comes to lie in a ε-neighborhood of i_0. Hence, when $i_n \xrightarrow{\Pr} i_0$,

$$\Omega_{G_n}(\theta) = \int_{i \in I \backslash \{\underline{i}_\theta\}} \bar{F}\left[\frac{\theta - \hat{x}(\theta, i)}{\sigma}\right] dG_n(i)$$

$$\xrightarrow{i_n \xrightarrow{\Pr} i_0} \begin{cases} \bar{F}\left[\frac{\theta - x_s(\theta, i_0)}{\sigma}\right] & \text{if} \quad \hat{x}(\theta, i_0) = x_s(\theta, i_0) \\ 0 & \text{if} \quad \hat{x}(\theta, i_0) = -\infty \end{cases}$$

$$= \Omega_{\mathbf{1}_{\{i = i_0\}}}(\theta) .$$

Alternatively, suppose that $\underline{i}_\theta = i_0$ or, equivalently, $\theta = \bar{\theta}_{i_0}$. In that case, $\hat{x}(\bar{\theta}_{i_0}, i_0) = \{-\infty, x_s(\bar{\theta}_{i_0}, i_0)\}$. From Lemma 24 we then know that, in a left neighborhood of i_0, $\hat{x}(\bar{\theta}_{i_0}, i) = -\infty$, while in a right neighborhood $\hat{x}(\bar{\theta}_{i_0}, i) = x_s(\bar{\theta}_{i_0}, i)$. Therefore,

$$\int_{i \in I \backslash \{\underline{i}_0\}} \bar{F}\left[\frac{\theta - \hat{x}(\bar{\theta}_{i_0}, i)}{\sigma}\right] dG_n(i) = \int_{i = i_0}^{1} \bar{F}\left[\frac{\theta - \hat{x}(\bar{\theta}_{i_0}, i)}{\sigma}\right] dG_n(i)$$

$$\xrightarrow{i_n \xrightarrow{\Pr} i_0} \gamma \bar{F}\left[\frac{\bar{\theta}_{i_0} - x_s(\bar{\theta}_{i_0}, i_0)}{\sigma}\right] ,$$

where $\gamma = \lim_{n \to \infty} 1 - G_n(i_0)$. When $i_n \xrightarrow{\Pr} i_0$, $\lim_{n \to \infty} G_n(i_0)$ can take on any value in $[0, 1]$. It follows that

$$\Omega_{G_n}(\bar{\theta}_{i_0}) \xrightarrow{i_n \xrightarrow{\Pr} i_0} \left[0, \bar{F}\left[\frac{\theta - x_s(\bar{\theta}_{i_0}, i_0)}{\sigma}\right]\right] = \Omega_{i_0}(\bar{\theta}_{i_0}) .$$

This completes the proof of Claim 1.

Claim 2: $\theta_{G_n}^* \xrightarrow{i_n \xrightarrow{\Pr} i_0} \theta_{\mathbf{1}_{\{i = i_0\}}}^* \leq \bar{\theta}_{i_0}$

Proof of Claim 2: From Claim 1 it immediately follows that the equilibrium standards converge. That $\theta_{\mathbf{1}_{\{i = i_0\}}}^* \leq \bar{\theta}_{i_0}$ follows from the fact that for $\theta > \bar{\theta}_{i_0}$ the mass of winners is zero, violating market clearing. This proves Claim 2.

Claim 3: $O_{G_n}^* \xrightarrow{i_n \xrightarrow{\Pr} i_0} O_{\mathbf{1}_{\{i = i_0\}}}^*$

Proof of Claim 3: First suppose that $\theta_{\mathbf{1}_{\{i = i_0\}}}^* < \bar{\theta}_{i_0}$. Then the equilibrium of the homogeneous contest is in pure strategies, with all contestants playing

$x_s\left(\theta^*_{\mathbf{1}_{\{i=i_0\}}}, i_0\right)$. For a heterogeneous contest whose population converges to i_0 in probability, Lemma 24 implies that, for n sufficiently large, in a neighborhood of i_0, $\hat{x}\left(\theta^*_{G_n}, i\right) = x_s\left(\theta^*_{G_n}, i\right)$. Joint continuity of $x_s\left(\cdot, \cdot\right)$ in θ and i then implies that

$$O^*_{G_n} = \int_0^1 \exp\left[\hat{x}\left(\theta^*_{G_n}, i\right)\right] \mathrm{d}G_n\left(i\right) \xrightarrow{i_n \overset{\mathrm{Pr.}}{\to} i_0} \exp\left[x_s\left(\theta^*_{\mathbf{1}_{\{i=i_0\}}}, i_0\right)\right] = O^*_{\mathbf{1}_{\{i=i_0\}}}.$$

Alternatively, suppose that $\theta^*_{\mathbf{1}_{\{i=i_0\}}} = \bar{\theta}_{i_0}$. Then the equilibrium of the homogeneous contest is in mixed strategies, with a fraction $\gamma \equiv m/\bar{F}\left[\theta^*_{\mathbf{1}_{\{i=i_0\}}} - x_s\left(\theta^*_{\mathbf{1}_{\{i=i_0\}}}, i_0\right)/\sigma\right]$ playing $x_s\left(\theta^*_{\mathbf{1}_{\{i=i_0\}}}, i_0\right)$, and $1 - \gamma$ dropping out. For a heterogeneous contest whose population converges to i_0 in probability, consider the sequence of boundary types $\left\{\underline{i}_{\theta^*_{G_n}}\right\}_{n \in \mathbb{N}}$ for whom, by definition, $\hat{x}\left(\theta^*_{G_n}, \underline{i}_{\theta^*_{G_n}}\right) = \left\{-\infty, x_s\left(\theta^*_{G_n}, \underline{i}_{\theta^*_{G_n}}\right)\right\}$. Lemma 24 implies that, in a right neighborhood of $\underline{i}_{\theta^*_{G_n}}$, $\hat{x}\left(\theta^*_{G_n}, i\right) = x_s\left(\theta^*_{G_n}, i\right)$, while in a left neighborhood of $\underline{i}_{\theta^*_{G_n}}$, $\hat{x}\left(\theta^*_{G_n}, i\right) = -\infty$.

By continuity of $\underline{i}_\theta$ in θ and $\theta^*_{G_n} \xrightarrow{i_n \overset{\mathrm{Pr.}}{\to} i_0} \theta^*_{\mathbf{1}_{\{i=i_0\}}}$ (Claim 2), we have $\underline{i}_{\theta^*_{G_n}} \xrightarrow{i_n \overset{\mathrm{Pr.}}{\to} i_0} \underline{i}_{\theta^*_{\mathbf{1}_{\{i=i_0\}}}}$. Because $\theta^*_{\mathbf{1}_{\{i=i_0\}}} = \bar{\theta}_{i_0}$ by assumption, this implies that

$$\underline{i}_{\theta^*_{G_n}} \xrightarrow{i_n \overset{\mathrm{Pr.}}{\to} i_0} \underline{i}_{\bar{\theta}_{i_0}} = i_0,$$

where the equality holds by definition of $\underline{i}_\theta$ and $\bar{\theta}_{i_0}$.

Next notice that the mass of winners converges to

$$\int_0^1 \bar{F}\left[\frac{\theta^*_{G_n} - \hat{x}\left(\theta^*_{G_n}, i\right)}{\sigma}\right] \mathrm{d}G_n\left(i\right)$$

$$= \int_{\underline{i}_{\theta^*_{G_n}}}^1 \bar{F}\left[\frac{\theta^*_{G_n} - x_s\left(\theta^*_{G_n}, i\right)}{\sigma}\right] \mathrm{d}G_n\left(i\right)$$

$$\xrightarrow{i_n \overset{\mathrm{Pr.}}{\to} i_0} \lim_{n \to \infty}\left[1 - G_n\left(i_0\right)\right] \cdot \bar{F}\left[\frac{\theta^*_{\mathbf{1}_{\{i=i_0\}}} - x_s\left(\theta^*_{\mathbf{1}_{\{i=i_0\}}}, i_0\right)}{\sigma}\right].$$

When $i_n \xrightarrow{\mathrm{Pr}} i_0$, in principle, $\lim_{n \to \infty} 1 - G_n\left(i_0\right)$ can take on any value in $[0, 1]$. However, market clearing requires that $\lim_{n \to \infty} 1 - G_n\left(i_0\right) = \gamma$, where γ is as defined above.

Together, these observations imply

$$
\begin{aligned}
O_{G_n}^* \;=\;& \int_0^1 \exp\left[\hat{x}\left(\theta_{G_n}^*,i\right)\right] \mathrm{d}G_n\left(i\right) = \int_{\underline{i}_{\theta_{G_n}^*}}^1 \exp\left[x_s\left(\theta_{G_n}^*,i\right)\right]\mathrm{d}G_n\left(i\right) \\
& \overset{i_n \overset{\mathrm{Pr}}{\to} i_0}{\longrightarrow} \lim_{n\to\infty}\left[1 - G_n\left(i_0\right)\right]\exp\left[x_s\left(\bar{\theta}_{i_0},i_0\right)\right] \\
=\;& \gamma\exp\left[x_s\left(\bar{\theta}_{i_0},i_0\right)\right] = O_{\mathbf{1}_{\{i=i_0\}}}^* \;.
\end{aligned}
$$

This completes the proof of Claim 3.

Claim 4: $\hat{\sigma} \overset{i \overset{\mathrm{Pr}}{\to} i_0}{\longrightarrow} \underline{\sigma}_{i_0} > 0$

Proof of Claim 4: From Claim 3 we know that, for every σ, the aggregate output in the heterogeneous contest converges to that in the homogeneous contest. Furthermore, because $C_{XX}\left(\cdot,\cdot\right) > 0$, Proposition 1 implies that the homogeneous contest has a unique output maximizing $\sigma = \underline{\sigma}_{i_0}$. Together these observations prove Claim 4.

This completes the proof of Proposition 5.   ∎

Very Heterogeneous Contests   Proof of Theorem 2: Let $-C_{Xi}/C_X > \beta$ for all $\left(X,i\right) \in \left[0,\infty\right) \times I$, and some $\beta > 0$. That is,

$$
\eta_{C_X,i} = \frac{-C_{Xi}}{C_X} = \frac{-XC_{Xi}}{XC_X} = \frac{-c_{xi}}{c_x} > \beta \;.
$$

Let $c^{-1}\left[\cdot,1-m\right]$ denote the inverse of $c\left[x,1-m\right]$ with respect to x.

Claim 1:
$$
\lim_{\sigma\to 0} O^*\left(\sigma\right) = me^{c^{-1}\left[v,1-m\right]} \;.
$$

Proof of Claim 1: From Lemma 28 we know that $\lim_{\sigma\to 0}\underline{i} = 1 - m$. Furthermore, an argument identical to that in Lemma 15 implies that the probability of winning for type $i = 1 - m$ goes to 1 when $\sigma \to 0$. Hence, the limit of the output of the marginal type, $\lim_{\sigma\to 0}x_s\left(1-m\right)$, is characterized by the zero-profit condition

$$
v = c\left[\lim_{\sigma\to 0} x_s\left(1-m\right),1-m\right] \;,
$$

which in turn implies that

$$
\lim_{\sigma\to 0} x_s\left(1-m\right) = c^{-1}\left(v,1-m\right) \;.
$$

Next, an argument identical to that in Lemma 13 yields that, for all $i \geq 1 - m$,

$$\lim_{\sigma \to 0} \theta^*_\sigma - x_s(i) = 0 .$$

Together, these observations imply

$$\lim_{\sigma \to 0} x_s(i) = \lim_{\sigma \to 0} x_s(1 - m) = c^{-1}(v, 1 - m) ,$$

for all $i \geq 1 - m$. Hence,

$$\lim_{\sigma \to 0} O^*(\sigma) = m e^{c^{-1}(v, 1 - m)} .$$

Claim 2: Fix some $\sigma > 0$ but small. Then, for any $i > 1 - m$,

$$\lim_{\beta \to \infty} [x_s(i) - x_s(1 - m)] = \infty .$$

Proof of Claim 2: Dividing the FOCs for $x_s(i)$ and $x_s(1 - m)$ yields

$$\frac{f\left[\frac{\theta^* - x_s(i)}{\sigma}\right]}{f\left[\frac{\theta^* - x_s(1-m)}{\sigma}\right]} = \frac{c_x[x_s(i), i]}{c_x[x_s(1 - m), 1 - m]} .$$

Taking logs on both sides and subtracting $\ln c_x[x_s(i), 1 - m]$, this equality is equivalent to

$$\left\{ \begin{array}{c} \ln f\left[\frac{\theta^* - x_s(i)}{\sigma}\right] - \ln f\left[\frac{\theta^* - x_s(1-m)}{\sigma}\right] \\ + \ln c_x[x_s(1 - m), 1 - m] - \ln c_x[x_s(i), 1 - m] \end{array} \right\} = \ln c_x[x_s(i), i] - \ln c_x[x_s(i), 1 - m] .$$

From log-concavity of f we know that $f'(\cdot)/f(\cdot)$ is strictly decreasing in its argument. Hence,

$$\ln f\left[\frac{\theta^* - x_s(i)}{\sigma}\right] - \ln f\left[\frac{\theta^* - x_s(1 - m)}{\sigma}\right] = \int_{\frac{\theta^* - x_s(1-m)}{\sigma}}^{\frac{\theta^* - x_s(i)}{\sigma}} \frac{f'\left[\frac{\theta^* - x_s(\iota)}{\sigma}\right]}{f\left[\frac{\theta^* - x_s(\iota)}{\sigma}\right]} d\frac{\theta^* - x_s(\iota)}{\sigma}$$

$$> -\frac{\frac{1}{\sigma} f'\left[\frac{\theta^* - x_s(1-m)}{\sigma}\right]}{f\left[\frac{\theta^* - x_s(1-m)}{\sigma}\right]} [x_s(i) - x_s(1 - m)] .$$

From log-convexity of c_x we know that c_{xx}/c_x is weakly increasing in x. Hence,

$$\ln c_x[x_s(1 - m), 1 - m] - \ln c_x[x_s(i), 1 - m] = \int_{x_s(i)}^{x_s(1-m)} \frac{c_{xx}(x, 1 - m)}{c_x(x, 1 - m)} dx$$

$$> -\frac{c_{xx}[x_s(i), 1 - m]}{c_x[x_s(i), 1 - m]} [x_s(i) - x_s(1 - m)] .$$

Furthermore, because $-c_{xi}/c_x > \beta$,

$$\ln c_x\left[x_s\left(i\right),i\right] - \ln c_x\left[x_s\left(i\right),1-m\right] = -\int_{1-m}^{i} \frac{-c_{xi}\left[x_s\left(i\right),\iota\right]}{c_x\left[x_s\left(i\right),\iota\right]} d\iota \leq -\beta\left[i-\left(1-m\right)\right] \ .$$

Together, these inequalities imply that

$$-\left\{\frac{\frac{1}{\sigma}f'\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right]}{f\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right]} + \frac{c_{xx}\left[x_s\left(i\right),1-m\right]}{c_x\left[x_s\left(i\right),1-m\right]}\right\}\left[x_s\left(i\right) - x_s\left(1-m\right)\right]$$

$$< \left\{\begin{array}{c} \ln f\left[\frac{\theta^*-x_s(i)}{\sigma}\right] - \ln f\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right] \\ + \ln c_x\left[x_s\left(1-m\right),1-m\right] - \ln c_x\left[x_s\left(i\right),1-m\right] \end{array}\right\}$$

$$= \ln c_x\left[x_s\left(i\right),i\right] - \ln c_x\left[x_s\left(i\right),1-m\right] \leq -\beta\left[i-\left(1-m\right)\right] \ .$$

Therefore,

$$x_s\left(i\right) - x_s\left(1-m\right) > \beta\frac{i-\left(1-m\right)}{\frac{\frac{1}{\sigma}f'\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right]}{f\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right]} + \frac{c_{xx}\left[x_s(i),1-m\right]}{c_x\left[x_s(i),1-m\right]}} \ .$$

Observe that the denominator of the RHS is positive but bounded from above. To see this, notice that:

1) For small σ, $x_s\left(1-m\right) > \theta^*$ and, hence, $f'/f > 0$;

2) The argument of f'/f, $\frac{\theta^*-x_s(1-m)}{\sigma}$, must remain bounded when $\beta \to \infty$. Otherwise, the probability of winning for all $i \geq 1-m$ goes to one and, because $\sigma > 0$, the mass of winners with abilities $i < 1-m$ remains bounded away from zero. As a result, the mass of winners strictly exceeds m, which violates market clearing. Contradiction.

3) $c_{xx}\left[x_s\left(i\right),1-m\right]/c_x\left[x_s\left(i\right),1-m\right] > 0$ and remains bounded as $\beta \to \infty$.

Together, 1)–3) imply that

$$\lim_{\beta\to\infty}\left[x_s\left(i\right) - x_s\left(1-m\right)\right] = \infty \ .$$

This completes the proof of Claim 2.

Let $x_{\sigma=0} \equiv \lim_{\sigma\to0} x_s\left(1-m\right) = c^{-1}\left(v,1-m\right)$. That is, $x_{\sigma=0}$ denotes the output of the marginal participant in the perfectly meritocratic contest—which is in fact equal to that of all other participants.

Claim 3: When $\beta \to \infty$, the difference $x_{\sigma=0} - x_s\left(1-m\right)$ remains bounded from above.

Proof of Claim 3: Dividing the zero-profit condition determining $x_{\sigma=0}$ by the FOC for $x_s(1-m)$, we find that

$$\frac{1}{\frac{1}{\sigma}f\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right]} = \frac{c(x_{\sigma=0},1-m)}{c_x[x_s(1-m),1-m]} \ .$$

By convexity of c,

$$c(x_{\sigma=0},1-m) > c[x_s(1-m),1-m] + c_x[x_s(1-m),1-m][x_{\sigma=0}-x_s(1-m)] \ .$$

Hence,

$$\frac{1}{\frac{1}{\sigma}f\left[\frac{\theta^*-x_s(1-m)}{\sigma}\right]} = \frac{c(x_{\sigma=0},1-m)}{c_x[x_s(1-m),1-m]} > \frac{c[x_s(1-m),1-m]}{c_x[x_s(1-m),1-m]}+[x_{\sigma=0}-x_s(1-m)] \ .$$

Recall that $\frac{\theta^*-x_s(1-m)}{\sigma}$ remains bounded when $\beta \to \infty$. Since $c/c_x > 0$, this implies that the difference $x_{\sigma=0}-x_s(1-m)$ must remain bounded from above when $\beta \to \infty$.

This completes the proof of Claim 3.

Finally, from Lemma 29 we know that $d\underline{i}/d\sigma < 0$. Since $\lim_{\sigma\to 0}\underline{i}=1-m$, this implies that $\underline{i} < 1-m$ for $\sigma > 0$. Together with Claims 1–3 this implies that, for β sufficiently large,

$$\frac{O^*(\sigma)}{\lim_{\rho\to 0}O^*(\rho)} = \frac{\int_{\underline{i}}^1 e^{x_s(i)}di}{me^{x_{\sigma=0}}} = \frac{\int_{\underline{i}}^1 e^{x_s(i)-x_s(1-m)}di}{me^{x_{\sigma=0}-x_s(1-m)}} > 1 \ .$$

Hence, for sufficiently heterogeneous contests, perfect meritocracy is suboptimal.

This completes the proof of Proposition 2.   ■

Marginal Benefit of Meritocracy   Proof of Remark 1:

$\Longleftarrow$: Trivial.

$\Longrightarrow$: Recall from Lemma 21 that

$$\eta_{C_X} = X\frac{C_{XX}(X,i)}{C_X(X,i)} = \frac{c_{xx}(x,i)}{c_x(x,i)} - 1 \ . \tag{35}$$

We first show that if $\frac{c_{xx}(x,i)}{c_x(x,i)}$ does not depend on i, then $c_x(x,i)$ must be multiplicatively separable in x and i. To this end, suppose that $u \equiv \frac{c_{xx}(x,i)}{c_x(x,i)}$ is a function of x alone. Integrating both sides yields

$$\int_{x_0}^x u(z)\,dz = \int_{x_0}^x \frac{c_{xx}(z,i)}{c_x(z,i)}dz = \ln\frac{c_x(x,i)}{l(i)} \ , \tag{36}$$

where $x_0 < x$ is some constant and $l(i)$ is some function that does not depend on x. Solving (36) for c_x, we find

$$c_x(x, i) = l(i) \exp\left[\int_{x_0}^x u(z)\, dz\right] ,$$

which is separable in x and i. Hence, we can write

$$c_x(x, i) = c_x(x) l(i) , \tag{37}$$

where $c_x(x) \equiv c_x(x, i)/l(i)$ does not depend on on i.

Integrating (37) over x we find that $c(x, i)$ is of the form

$$
\begin{aligned}
c(x, i) &= \int_{x_0}^x c_x(z, i)\, dz + d(i) = l(i) \int_{x_0}^x c_x(z)\, dz + d(i) \\
&= c(x) l(i) + d(i) ,
\end{aligned}
$$

where $d(i)$ is some function that does not depend on x and $c(x) \equiv \int_{x_0}^x c_x(z)\, dz$.

Notice that for cost functions $c(x, i)$ of this form, equation (35) simplifies to

$$c_{xx}(x) = \left(\eta_{C_X} + 1\right) c_x(x) .$$

Since η_{C_X} is constant in x by assumption, while $c(-\infty) = c'(-\infty) = 0$ (Lemma 1), solving this differential equation yields

$$c(x) = k e^{\alpha x} ,$$

where $k > 0$ and $\alpha = \eta_{C_X} + 1$. In turn, this implies that $C(X, i) = kX^\alpha \cdot l(i) + d(i)$.

Finally, the claimed properties of $l(\cdot)$ and $d(i) = 0$ follow from our assumptions on $C(\cdot, \cdot)$.   ∎

The remainder of this section contains a proof of Proposition 6.

Recall the additional notation introduced before Lemma 29. For future reference, we calculate the following semi-elasticities:

Lemma 30

$$\frac{dx_s(\theta_\sigma^*, i)}{d\sigma/\sigma} = \underbrace{-\frac{\frac{f'}{f}\frac{\theta^* - x_s(i)}{\sigma} + 1}{-soc}}_{\frac{\partial x_s(i)}{\partial \sigma/\sigma}} + \underbrace{\frac{\frac{1}{\sigma}\frac{f'}{f}}{-soc}}_{\frac{\partial x_s(i)}{\partial \theta^*}} \underbrace{\left(B - 1_{\{\underline{i}\}}\frac{\sigma \bar{F}_{\underline{i}} g(\underline{i})}{\alpha D}\frac{d\underline{i}}{d\sigma/\sigma}\right)}_{\frac{d\theta^*}{d\sigma/\sigma}} ,$$

$$\frac{d\underline{i}}{d\sigma/\sigma} = -\frac{E}{\frac{-l'(\underline{i})}{l(\underline{i})}D + \sigma \bar{F}_{\underline{i}} g(\underline{i})} < 0 ,$$

where: 1) $B \equiv \frac{1}{D}\int_{\max\{0, \underline{i}\}}^1 \frac{\theta^* - x_s(i)}{-soc} fg\, di - \frac{1}{\alpha}$; *2)* $D \equiv \int_{\max\{0, \underline{i}\}}^1 \frac{1}{-soc} fg\, di > 0$; *and 3)*
$E \equiv \int_{\max\{0, \underline{i}\}}^1 \frac{\alpha[x_s(i) - x_s(\underline{i})] + 1}{-soc} fg\, di > 0.$

Proof. The calculations of $\mathrm{d}x_s\left(\theta_\sigma^*, i\right)/\mathrm{d}\sigma$ and $\mathrm{d}\underline{i}/\mathrm{d}\sigma$ are essentially the same as in the proof of Lemma 29 above. The only difference is that we now include the indicator function $1_{\{\underline{i}\}}$ to condition on whether $\underline{i}$ exists. $\blacksquare$

Recall from eqn. (13) that

$$\tilde{A}_{dir} \equiv 1_{\{\underline{i}\}}\frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma}\frac{e^{x_s(i)}}{O^*}g\left(i\right) , \tag{38}$$

while from eqn. (16)

$$\tilde{A}_{ind} \equiv 1_{\{\underline{i}\}}\frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma}\int_{\underline{i}}^1 \frac{e^{x_s(i)}}{O^*}\frac{\partial x_s\left(i\right)}{\partial\theta^*}g\left(i\right)\mathrm{d}i$$

$$= 1_{\{\underline{i}\}}\frac{\sigma\bar{F}_{\underline{i}}g\left(\underline{i}\right)}{\alpha D}\frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma}\int_{\underline{i}}^1 \frac{e^{x_s(i)}}{O^*}\frac{\frac{1}{\sigma}f'/f}{-soc}g\left(i\right)\mathrm{d}i . \tag{39}$$

Lemma 31 *As the distribution of abilities becomes degenerate, the sum of the direct and indirect attrition effects converges to the homogenous attrition effect. Formally,*

$$\tilde{A}_{dir} \xrightarrow{i\overset{\mathrm{Pr}}{\to}i_0} \begin{cases} \frac{\alpha}{-soc_{i_0}} > 0 & \textit{if } \underline{i} \xrightarrow{i\overset{\mathrm{Pr}}{\to}i_0} i_0 \\ 0 & \textit{otherwise} \end{cases} ,$$

$$\tilde{A}_{ind} \xrightarrow{i\overset{\mathrm{Pr}}{\to}i_0} \begin{cases} \frac{1}{\alpha}\left(1 - \frac{\alpha}{-soc_{i_0}}\right) & \textit{if } \underline{i} \xrightarrow{i\overset{\mathrm{Pr}}{\to}i_0} i_0 \\ 0 & \textit{otherwise} \end{cases} .$$

Proof. If $\underline{i}$ does not exist when $i\xrightarrow{\mathrm{Pr}}i_0$, then $1_{\{\underline{i}\}} = 0$ and, hence, $\tilde{A}_{dir} = \tilde{A}_{ind} = 0$.

If $\underline{i}$ does exist along the sequence, then $\underline{i}$ must converge to some $\underline{i}' \in I$. To see this notice that, otherwise, there would exist (at least) two convergence points for boundary ability $\underline{i}$, which is inconsistent with $c_i\left(x, i\right) < 0$ for all i.

Because $\underline{i} \to \underline{i}' > i_0$ when $i\xrightarrow{\mathrm{Pr}}i_0$ is inconsistent with equilibrium, it remains to consider the two cases: $\underline{i} \to \underline{i}' < i_0$ and $\underline{i} \to i_0$. In both cases

$$g\left(\underline{i}\right)\frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} = g\left(\underline{i}\right)\frac{\int_{\max\{0,\underline{i}\}}^1 \frac{\alpha[x_s(i)-x_s(\underline{i})]+1}{-soc}fg\mathrm{d}i}{\frac{-l'(\underline{i})}{l(\underline{i})}\int_{\max\{0,\underline{i}\}}^1 \frac{1}{-soc}fg\mathrm{d}i + \sigma\bar{F}_{\underline{i}}g\left(\underline{i}\right)} . \tag{40}$$

Case 1 ($\underline{i} \to \underline{i}' < i_0$): If $\underline{i} \to \underline{i}' < i_0$ when $i\xrightarrow{\mathrm{Pr}}i_0$, then $\lim_{i\overset{\mathrm{Pr}}{\to}i_0}g\left(\underline{i}\right) = 0$. Therefore, from (40),

$$g\left(\underline{i}\right)\frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \xrightarrow{i\overset{\mathrm{Pr}}{\to}i_0} \left[\lim_{i\overset{\mathrm{Pr}}{\to}i_0}g\left(\underline{i}\right)\right]\frac{\frac{\alpha[x_s(i_0)-x_s(\underline{i}')]+1}{-soc_{i_0}}f_{i_0}}{\frac{-l'(\underline{i}')}{l(\underline{i}')}\frac{1}{-soc_{i_0}}f_{i_0} + \sigma\bar{F}_{\underline{i}'}\left\{\lim_{i\overset{\mathrm{Pr}}{\to}i_0}\left[g\left(\underline{i}\right)\right]\right\}} = 0 .$$

It then follows from the definition of $\tilde{A}_{dir}$ that

$$\tilde{A}_{dir} \xrightarrow{i \overset{\text{Pr}}{\to} i_0} \frac{e^{x_s(\underline{i}')}}{O^*\mathbf{1}_{\{i=i_0\}}} \left[\lim_{i \overset{\text{Pr}}{\to} i_0} g\left(\underline{i}'\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \right] = 0 \ ,$$

and, similarly,

$$\tilde{A}_{ind} = \mathbf{1}_{\{\underline{i}\}} \frac{\bar{F}_{\underline{i}} \int_{\underline{i}}^1 \frac{e^{x_s(i)}}{O^*} \frac{f'/f}{-soc} g \mathrm{d}i}{\alpha \int_{\max\{0,\underline{i}\}}^1 \frac{1}{-soc} fg \mathrm{d}i} g\left(\underline{i}\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \xrightarrow{i \overset{\text{Pr}}{\to} i_0} \frac{\bar{F}_{\underline{i}} \frac{e^{x_s(i_0)}}{O^*\mathbf{1}_{\{i=i_0\}}} \frac{f'_{i_0}/f_{i_0}}{-soc_{i_0}}}{\frac{\alpha}{-soc_{i_0}} f_{i_0}} \left[\lim_{i \overset{\text{Pr}}{\to} i_0} g\left(\underline{i}'\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \right] = 0 \ .$$

Case 2 $(\underline{i} \to i_0)$: If $\underline{i} \to i_0$ when $i \xrightarrow{\text{Pr}} i_0$, then $g\left(\underline{i}\right) \xrightarrow{i \overset{\text{Pr}}{\to} i_0} \infty$. Therefore, from (40)

$$g\left(\underline{i}\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \xrightarrow{i \overset{\text{Pr}}{\to} i_0} \frac{\int_{\max\{0,\underline{i}\}}^1 \frac{\alpha[x_s(i)-x_s(\underline{i})]+1}{-soc} fg \mathrm{d}i}{\frac{1}{g(\underline{i})} \frac{-l'(\underline{i})}{l(\underline{i})} \int_{\max\{0,\underline{i}\}}^1 \frac{1}{-soc} fg \mathrm{d}i + \sigma \bar{F}_{\underline{i}}} = \frac{\frac{1}{-soc} f_{i_0}}{\sigma \bar{F}_{\underline{i}}} \ .$$

Using the FOC and the zero-profit condition at $\underline{i} = i_0$, we find that

$$\frac{\frac{v}{\sigma} f_{i_0}}{v \bar{F}_{i_0}} = \frac{c_x\left(i_0\right)}{c\left(i_0\right)} = \alpha \ .$$

Hence, $f_{i_0}/\bar{F}_{i_0} = \alpha\sigma$, such that

$$\tilde{A}_{dir} \xrightarrow{i \overset{\text{Pr}}{\to} i_0} \frac{e^{x_s(i_0)}}{O^*\mathbf{1}_{\{i=i_0\}}} \left[\lim_{i \overset{\text{Pr}}{\to} i_0} g\left(\underline{i}'\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \right] = \frac{\alpha}{-soc_{i_0}} \ ,$$

and

$$\tilde{A}_{ind} \xrightarrow{i \overset{\text{Pr}}{\to} i_0} \frac{\bar{F}_{i_0} \frac{e^{x_s(i_0)}}{O^*\mathbf{1}_{\{i=i_0\}}} \frac{f'_{i_0}/f_{i_0}}{-soc_{i_0}}}{\frac{\alpha}{-soc_{i_0}} f_{i_0}} \left[\lim_{i \overset{\text{Pr}}{\to} i_0} g\left(\underline{i}'\right) \frac{-\mathrm{d}\underline{i}}{\mathrm{d}\sigma/\sigma} \right]$$

$$= \frac{1}{\alpha} \frac{\bar{F}_{i_0}}{f_{i_0}} \frac{f'_{i_0}}{f_{i_0}} \frac{1}{\sigma - soc_{i_0}} \frac{f_{i_0}}{\bar{F}_{i_0}} = \frac{1}{\alpha\sigma} \frac{f'_{i_0}/f_{i_0}}{- soc_{i_0}} e^{x_s(i_0)}$$

$$= \frac{1}{\alpha} \left[\frac{\frac{1}{\sigma} f'_{i_0}/f_{i_0}}{-\left(\frac{1}{\sigma} f'_{i_0}/f_{i_0} + \alpha\right)} - 1 + 1 \right] = \frac{1}{\alpha} \left(1 - \frac{\alpha}{-soc_{i_0}} \right) > 0 \ .$$

∎

Next,

Lemma 32 *(i) If $\underline{i}$ exists, then the aggregate attrition effect is positive: $\tilde{A} > 0$.*

(ii) Furthermore, $\tilde{A} \xrightarrow{i \overset{\text{Pr}}{\to} i_0} A\mathbf{1}_{\{i=i_0\}}$.

Proof.

Part (i): Observe from (38) and (39) that

$$\tilde{A} = \tilde{A}_{dir} + \tilde{A}_{ind} = 1_{\{\underline{i}\}} \frac{e^{x_s(\underline{i})}}{O^*} \left[1 + \frac{\bar{F}_{\underline{i}} \int_{\underline{i}}^1 \frac{e^{x_s(i) - x_s(\underline{i})} f'/f}{-soc} g di}{\alpha \int_{\underline{i}}^1 \frac{1}{-soc} f g di} \right] g(\underline{i}) \frac{-d\underline{i}}{d\sigma/\sigma} .$$

Hence, provided $\underline{i}$ exists, $\tilde{A} > 0$ iff the expression in brackets is strictly positive. In turn, this is equivalent to

$$\int_{\underline{i}}^1 \frac{\frac{e^{x(i) - x(\underline{i})} f'/f}{f/\bar{F}_{\underline{i}}} + \alpha}{-soc} f g di > 0 .$$

Next, notice that

$$\int_{\underline{i}}^1 \frac{\frac{e^{x(i) - x(\underline{i})} f'/f}{f/\bar{F}_{\underline{i}}} + \alpha}{-soc} f g di > \int_{\underline{i}}^1 \frac{\frac{f'/f}{f/\bar{F}_{\underline{i}}} + \alpha}{-soc} f g di = \int_{\underline{i}}^1 \frac{\frac{f_{\underline{i}}}{f} \frac{f'/f}{f_{\underline{i}}/\bar{F}_{\underline{i}}} + \alpha}{-soc} f g di ,$$

where for the inequality we have used that $e^{x(i) - x(\underline{i})} > 1$ for all $i > \underline{i}$.

Use the FOC and the zero-profit condition at $\underline{i}$ to find that $\frac{\frac{v}{\sigma} f\left(\frac{\theta - x(\underline{i})}{\sigma}\right)}{v \bar{F}_{\underline{i}}} = \frac{c_x(x(\underline{i}), \underline{i})}{c(x(\underline{i}), \underline{i})} = \alpha$. Hence, $f_{\underline{i}}/\bar{F}_{\underline{i}} = \alpha \sigma$. To prove that $\tilde{A} > 0$, it therefore suffices to show that

$$\frac{f_{\underline{i}}}{f} \frac{f'/f}{\alpha \sigma} + \alpha = \frac{\frac{f_{\underline{i}}}{f} f'/f + \alpha^2 \sigma}{\alpha \sigma} > 0 .$$

Clearly, the inequality can only fail if $f'/f < 0$. In that case, $x(i) < \theta$. Because $\underline{i} \leq i$, $x(\underline{i}) \leq x(i) < \theta$. In turn, this implies that $0 < \frac{f_{\underline{i}}}{f} < 1$ and, therefore,

$$\frac{f_{\underline{i}}}{f} f'/f + \alpha^2 \sigma > f'/f + \alpha^2 \sigma \geq f'/f + \alpha \sigma > 0 ,$$

where second inequality uses the fact that $\alpha \geq 1$, while the third inequality follows from the SOC.

We may therefore conclude that $\tilde{A} > 0$.

Part (ii): It remains to prove that $\tilde{A} \xrightarrow{i \overset{\text{Pr.}}{\to} i_0} A \mathbf{1}_{\{\underline{i} = i_0\}}$. For all cases other than $\underline{i} \xrightarrow{i \overset{\text{Pr.}}{\to} i_0} i_0$, this is trivially true, because $\tilde{A} \xrightarrow{i \overset{\text{Pr.}}{\to} i_0} 0 = A \mathbf{1}_{\{\underline{i} = i_0\}}$. If $\underline{i} \xrightarrow{i \overset{\text{Pr.}}{\to} i_0} i_0$, then using Lemma 31,

$$\begin{aligned} \lim_{i \overset{\text{Pr.}}{\to} i_0} \tilde{A} &= \lim_{i \overset{\text{Pr.}}{\to} i_0} \tilde{A}_{dir} + \tilde{A}_{ind} = \frac{\alpha}{-soc_{i_0}} + \frac{1}{\alpha}\left(1 - \frac{\alpha}{-soc_{i_0}}\right) \\ &= \frac{\alpha - 1}{-soc_{i_0}} + \frac{1}{\alpha} = \frac{\eta_C - 1}{-soc_{i_0}} + \frac{1}{\eta_{C_X} + 1} . \end{aligned}$$

where we have used $c'(x)/c(x) = \eta_C = \alpha$ and $\eta_{C_X} = \alpha - 1$.

Finally, when $\underline{i} \xrightarrow{\text{Pr}} i_0$, the attrition effect in the corresponding homogeneous contest is

$$A_{\mathbf{1}_{\{i=i_0\}}} \equiv \frac{\mathrm{d}O^*_{\mathbf{1}_{\{i=i_0\}}}/O^*_{\mathbf{1}_{\{i=i_0\}}}}{\mathrm{d}\sigma/\sigma} - K = \frac{\eta_C - 1}{-soc} + \frac{1}{\eta_{C'} + 1} \ .$$

Hence, also in this case, $\tilde{A} \xrightarrow{i \xrightarrow{\text{Pr}} i_0} A_{\mathbf{1}_{\{i=i_0\}}}$. $\blacksquare$

Lemma 33 *The heterogeneity effect is strictly positive. In particular,*

$$\tilde{H} = \int_{\max\{0,\underline{i}\}}^{1} x_s(i) \frac{\left[e^{x_s(i)}/O^*\right]\frac{1}{\sigma}f'/f - \psi f}{-soc}g\,\mathrm{d}i > 0 \ ,$$

where $\psi = \frac{1}{D}\int_{\max\{0,\underline{i}\}}^{1}\frac{e^{x_s(i)}}{O^*}\frac{\frac{1}{\sigma}f'/f}{-soc}g\,\mathrm{d}i$. Furthermore, $\tilde{H} \xrightarrow{i \xrightarrow{\text{Pr}} i_0} 0$.

Proof. Recall that the heterogeneity effect $\tilde{H}$ is defined as the elasticity of output to noise minus the attrition ($\tilde{A}_{dir} + \tilde{A}_{ind}$) and competition effects ($\tilde{K}$):

$$\tilde{H} \equiv \frac{\mathrm{d}O^*/O^*}{\mathrm{d}\sigma/\sigma} - \tilde{A}_{dir} - \tilde{A}_{ind} - \tilde{K} \ .$$

Using Lemma 30 and (39) this can be expanded and simplified as follows.

$$
\begin{aligned}
\tilde{H} &= \int_{\max\{0,\underline{i}\}}^{1} \frac{e^{x_s(\theta^*_\sigma,i)}}{O^*}\frac{\mathrm{d}x_s(\theta^*_\sigma,i)}{\mathrm{d}\sigma/\sigma}g(i)\,\mathrm{d}i - \tilde{A}_{ind} + \frac{1}{\alpha} \\
&= \int_{\max\{0,\underline{i}\}}^{1} \frac{e^{x_s(i)}}{O^*}\left\{-\frac{\frac{f'}{f}\frac{\theta^*-x_s(i)}{\sigma}+1}{-soc} + \frac{\frac{1}{\sigma}\frac{f'}{f}}{-soc}\left[\frac{\int_{\max\{0,\underline{i}\}}^{1}\frac{\theta^*-x_s(i)}{-soc}fg\,\mathrm{d}i}{\int_{\max\{0,\underline{i}\}}^{1}\frac{1}{-soc}fg\,\mathrm{d}i}-\frac{1}{\alpha}\right]\right\}g(i)\,\mathrm{d}i + \frac{1}{\alpha} \\
&= \int_{\max\{0,\underline{i}\}}^{1} \frac{e^{x_s(i)}}{O^*}\frac{\frac{f'}{f}\frac{x_s(i)}{\sigma}}{-soc}g(i)\,\mathrm{d}i + \int_{\max\{0,\underline{i}\}}^{1} \frac{e^{x_s(i)}}{O^*}\frac{\frac{1}{\sigma}\frac{f'}{f}}{-soc}\frac{\int_{\max\{0,\underline{i}\}}^{1}\frac{-x_s(i)}{-soc}fg\,\mathrm{d}i}{\int_{\max\{0,\underline{i}\}}^{1}\frac{1}{-soc}fg\,\mathrm{d}i}g(i)\,\mathrm{d}i \\
&\quad + \int_{\underline{i}}^{1} \frac{e^{x_s(i)}}{O^*}\left[-\frac{1}{-soc}\left(\frac{\frac{1}{\sigma}f'/f+\alpha}{\alpha}\right)\right]g\,\mathrm{d}i + \frac{1}{\alpha} \ .
\end{aligned}
$$

Since $\frac{1}{\sigma}f'/f + \alpha = -soc$ and $\int_{\underline{i}}^{1}e^{x_s(i)}g\,\mathrm{d}i = O^*$, the last two terms cancel. Hence,

$$
\begin{aligned}
\tilde{H} &= \int_{\max\{0,\underline{i}\}}^{1} \frac{e^{x_s(i)}}{O^*}\frac{\frac{f'}{f}\frac{x_s(i)}{\sigma}}{-soc}g\,\mathrm{d}i + \int_{\max\{0,\underline{i}\}}^{1} \frac{-x_s(i)}{-soc}fg\,\mathrm{d}i \cdot \psi \\
&= \int_{\max\{0,\underline{i}\}}^{1} x_s(i)\left[\frac{\frac{e^{x_s(i)}}{O^*}\frac{1}{\sigma}\frac{f'}{f}-\psi f}{-soc}\right]g\,\mathrm{d}i \ ,
\end{aligned}
$$

where

$$\psi = \frac{\int_{\max\{0,\underline{i}\}}^1 \frac{e^{x_s(i)}}{O^*} \frac{\frac{1}{\sigma} f'/f}{-soc} g \mathrm{d}i}{\int_{\max\{0,\underline{i}\}}^1 \frac{1}{-soc} f g \mathrm{d}i} \ .$$

Now notice that

$$\tilde{H} \xrightarrow{i \xrightarrow{\text{Pr.}} i_0} x_s(i_0) \frac{\left[e^{x_s(i)}/O^*\right] \frac{1}{\sigma} f'_{i_0}/f_{i_0} - \left[e^{x_s(i)}/O^*\right] \frac{1}{\sigma} f'_{i_0}/f_{i_0}}{-soc_{i_0}} = 0 \ .$$

Finally, we show that $\tilde{H} > 0$. Integrating by parts

$$\tilde{H} = [x_s(i) q(i)]_{\max\{0,\underline{i}\}}^1 - \int_{\max\{0,\underline{i}\}}^1 x'_s(i) q(i) \mathrm{d}i \ , \tag{41}$$

where

$$q(i) = \int_{\max\{0,\underline{i}\}}^i \frac{\left[e^{x_s(\iota)}/O^*\right] \frac{1}{\sigma} f'/f - \psi f}{-soc_\iota} g \mathrm{d}\iota \ .$$

Observe that $q(\max\{0,\underline{i}\}) = 0$ and $q(1) = 0$. Hence, the first term of (41) equals zero, and

$$\tilde{H} = -\int_{\max\{0,\underline{i}\}}^1 x'_s(i) q(i) \mathrm{d}i \ .$$

Clearly, $x'_s(i) > 0$. It remains to show that $q(i) < 0$ for $i \in (\max\{0,\underline{i}\}, 1)$. Since, $q(\max\{0,\underline{i}\}) = 0$ and $q(1) = 0$, it remains to observe that $\left[e^{x_s(i)}/O^*\right] \frac{1}{\sigma} f'/f$ crosses $-\psi f$ in i just once from below. To see this, notice that: 1) $e^{x_s(i)} f'/f$ is strictly increasing in i and crosses the horizontal axis where the argument of f and f'—i.e., $[\theta - x_s(i)]/\sigma$—is zero; and 2) f is single-peaked around zero, and hence, regardless of whether ψ is positive or negative, $-\psi f$ is downward sloping where $\left[e^{x_s(i)}/O^*\right] \frac{1}{\sigma} f'/f = -\psi f$. $\blacksquare$

This completes the proof of Proposition 6.

Discrete Number of Contestants

2-Player LR Contests

Lemma 34 *If $\rho < \underline{\rho}$ then, in all symmetric MSEs, the supports of players' equilibrium strategies are unbounded from below.*

Proof. Let $\underline{s}$ denote the infimum of a symmetric MSE's support. Suppose, contrary to the lemma, that for some $\rho < \underline{\rho}$ there exists a MSE with $\underline{s} > -\infty$. Initially, we

consider strategies on a grid, such that the support can be written as $\{\underline{s} + \Delta_k\}_{k \in \mathbb{N}}$, $\Delta_k > 0$. Eventually, we let the step-size $\Delta_{k+1} - \Delta_k \to 0$.

Define $\Delta_0 \equiv 0$ and let cumulative step-size Δ_k be variable in $k \in \mathbb{N}$, such that the cost of raising log-output x from $\underline{s} + \Delta_k$ to $\underline{s} + \Delta_{k+1}$ is constant and equal to $\delta > 0$ for all k. On such an "iso-marginal-cost grid,"

$$c(\underline{s} + \Delta_k) = c(\underline{s}) + k\delta, \ k \in \mathbb{N} .$$

(Since $c(\cdot)$ is strictly convex, $\Delta_{k+1} - \Delta_k$ is strictly decreasing in k.)

Denote by p_k^i, $i \in \{1, 2\}$, the equilibrium probability that player i produces $\underline{s} + \Delta_k$. Since we restrict attention to symmetric equilibria, $p_k^1 = p_k^2$ for all k. However, for purposes of exposition, in this case we do not drop the player-superscripts.

Let $\underline{k}$ index of the smallest point on the grid for which $p_k^i > 0$. Notice that $\underline{k}$ and p_k^i, as well as $\Delta_{\underline{k}}$ and $\Delta_{\underline{k}+l}$, $l \in \mathbb{N}$, all depend on δ. Since $\underline{s} > -\infty$ is the infimum of the MSE, $\underline{s} + \Delta_{\underline{k}} \to \underline{s}$ as $\delta \to 0$, i.e., $\Delta_{\underline{k}} \xrightarrow{\delta \to 0} 0$. Also, $\Delta_{\underline{k}+l} \xrightarrow{\delta \to 0} 0$, for all l.

For future reference, we show:

Claim 1: $\lim_{\delta \to 0} \dfrac{\mathrm{d}\Delta_{\underline{k}+l}}{\mathrm{d}\delta} = \dfrac{\underline{k}+l}{c'(\underline{s})}$, $l \in N$.

Proof of Claim 1: From the definition of δ it follows that

$$(\underline{k} + l)\,\delta = c(\underline{s} + \Delta_{\underline{k}+l}) - c(\underline{s}) .$$

Differentiating both sides with respect to δ, we find

$$\underline{k} + l = c'(\underline{s} + \Delta_{\underline{k}+l}) \frac{\mathrm{d}\Delta_{\underline{k}+l}}{\mathrm{d}\delta} .$$

Since $\Delta_{\underline{k}+l} \xrightarrow{\delta \to 0} 0$, this implies

$$\lim_{\delta \to 0} \frac{\mathrm{d}\Delta_{\underline{k}+l}}{\mathrm{d}\delta} = \lim_{\delta \to 0} \frac{\underline{k} + l}{c'(\underline{s} + \Delta_{\underline{k}+l})} = \frac{\underline{k} + l}{c'(\underline{s})} ,$$

which completes the proof of Claim 1.

All actions $\underline{s} + \Delta_k$ that are played with strictly positive probability must yield the same expected payoff $q \geq 0$. Hence, we may write the following complementary slackness condition:

$$\forall l \in \mathbb{N} : p_l^i \left(\sum_{k=\underline{k}}^{\infty} \left\{ \bar{\bar{\Phi}} \left[\frac{\underline{s} + \Delta_k - (\underline{s} + \Delta_l)}{\rho} \right] v - c(\underline{s} + \Delta_l) \right\} p_k^j - q \right) = 0 , \ i \neq j .$$

For future reference, also note that the bracketed factor is ≤ 0 since, otherwise, the payoff from playing $\underline{s} + \Delta_l$ would be strictly greater than q.

Because $p_{\underline{k}}^i > 0$,

$$\sum_{k=\underline{k}}^{\infty} \bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right) p_k^j = \frac{q + c\left(\underline{s} + \Delta_{\underline{k}}\right)}{v} \ .$$

Isolating $p_{\underline{k}}^j$,

$$p_{\underline{k}}^j = \frac{1}{\bar{\Phi}\left(0\right)}\left[\frac{q + c\left(\underline{s} + \Delta_{\underline{k}}\right)}{v} - \sum_{k=\underline{k}+1}^{\infty} \bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right) p_k^j\right] \ . \tag{42}$$

Since the bracketed factor in the complementary slackness condition is ≤ 0, the following inequality holds for $l = \underline{k} + 1$.

$$\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right) p_{\underline{k}}^j + \sum_{k=\underline{k}+1}^{\infty} \bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right) p_k^j \leq \frac{q + c\left(\underline{s} + \Delta_{\underline{k}+1}\right)}{v} \ . \tag{43}$$

Substituting (42) into (43) yields

$$\left\{\begin{array}{c} \frac{\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right)}{\bar{\Phi}(0)}\left[\frac{q + c\left(\underline{s} + \Delta_{\underline{k}}\right)}{v} - \sum_{k=\underline{k}+1}^{\infty} \bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right) p_k^j\right] \\ + \sum_{k=\underline{k}+1}^{\infty} \bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right) p_k^j \end{array}\right\} \leq \frac{q + c\left(\underline{s} + \Delta_{\underline{k}+1}\right)}{v} \ .$$

Using that $c\left(\underline{s} + \Delta_{\underline{k}+1}\right) = c\left(\underline{s} + \Delta_{\underline{k}}\right) + \delta$ and simplifying, we find that

$$\sum_{k=\underline{k}+1}^{\infty} \left[\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right) - 2\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right)\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)\right] p_k^j \tag{44}$$

$$\leq \frac{q + c\left(\underline{s} + \Delta_{\underline{k}}\right) + \delta}{v} - 2\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right)\frac{q + c\left(\underline{s} + \Delta_{\underline{k}}\right)}{v} \ .$$

Claim 2: The LHS of (44) is strictly positive.

Proof of Claim 2: We prove the claim by showing that, for every k, the term in square brackets is strictly positive. This requires that

$$\frac{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right)}{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)} > 2\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right) \ .$$

For this inequality to hold, it suffices to show that the derivative of the LHS with respect to Δ_k is strictly positive, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}\Delta_k}\left[\frac{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right)}{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)}\right] = \frac{-\frac{1}{\rho}\varphi\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right)\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right) + \frac{1}{\rho}\varphi\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right)}{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)^2} > 0 .$$

This is equivalent to

$$\frac{\varphi\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)}{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}}}{\rho}\right)} > \frac{\varphi\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right)}{\bar{\Phi}\left(\frac{\Delta_k - \Delta_{\underline{k}+1}}{\rho}\right)} .$$

Since log-concavity of the density implies that the hazard rate is strictly increasing, and $\Delta_k - \Delta_{\underline{k}} > \Delta_k - \Delta_{\underline{k}+1}$ for all $k \geq \underline{k}$, the inequality is indeed satisfied.

This proves Claim 2.

Claim 3: For δ sufficiently small, the RHS of (44) is negative.

Proof of Claim 3: Notice that $q \geq \delta$, i.e., the expected payoff must weakly exceed the cost of an output increment. Otherwise, it would be profitable to move down one step and save δ in costs, even if losing were certain.

Also notice that the RHS of (44) is strictly decreasing in the expected equilibrium payoff q, since $\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right) > 1/2$.

Together, these observations imply that, to prove Claim 3, it suffices to show that the RHS of (44) is negative for small δ when $q = \delta$. In turn, this entails showing that, for small δ,

$$\frac{2\delta + c\left(\underline{s} + \Delta_{\underline{k}}\right)}{\delta + c\left(\underline{s} + \Delta_{\underline{k}}\right)} - 2\bar{\Phi}\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right) \leq 0 . \tag{45}$$

Since $\Delta_{\underline{k}} - \Delta_{\underline{k}+1} \xrightarrow{\delta \to 0} 0$, (45) holds with equality in the limit.

Differentiating the LHS of (45) with respect to δ yields

$$\frac{c\left(\underline{s} + \Delta_{\underline{k}}\right) - \delta c'\left(\underline{s} + \Delta_{\underline{k}}\right)\frac{\mathrm{d}\Delta_{\underline{k}}}{\mathrm{d}\delta}}{\left[\delta + c\left(\underline{s} + \Delta_{\underline{k}}\right)\right]^2} + 2\varphi\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right)\frac{1}{\rho}\frac{\mathrm{d}\left(\Delta_{\underline{k}} - \Delta_{\underline{k}+1}\right)}{\mathrm{d}\delta} .$$

Write $\rho = \alpha\underline{\rho} = \alpha\frac{\varphi(0)v}{c'(\bar{x})}$, where $0 < \alpha < 1$, and substitute $\frac{v}{2} = c\left(\bar{x}\right)$. The derivative of the LHS of (45) then becomes

$$\frac{c\left(\underline{s} + \Delta_{\underline{k}}\right) - \delta c'\left(\underline{s} + \Delta_{\underline{k}}\right)\frac{\mathrm{d}\Delta_{\underline{k}}}{\mathrm{d}\delta}}{\left[\delta + c\left(\underline{s} + \Delta_{\underline{k}}\right)\right]^2} + \frac{\varphi\left(\frac{\Delta_{\underline{k}} - \Delta_{\underline{k}+1}}{\rho}\right)}{\varphi\left(0\right)}\frac{1}{\alpha}\frac{c'\left(\bar{x}\right)}{c\left(\bar{x}\right)}\frac{\mathrm{d}\left(\Delta_{\underline{k}} - \Delta_{\underline{k}+1}\right)}{\mathrm{d}\delta}$$

Letting $\delta \to 0$, this expression reduces to

$$\frac{1}{c\left(\underline{s}\right)} - \frac{1}{\alpha}\frac{c'\left(\bar{x}\right)}{c\left(\bar{x}\right)}\frac{1}{c'\left(\underline{s}\right)} ,\tag{46}$$

where we have used Claim 1 and the fact that $\Delta_{\underline{k}}, \Delta_{\underline{k}+1} \xrightarrow{\delta \to 0} 0$.

The expression in (46) is strictly negative (implying that it is also negative in a neighborhood of $\delta = 0$) iff

$$\alpha\frac{c'\left(\underline{s}\right)}{c\left(\underline{s}\right)} < \frac{c'\left(\bar{x}\right)}{c\left(\bar{x}\right)} .\tag{47}$$

This inequality in (47) holds, because: 1) $0 < \alpha < 1$; 2) $c'\left(\cdot\right)/c\left(\cdot\right)$ is weakly increasing in its argument due to weak log-convexity of $c\left(\cdot\right)$; and 3) $\underline{s} < \bar{x}$.

Thus we may conclude that the RHS of (44) is negative for small δ.

This proves Claim 3.

Together, Claims 2 and 3 contradict (44). Hence, on a sufficiently fine grid, there cannot exist a symmetric MSE with support bounded from below when $\rho < \underline{\rho}$. ∎

Lemma 35 *In any symmetric MSE with unbounded lower support, players' equilibrium payoffs are zero.*

Proof. When a symmetric MSE's lower support is unbounded, notice that there can be no mass point in players' strategy at $x = -\infty$. Otherwise, it would be a profitable deviation to exert a small amount of effort at infinitesimal cost and win with certainty whenever the opponent played $x_j = -\infty$.

Since there is no mass point at $-\infty$, it follows that player i's probability of winning goes to zero when $x_i \to -\infty$. Furthermore, $\lim_{x_i \to -\infty} c\left(x_i\right) = 0$. Thus, in equilibrium, player i's payoff goes to zero when $x_i \to -\infty$. Finally, since all actions in the support of an MSE must have the same expected payoff, this implies the result. ∎

Proof of Lemma 5:

Part (i): First we show that for all $\rho > 0$, there exists a (symmetric) equilibrium.

Let $H_i^*\left(X\right)$ denote the CDF of agent i's equilibrium strategy. Observe that $X_i > \bar{X}$, $i \in \{1, 2\}$, is never a best response to any effort of a rival, since it is strictly dominated by $X_i = 0$. Hence, a strategy profile (H_1^*, H_2^*) is an equilibrium of the contest with $X_i \in [0, \infty)$, iff it is an equilibrium of the contest with efforts restricted to the non-empty and compact set $[0, \bar{X}]$.

For the restricted contest—and, hence, also for the unrestricted one—the claim now follows from Theorem 6 of Dasgupta and Maskin (1986a, DM86). To see this, first notice that the contest is a symmetric game according to Definition 7 in DM86.

Second, the sum of the agents' payoff functions, $v - C(X_1) - C(X_2)$, is continuous, and hence upper semi-continuous.

Third, the set $A^{**}(i)$ of discontinuity points of the payoff function $\pi_i(X_i, X_j)$, where $i, j \in \{1, 2\}$ and $i \neq j$, consists of a single point, namely, $\{(0,0)\}$.

Fourth, the set $A^{**}(i) = \{(0,0)\}$ is a subset of $A^*(i)$, where $A^*(i)$ is defined by equation (2) on page 7 of DM86. To see this, notice that we can choose $D(1) = D(2) = 1$ and $f_{12}^1(0) = f_{21}^1(0) = 0$.

Fifth, the payoff function $\pi_i(X_i, X_j)$ is bounded and satisfies *Property (α)* of DM86. In particular, for $\lambda = 0$

$$\lambda \lim_{X_i \xrightarrow{-} 0} \inf \pi_i(X_i, 0) + (1 - \lambda) \lim_{X_i \xrightarrow{+} 0} \inf \pi_i(X_i, 0) = v > \frac{v}{2} = \pi_i(0, 0) \ .$$

Since we have verified that all predicates of Theorem 6 of DM86 are satisfied, the result follows.

Next we show that a PSE (x_1^*, x_2^*) exists iff $\rho \geq \underline{\rho}$, that it is unique, and that it is symmetric, with $x_1^* = x_2^* = x^* = c'^{-1}\left[\frac{v}{\rho}\varphi(0)\right]$.

To establish the form of contestants' best response curves, $\hat{x}_i(x_j)$, where $i, j \in \{1, 2\}$ and $i \neq j$, we follow essentially the same steps as in the proof of Proposition 1. However, here, x_j takes the place of θ, and ρ takes the places of σ.

Claim 1: All PSEs are interior.

Proof of Claim 1: First observe that $x_1 = x_2 = -\infty$ is not an equilibrium, since both players have an incentive to deviate: each player would be strictly better off producing an infinitesimally small amount and win with certainty.

Second, notice that $(x_1, x_2) = (b, -\infty)$, $b > -\infty$, is not an equilibrium either, since player 1 would be strictly better off deviating to some $b' \in (-\infty, b)$. Mutatis mutandis, the same is true for $(x_1, x_2) = (-\infty, b)$.

Claim 2: The symmetric pure strategy profile $\left(x_\rho^0, x_\rho^0\right)$, where $x_\rho^0 = c'^{-1}\left[\frac{v}{\rho}\varphi(0)\right]$, satisfies the FOCs and SOCs.

Proof of Claim 2: Trivial.

Claim 3: $\left(x_\rho^0, x_\rho^0\right)$ *is the only one solution to the system of FOCs and SOCs.*

Proof of Claim 3: Ignoring the participation constraint for now, denote by $\hat{x}_1\left(x_2\right)$ and $\hat{x}_2\left(x_1\right)$ the players *interior* best responses, i.e., all solutions to players' FOCs that satisfy the SOC for a local maximum. Since $\hat{x}_i\left(x_\rho^0\right) = x_\rho^0$, $i \in \{1, 2\}$, $\hat{x}_1\left(x_2\right)$ and $\hat{x}_2\left(x_1\right)$ intersect at $\left(x_\rho^0, x_\rho^0\right)$. We will show that this is in fact the only point where the interior best responses intersect.

Figure 6. The figure plots the interior best response curves for the two contestants. Lemma <ref>L: dx/dsig</ref> bounds the slope of the curves in each of the quadrants around the unique intersection point (x_ρ^0, x_ρ^0). This point constitutes a PSE iff the participation constraints of the agents are satisfied, which requires that $x_\rho^0 = \bar{x}$.
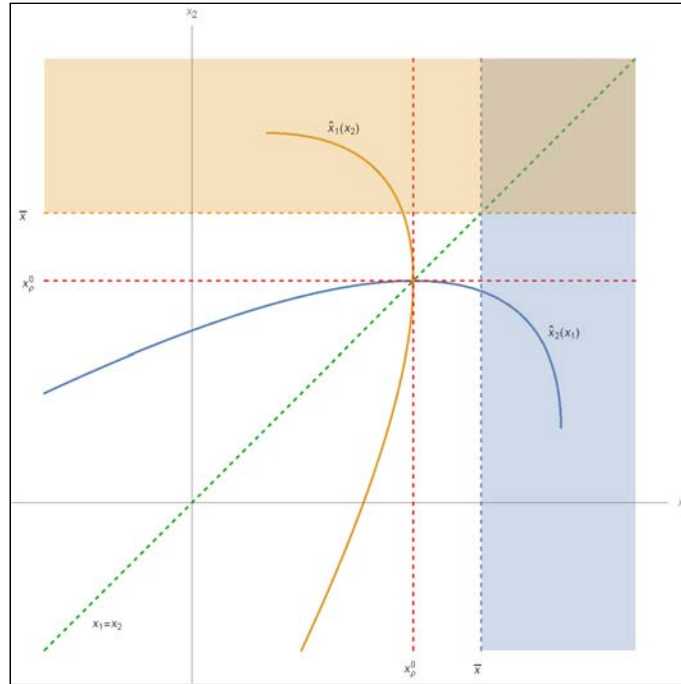
Figure 6 facilitates the understanding of the following analytical argument. From Lemma 18 it follows that, for all $x_1 < x_\rho^0$, $x_1 < \hat{x}_2\left(x_1\right) < x_\rho^0$, while for all $x_1 > x_\rho^0$, $\hat{x}_2\left(x_1\right) < x_\rho^0$. Similarly, for all $x_2 < x_\rho^0$, $x_2 < \hat{x}_1\left(x_2\right) < x_\rho^0$, while for all $x_2 > x_\rho^0$, $\hat{x}_1\left(x_2\right) < x_\rho^0$. Thus, the Cartesian plane can be partitioned into five regions about the point $\left(x_\rho^0, x_\rho^0\right)$, and one can compute whether the reaction functions $\hat{x}_1\left(x_2\right)$ and

$\hat{x}_2\left(x_1\right)$ pass through each of the regions:

Quadrant	Region Definition	$\hat{x}_1\left(x_2\right)$	$\hat{x}_2\left(x_1\right)$
I	$x_1 > x_\rho^0 \cap x_2 > x_\rho^0$	No	No
II	$x_1 < x_\rho^0 \cap x_2 > x_\rho^0$	Yes	No
III.a	$x_1 < x_\rho^0 \cap x_2 < x_\rho^0 \cap x_1 < x_2$	No	Yes
III.b	$x_1 < x_\rho^0 \cap x_2 < x_\rho^0 \cap x_1 > x_2$	Yes	No
IV	$x_1 > x_\rho^0 \cap x_2 < x_\rho^0$	No	Yes

Since the inequalities in Lemma 18 are strict, the reaction curves remain strictly in the interior of the regions, except at $\left(x_\rho^0, x_\rho^0\right)$.

Because $\hat{x}_1\left(x_2\right)$ and $\hat{x}_2\left(x_1\right)$ never occupy the same region except at $\left(x_\rho^0, x_\rho^0\right)$, they cannot intersect anywhere else. Thus, $\left(x_\rho^0, x_\rho^0\right)$ is the unique solution to the system of FOCs and SOCs.

Claim 4: $\left(x_\rho^0, x_\rho^0\right)$ *is a PSE if and only if* $\rho \geq \underline{\rho}$.

Proof of Claim 4: The unique solution $\left(x_\rho^0, x_\rho^0\right)$ to the system of FOCs and SOCs constitutes a PSE if and only if it satisfies the participation constraint, namely,

$$v\overline{\Phi}\left(0\right) \geq c\left(x_\rho^0\right) \ .$$

In Figure 6, this corresponds to the area outside the shaded regions.

Substituting in $x_\rho^0 = c'^{-1}\left[\frac{v}{\rho}\varphi\left(0\right)\right]$ and recognizing that $\overline{\Phi}\left(0\right) = 1/2$, this constraint becomes

$$v/2 \geq c\left(c'^{-1}\left[\frac{v}{\rho}\varphi\left(0\right)\right]\right) \ .$$

Finally, solving for ρ yields

$$\rho \geq \frac{v\varphi\left(0\right)}{c'\left[c^{-1}\left(v/2\right)\right]} = \underline{\rho} \ .$$

This completes the proof of the claim.

Part (ii): The proof that individual effort increases with meritocracy carries over directly from the interior case of the continuum model—in particular from Lemma 18.(i).

To see that $x_{\underline{\rho}}^* = \bar{x} = c^{-1}\left(v/2\right)$—and thus $O_{\underline{\rho}}^* = 2\bar{X}$—recall that $x_\rho^* = c'^{-1}\left[\frac{v}{\rho}\varphi\left(0\right)\right]$ and $\underline{\rho} = \frac{v\varphi(0)}{c'[\bar{x}]}$. Hence,

$$x_\rho^*\big|_{\rho=\underline{\rho}} = c'^{-1}\left[\frac{v}{\frac{v\varphi(0)}{c'[\bar{x}]}}\varphi\left(0\right)\right] = \bar{x} \ .$$

Finally, observe that at $(\bar{x}, \bar{x})$, the expected payoff is zero:

$$\pi\left(\bar{x}, \bar{x}\right) = \bar{\Phi}\left(0\right) v - c\left(\bar{x}\right) = \frac{v}{2} - c\left(c^{-1}\left(\frac{v}{2}\right)\right) = 0 \ .$$

Part (iii): From Part (i), if $\rho < \underline{\rho}$, then only MSEs exist. From Lemma 34 we know that, in any such *symmetric* MSE, the supports of the players' equilibrium strategies are unbounded from below. It then follows from Lemma 35 that, in these symmetric MSEs, players' equilibrium payoffs are zero.

Part (iv): Suppose $C''\left(\cdot\right) \overset{(=)}{>} 0$. Individual rationality requires that, in any equilibrium, the probability of winning times the prize value must weakly exceed the expected cost. That is, for $i = 1, 2$,

$$v \Pr\left\{i \text{ wins}\right\} \geq E_{H_i^*}\left[C\left(X\right)\right] \overset{(=)}{>} C\left[E_{H_i^*}\left(X\right)\right] \ ,$$

where the second inequality (equality) follows from the convexity (linearity) of C and Jensen's inequality.

Because C is strictly monotone, this implies

$$E_{H_i^*}\left(X\right) \overset{(\leq)}{<} C^{-1}\left(v \Pr\left\{i \text{ wins}\right\}\right) \ .$$

Summing the inequalities across both players yields

$$E_{H_1^*}\left(X\right) + E_{H_2^*}\left(X\right) \overset{(\leq)}{<} C^{-1}\left(v \Pr\left\{1 \text{ wins}\right\}\right) + C^{-1}\left(v \Pr\left\{2 \text{ wins}\right\}\right) \ .$$

By concavity (linearity) of $C^{-1}\left(\cdot\right)$,

$$\frac{C^{-1}\left(v \Pr\left\{1 \text{ wins}\right\}\right) + C^{-1}\left(v \Pr\left\{2 \text{ wins}\right\}\right)}{2} \overset{(\leq)}{<} C^{-1}\left(\frac{v \Pr\left\{1 \text{ wins}\right\} + v \Pr\left\{2 \text{ wins}\right\}}{2}\right) = C^{-1}\left(\frac{v}{2}\right) \ ,$$

and, therefore,

$$E_{H_1^*}\left(X\right) + E_{H_2^*}\left(X\right) \overset{(\leq)}{<} 2C^{-1}\left(\frac{v}{2}\right) = 2\bar{X} \ .$$

This completes the proof of the first claim in Part (iv).

Next, suppose that $C''\left(\cdot\right) = 0$ and $\rho < \underline{\rho}$. From Part (iii), agent payoffs are zero in a symmetric MSE. Thus,

$$v \Pr\left\{i \text{ wins}\right\} = E_{H_i^*}\left[C\left(X\right)\right] = C\left[E_{H_i^*}\left(X\right)\right] \ ,$$

where the second equality follows from the linearity of C. Tracing the argument for $C''(\cdot) \overset{(=)}{>} 0$ above, and replacing inequalities with equalities, yields the result.
■

Proof of Proposition 1.LR: From Lemma 5.(ii.a,b) we know that, in the unique PSE, agents' outputs, x^*, strictly exceed $-\infty$ and strictly increase in meritocracy. Hence, among PSEs, output is uniquely maximized at $\rho = \underline{\rho}$, where aggregate output is $O^* = 2\bar{X}$ (see Lemma 5.(ii.c)).

From Lemma 5.(iv.a) we know that, in all MSEs, aggregate output is no greater than $2\bar{X}$. Hence, the PSE with $\rho = \underline{\rho}$ maximizes total output.

Furthermore, if $C''(\cdot) > 0$, we know from Lemma 5.(iv.a) that in all MSEs aggregate output is strictly smaller than $2\bar{X}$. Hence, in this case, the PSE with $\rho = \underline{\rho}$ is the unique maximizer.

Finally, for $C''(\cdot) = 0$ and $\rho \in (0, \underline{\rho}]$, we know from Lemma 5.(iv.b) that, in symmetric MSEs, aggregate output equals $2\bar{X}$.   ■

n-Player Tullock Contests   Proof of Lemma 6: First, recall that EVTIM random variables are closed under maximization. This implies that, for $y \in \mathbb{R}$,

$$\Pr\left\{\max_j \{x_j + \varepsilon_j\} \leq y\right\} = \prod_{j=1}^n \Pr\{x_j + \varepsilon_j \leq y\} = \prod_{j=1}^n e^{-e^{-\frac{y-x_j}{\sigma}}} = e^{-e^{-\frac{y}{\sigma}}\sum_{j=1}^n e^{\frac{x_j}{\sigma}}} = e^{-e^{-\frac{y-\hat{x}}{\sigma}}},$$

which is the CDF of an EVTIM random variable with location $\hat{x} = \sigma \ln \sum_{j=1}^n e^{\frac{x_j}{\sigma}}$ and scale σ. Hence, the probability of beating all other agents with outputs x_j and noise $\varepsilon_j \sim EVTIM(0, \sigma)$, $j \neq i$, is the same as beating a single synthetic agent with output $\hat{x}$ and noise $\hat{\varepsilon} \sim EVTIM(0, \sigma)$.

Now recall that the difference of two EVTIM random variables with the same location parameter $\mu \in \mathbb{R}$ and the same scale parameter $\sigma > 0$ is distributed $Logistic(0, \sigma)$. Hence, the probability that player i beats all other players $j \neq i$, or equivalently, the single synthetic opponent, is

$$\Pr\left\{\max_{j \neq i} \{x_j + \varepsilon_j\} \leq x_i + \varepsilon_i\right\} = \Pr\{\hat{\varepsilon} - \varepsilon_i \leq x_i - \hat{x}_{-i}\},$$

where $\hat{x}_{-i} = \sigma \ln \sum_{j=1; j \neq i}^n e^{\frac{x_j}{\sigma}}$, and $\hat{\varepsilon}, \varepsilon_i$ are EVTIMs with location 0 and scale σ, whose difference is distributed $Logistic(0, \sigma)$.

The probability that contestant i is ranked first is then

$$\Pr\{\hat{\varepsilon} - \varepsilon_i \leq x_i - \hat{x}_{-i}\} = \frac{1}{1 + e^{-\frac{x_i - \hat{x}_{-i}}{\sigma}}} = \frac{e^{\frac{x_i}{\sigma}}}{e^{\frac{x_i}{\sigma}} + e^{\frac{\bar{x}_{-i}}{\sigma}}}$$

$$= \frac{e^{\frac{x_i}{\sigma}}}{\sum_{j=1}^{n} e^{\frac{x_j}{\sigma}}} = \frac{X_i^r}{\sum_{j=1}^{n} X_j^r} \ .$$

This completes the proof. ∎

Lemma 36 *The PSE of an n-player Tullock contest is unique.*

Proof. The payoff of an n-player Tullock contest is

$$v \frac{X_i^{\frac{1}{\sigma}}}{\sum_{j=1}^{n} X_j^{\frac{1}{\sigma}}} - C(X_i) \ .$$

The FOC is given by

$$v \frac{\frac{1}{\sigma} X_i^{\frac{1}{\sigma} - 1}}{Q} \left(1 - \frac{X_i^{\frac{1}{\sigma}}}{Q}\right) = C'(X_i) \ ,$$

where $Q = \sum_{j=1}^{n} X_j^{\frac{1}{\sigma}}$.

Let $S_i = X_i^{\frac{1}{\sigma}}/Q$ be a share function (à la Cornes and Hartley 2005, Jensen 2016), such that $\sum_{j=1}^{n} S_j = 1$. The FOC of player i can be written as

$$FOC_i \equiv \frac{v}{\sigma} \frac{S_i(1 - S_i)}{(QS_i)^{\sigma}} - C'[(QS_i)^{\sigma}] = 0 \ .$$

Implicitly differentiating the FOC yields $dS_i/dQ = -\frac{\partial FOC_i}{\partial Q} \Big/ \frac{\partial FOC_i}{\partial S_i}$, where

$$\frac{\partial FOC_i}{\partial Q} = -\frac{\partial [(QS_i)^{\sigma}]}{\partial Q} \frac{v}{\sigma} \frac{S_i(1 - S_i)}{(QS_i)^{2\sigma}} - \frac{\partial [(QS_i)^{\sigma}]}{\partial Q} C''[(QS_i)^{\sigma}]$$

$$= -\sigma (QS_i)^{\sigma - 1} S_i \left(\frac{v}{\sigma} \frac{S_i(1 - S_i)}{(QS_i)^{2\sigma}} + C''[(QS_i)^{\sigma}]\right)$$

$$= -\sigma \frac{(QS_i)^{\sigma}}{Q} \left(\frac{v}{\sigma} \frac{S_i(1 - S_i)}{(QS_i)^{2\sigma}} + C''[(QS_i)^{\sigma}]\right) < 0 \ .$$

Furthermore, $\frac{\partial FOC_i}{\partial S_i} < 0$, because it is the SOC. Thus, $dS_i/dQ < 0$. Note that this implies that there exists at most one solution to FOC_i for any Q.

Now suppose there are two equilibria, Q^* and Q^{**}, such that $Q^* < Q^{**}$. Since $dS_i/dQ < 0$, the optimal responses $S_i(Q^*) > S_i(Q^{**})$ for all i. Since $\sum_{j=1}^{n} S_j(Q) = \sum_{j=1}^{n} S_j = 1$, this is a contradiction. Hence, there can be only one PSE. ∎

Proof of Proposition 1.T: The only thing that still needs to be proved is the expression for $\bar{r}$. It is pinned down by the individual rationality constraint holding with equality in the unique and symmetric PSE.

Together, zero profit and symmetry imply

$$\frac{v}{n} = c(x) \ ,$$

while the FOC yields

$$v \frac{r X^{r-1}}{Q} \left(1 - \frac{X^r}{Q}\right) = C'(X) \ .$$

Since $Q = nX^{\frac{1}{\sigma}}$, this reduces to

$$r \frac{v}{n} \left(1 - \frac{1}{n}\right) = XC'(X) = c'(x) \ .$$

Combining these expressions gives

$$rc(x)\left(1 - \frac{1}{n}\right) = c'(x) \ .$$

Substituting $x = c^{-1}(v/n)$,

$$rc\left[c^{-1}(v/n)\right]\left(1 - \frac{1}{n}\right) = c'\left[c^{-1}(v/n)\right] \ .$$

Hence,

$$\bar{r} = \frac{c'\left[c^{-1}(v/n)\right]}{c\left[c^{-1}(v/n)\right]} \bigg/ \left(1 - \frac{1}{n}\right) \ .$$

∎

Allowing for Concave Costs and $\eta_{C''/C'} \not\geq -1$

Model and Results

In the main text we assumed that $C''(\cdot) \geq 0$ and $\eta_{C''/C'} \geq -1$. Here we relax these assumptions. As we shall see, $\eta_{C''/C'} \not\geq -1$ does not materially change our findings. By contrast, $C''(\cdot) < 0$ does lead to fundamentally different conclusions as to the optimal level of meritocracy: for strictly concave costs, perfect meritocracy is uniquely output maximizing. It is also Pareto efficient.

Suppose that the model is as in the main text, except for the properties of $C(\cdot)$. We now assume that: 1) $C(\cdot)$ is analytic; 2) $C(0) = 0$; and 3) $C'(X) > 0$ for $X > 0$. Hence, unlike in the baseline model, we allow for increasing as well as decreasing marginal costs. Also, regularity condition (1) need not be satisfied. However, to ensure that a rise in the performance standard reduces the number of winners, we do require that marginal costs do not fall "too fast."

Specifically, let $\eta_{C'}$ denote the output elasticity of marginal costs. We assume that

$$\eta_{C'} \equiv \frac{\mathrm{d}\ln C'(X)}{\mathrm{d}\ln X} > -1 \ . \tag{48}$$

It is easily verified that all costs function of the form $C(X) = kX^{\alpha}$, $k > 0$, satisfy condition (48), both in the convex region ($\alpha \geq 1$) as well as in the concave region ($\alpha < 1$).

Since we no longer assume that (1) holds, marginal cost and marginal benefit curves may intersect any number of times. As a result, the best-response correspondence, $\hat{x}(\theta)$, can be more complicated than in the baseline model. A (mixed) best response $H_i(\cdot; \theta)$ for agent i is now a function that assigns to each standard θ a CDF with support in $\hat{x}(\theta)$, i.e., $H_i(\cdot; \theta) : \mathbb{R} \to \Delta\hat{x}(\theta)$. Let $\mathcal{H}(\cdot; \theta) \equiv \{H_i(\cdot; \theta)\}_{i \in I}$ denote a profile of best responses. A best-response profile is symmetric if $H_i(\cdot; \theta) = H_j(\cdot; \theta) \equiv H(\cdot; \theta)$ for all $i, j \in I$. Symmetric profiles are denoted by $\mathcal{H}^{sym}(\cdot; \theta)$.

Let $D(\cdot)$ denote the CDF of outputs for the population as a whole. For a best-response profile $\mathcal{H}(\cdot; \theta)$, this distribution is given by $D[\cdot; \mathcal{H}(\cdot; \theta)] = \int_{i \in I} H_i(\cdot; \theta) \, \mathrm{d}i$. As D is a simple average of $H_i(\cdot; \theta)$, $i \in I$, the set of all $D[\cdot; \mathcal{H}(\cdot; \theta)]$ is equal to the set of all $H_i(\cdot; \theta)$, namely, $\Delta\tilde{x}(\theta)$. Symmetric best-response profiles $\mathcal{H}^{sym}(\cdot; \theta)$ are characterized by a single CDF $H(\cdot; \theta)$. Slightly abusing notation, we use $H(\cdot; \theta)$ to refer to the profile $\{H(\cdot; \theta)\}_{i \in I}$, as well as to the CDF itself. For a symmetric profile $\mathcal{H}^{sym}(\cdot; \theta) = \{H(\cdot; \theta)\}_{i \in I}$, we have $D[\cdot; \mathcal{H}^{sym}(\cdot; \theta)] = H(\cdot; \theta)$.

CDF D pins down aggregate output and the mass of winners as follows:

$$O = \int_{x \in [-\infty, \infty)} e^x \, dD \left[x; \mathcal{H} \left(\cdot; \theta \right) \right] ,$$

$$W = \int_{x \in [-\infty, \infty)} \bar{F} \left(\frac{\theta - x}{\sigma} \right) dD \left[x; \mathcal{H} \left(\cdot; \theta \right) \right] .$$

Here, $\bar{F} \left(\infty \right) \equiv 0$.

A tuple $\{\theta, \mathcal{H} \left(\cdot; \theta \right)\}$ constitutes an equilibrium if $W \left[\theta, \mathcal{H} \left(\cdot; \theta \right) \right] = m$. It is denoted by $\{\theta^*, \mathcal{H} \left(\cdot; \theta^* \right)\}$. An equilibrium is symmetric, and denoted by $\{\theta^*, H \left(\cdot; \theta^* \right)\}$, if $\mathcal{H}$ is symmetric.

As in the baseline model, we may restrict attention to symmetric equilibria since, for any equilibrium $\{\theta^*, \mathcal{H} \left(\cdot; \theta^* \right)\}$, the tuple $\{\theta^*, H \left(\cdot; \theta^* \right)\}$ with $H \left(\cdot; \theta \right) \equiv \int_{i \in I} H_i \left(\cdot; \theta \right) di$ constitutes a symmetric equilibrium that generates the same aggregate output. This follows from

$$D \left(\cdot; \mathcal{H} \left(\cdot; \theta \right) \right) = \int_{i \in I} H_i \left(\cdot; \theta \right) di = H \left(\cdot; \theta \right) = D \left(\cdot; \mathcal{H}^{sym} \left(\cdot; \theta \right) \right) .$$

The following proposition establishes existence of (symmetric) equilibrium for all σ, as well as uniqueness for small σ.

Proposition 7 *Equilibrium exists. For σ sufficiently small equilibrium is unique, and $\hat{x} \left(\theta^* \right) = \{-\infty, x_s \left(\theta^* \right)\}$.*

We now show that whether the contest organizer should strive for perfect meritocracy (i.e., $\sigma \to 0$) turns on the convexity or concavity of $C \left(\cdot \right)$.

Proposition 8 *Fix (m, v).*

1. *For $C'' \left(\cdot \right) > 0$, perfect meritocracy is neither output maximizing nor Pareto efficient.*

2. *If $C'' \left(\cdot \right) = 0$, then all $\sigma \in (0, \underline{\sigma}]$ are output maximizing.*

3. *For $C'' \left(\cdot \right) < 0$, only perfect meritocracy is output maximizing. It is also Pareto efficient.*

Proposition 8 shows that merely relaxing $\eta_{C''/C'} \geq -1$ leaves our conclusions essentially unchanged: For strictly convex costs, perfect meritocracy is suboptimal, while for linear costs, any $\sigma \leq \underline{\sigma}$ is optimal. By contrast, assuming that costs are strictly concave, makes perfect meritocracy uniquely output maximizing. To see

why, notice that decreasing marginal costs give rise to economies of scale. This make the output loss from drop-outs strictly *smaller* than the output gain from the remaining contestants working harder. Hence, rather than stopping at $\underline{\sigma}$, the contest organizer should strive to fully eliminate noise in performance ranking.

Economies of scale also imply that the contest should be structured as winner-take-all: abstracting from the open set problem, in the optimal contest, a minimum number of contestants produce maximum output under perfect meritocracy—i.e., $m \to 0$, $v \to \infty$, and $\sigma \to 0$.

Proofs

In the current set-up, the function $c(\cdot)$ has the following properties.

Lemma 37 *1)* $c(-\infty) = c'(-\infty) = 0$; *2)* $c'(\cdot) > 0$, $c''(\cdot) > 0$, *for* $x > -\infty$; *3) Convexity (concavity) of* $C(\cdot)$ *is equivalent to:*

$$\forall X > 0 : C''(X) \overset{(<)}{>} 0 \iff \forall x > -\infty : c'(x)/c(x) \overset{(<)}{>} 1 \iff \forall x > -\infty : c''(x)/c'(x) \overset{(<)}{>} 1 \ .$$

Proof. The proofs of Part 1 and $c'(\cdot) > 0$ are the same as in the baseline model.

To prove that $c''(\cdot) > 0$, recall that

$$c''(x) = XC'(X) + X^2 C'''(X) \ .$$

Hence, $c''(\cdot) > 0$ takes on the same sign as

$$\frac{XC''(X)}{C'(X)} + 1 = \frac{\mathrm{d}\ln C'(X)}{\mathrm{d}\ln X} + 1 = \eta_{C'} + 1 > 0 \ ,$$

where the inequality follows from our assumption that $\eta_{C'} > -1$.

Part 3: First observe that

$$c'(x)/c(x) = XC'(X)/C(X) \ . \tag{49}$$

Now recall the following characterization of convexity (concavity) for functions that go through the origin: if $C(0) = 0$, then

$$\forall X > 0 : C''(X) \overset{(<)}{>} 0 \iff \forall X > 0 : XC'(X)/C(X) \overset{(<)}{>} 1 \ . \tag{50}$$

Combining (49) and (50), we may conclude that

$$\forall X > 0 : C''(X) \overset{(<)}{>} 0 \iff \forall x > -\infty : c'(x)/c(x) \overset{(<)}{>} 1 \ .$$

This proves the first equivalence.

Next observe that

$$c''(x)/c'(x) = 1 + XC''(X)/C'(X) . \tag{51}$$

Because $C'(\cdot) > 0$ for $X > 0$, it immediately follows that

$$\forall X > 0 : C''(X) \overset{(<)}{>} 0 \iff \forall x > -\infty : c''(x)/c'(x) \overset{(<)}{>} 1 .$$

This proves the second equivalence. ∎

Next notice that Lemmas 7 and 8 carry over essentially unchanged.

The pure best-response correspondence $\hat{x}(\theta)$ has the following properties.

Lemma 38   *1. If $\hat{x}(\theta)$ is single-valued at $\theta = \theta'$, then $\hat{x}(\theta)$ is single-valued and differentiable in a neighborhood of θ'.*

  2. If $\hat{x}(\theta)$ is multi-valued at $\theta = \theta'$, then $\hat{x}(\theta)$ is single-valued and differentiable in a left and right neighborhood of θ'. Furthermore, $\lim_{\theta \uparrow \theta'} \hat{x}(\theta) = \max \hat{x}(\theta')$ and $\lim_{\theta \downarrow \theta'} \hat{x}(\theta) = \min \hat{x}(\theta')$.

Proof.
Part 1: If $\hat{x}(\theta') > -\infty$ and single-valued, then single-valuedness and differentiability of $\hat{x}(\theta)$ in a neighborhood of θ' follow from applying the implicit function theorem (IFT) to the FOC at the point $(\theta', \hat{x}(\theta'))$.

If $\hat{x}(\theta') = -\infty$ and single-valued, then it must be that $x = -\infty$ is strictly better than any other value of x. By Berge's maximum theorem, $\pi[\hat{x}(\theta), \theta]$ is continuous in θ. It follows that, in a neighborhood of θ', $x = -\infty$ remains the unique best response. Hence, $\hat{x}(\theta)$ is single-valued and differentiable around θ' (with zero derivative).

Part 2: First observe that, in a neighborhood of θ', only output levels x that converge to elements of $\hat{x}(\theta')$ when $\theta \to \theta'$ can be elements of $\hat{x}(\theta')$. Otherwise, there exists a sequence $\{\theta_n\}_{n \in \mathbb{N}} \to \theta'$ and a sequence $\{\hat{x}(\theta_n)\}_{n \in \mathbb{N}} \to x$ such that $x \notin \hat{x}(\theta')$. This contradicts UHC of $\hat{x}(\theta')$ at θ'.

Second we show that, upon a marginal increase in θ, the payoff from playing $\min \hat{x}(\theta')$ rises strictly faster than from playing any other $x \in \hat{x}(\theta')$. To see this,

denote by $\neg \min \hat{x}\left(\theta'\right)$ an element of $\hat{x}\left(\theta'\right)$ other than $\min \hat{x}\left(\theta'\right)$.

$$
\begin{aligned}
&\left. \frac{\mathrm{d}}{\mathrm{d}\theta} \left\{ \pi\left[\min \hat{x}\left(\theta'\right), \theta\right] - \pi\left[\neg \min \hat{x}\left(\theta'\right), \theta\right] \right\} \right|_{\theta=\theta'} \\
&= \left. \frac{\mathrm{d}}{\mathrm{d}\theta} \left\{ \bar{F}\left[\frac{\theta - \min \hat{x}\left(\theta'\right)}{\sigma}\right] v - c\left[\min \hat{x}\left(\theta'\right)\right] - \bar{F}\left[\frac{\theta - \neg \min \hat{x}\left(\theta'\right)}{\sigma}\right] v + c\left[\neg \min \hat{x}\left(\theta'\right)\right] \right\} \right|_{\theta=\theta'} \\
&= \frac{v}{\sigma} f\left[\frac{\theta' - \neg \min \hat{x}\left(\theta'\right)}{\sigma}\right] - \frac{v}{\sigma} f\left[\frac{\theta' - \min \hat{x}\left(\theta'\right)}{\sigma}\right] \\
&= c'\left[\neg \min \hat{x}\left(\theta'\right)\right] - c'\left[\min \hat{x}\left(\theta'\right)\right] > 0 \ .
\end{aligned}
$$

If $\neg \min \hat{x}\left(\theta'\right) > \min \hat{x}\left(\theta'\right) > -\infty$, the last equality follows from the FOC. If $\neg \min \hat{x}\left(\theta'\right) > \min \hat{x}\left(\theta'\right) = -\infty$ the equality also holds, because $c'\left(-\infty\right) = 0$ (Lemma 37). The *in*equality follows from $c''\left(\cdot\right) > 0$.

This proves that the payoff from playing $\min \hat{x}\left(\theta'\right)$ rises strictly faster than that from all other $x \in \hat{x}\left(\theta'\right)$.

Third, the envelope theorem implies that, in a neighborhood of θ', re-optimization changes in x due to small changes in θ only have second-order effects on π. We may conclude that, in a right-neighborhood of θ', only output levels x that converge to $\min \hat{x}\left(\theta'\right)$ when $\theta \downarrow \theta'$ can be elements of $\hat{x}\left(\theta\right)$. Similarly, in a left-neighborhood of θ', only output levels x that converge to $\max \hat{x}\left(\theta'\right)$ when $\theta \uparrow \theta'$ can be elements of $\hat{x}\left(\theta\right)$.

Finally, it remains to show that, in a sufficiently small left or right neighborhood of θ', $\hat{x}\left(\theta\right)$ cannot be multi-valued.

Suppose, by contradiction, that there exists an $\varepsilon > 0$ such that, for all $\theta \in \left(\theta', \theta' + \varepsilon\right)$, $x_1\left(\theta\right), x_2\left(\theta\right) \in \hat{x}\left(\theta\right)$, $x_1\left(\theta\right) \neq x_2\left(\theta\right)$. Now consider a standard $\theta'' \in \left(\theta', \theta' + \varepsilon\right)$. Because both $x_1\left(\theta''\right)$ and $x_2\left(\theta''\right)$ belong to $\hat{x}\left(\theta''\right)$, clearly, it is not true that only output levels x that converge to $\min \hat{x}\left(\theta''\right)$ when $\theta \downarrow \theta''$ are elements of $\hat{x}\left(\theta''\right)$. However, this contradicts our finding above. Hence, we may conclude that in a right-neighborhood of θ', $\hat{x}\left(\theta\right)$ cannot contain two (or more) elements. The argument for a left-neighborhood is analogous.

We may conclude that $\hat{x}\left(\theta\right)$ is indeed single-valued around θ', while differentiability again follows from the IFT.

This completes the proof. ∎

The following lemma characterizes $\Omega^{-1}\left(\cdot\right)$.

Lemma 39 *For all $w \in \left(0, 1\right)$, $\Omega^{-1}\left(\cdot\right)$ is non-empty, and single-valued.*

Proof. From the second part of Lemma 7, we know that $\lim_{\theta \to \infty} \theta - \hat{x}\left(\theta\right) = \infty$ and $\lim_{\theta \to -\infty} \theta - \hat{x}\left(\theta\right) = -\infty$. Hence, $\lim_{\theta \to \infty} \Omega\left(\theta\right) = \lim_{\theta \to \infty} \bar{F}\left[\frac{\theta - \hat{x}\left(\theta\right)}{\sigma}\right] = 0$, while

$\lim_{\theta \to -\infty} \Omega(\theta) =$
$= \lim_{\theta \to -\infty} \bar{F}\left[\frac{\theta - \hat{x}(\theta)}{\sigma}\right] = 1$. In turn, this implies that $\inf_{\theta \in \mathbb{R}} \Omega(\theta) = 0$ and
$\sup \Omega(\theta)_{\theta \in \mathbb{R}} = 1$.

Next, choose any $w \in (\inf_{\theta \in \mathbb{R}} \Omega(\theta), \sup_{\theta \in \mathbb{R}} \Omega(\theta)) = (0, 1)$. From the first part of Lemma 7 we know that $\hat{x}(\theta)$ is non-empty, compact-valued, and UHC for all θ. Compactness implies that $\max \hat{x}(\theta)$ and $\min \hat{x}(\theta)$ exist for all θ. Because $\hat{x}(\theta)$ is non-empty, there must exist a $\theta' \in \mathbb{R}$ such that $w \in \left[\bar{F}\left[\frac{\theta - \min \hat{x}(\theta')}{\sigma}\right], \bar{F}\left[\frac{\theta - \max \hat{x}(\theta')}{\sigma}\right]\right]$. In turn, this implies that there exists a mixed best-response $H(\cdot; \theta') \in \Delta \tilde{x}(\theta')$, where the mixing is between $\min \hat{x}(\theta')$ and $\max \hat{x}(\theta')$, such that $W[\theta', H(\cdot; \theta')] = w$ and, therefore, $w \in \Omega(\theta')$.

Notice that, together, these observations imply that $\Omega^{-1}(\cdot)$ is non-empty for all $w \in (0, 1)$.

It remains to prove single-valuedness of $\Omega^{-1}(\cdot)$. To do so, we show that $\Omega(\cdot)$ is strictly decreasing.

Fix some $\theta \in \mathbb{R}$. If $\hat{x}(\theta)$ is single-valued, then Lemma 38 and differentiability of $\bar{F}(\cdot)$ imply that $\Omega(\cdot)$ is differentiable at θ. Moreover,

$$\frac{d\Omega(\theta)}{d\theta} = -f \cdot \left(1 - \frac{d\hat{x}(\theta)}{d\theta}\right) < 0 \ ,$$

where the inequality follows from Lemma 8, where we established that $d\hat{x}(\theta)/d\theta < 1$.

If $\hat{x}(\theta)$ is multi-valued, then decreasingness of $\Omega(\cdot)$ at θ follows immediately from part 2 of Lemma 38.

This completes the proof of the lemma. ∎

Lemma 39 implies that, for every $m \in (0, 1)$, there exists a unique θ^* such that $\Omega(\theta^*) = m$. Hence, we have shown:

Lemma 40 *Equilibrium exists. The equilibrium standard is unique.*

The following sequence of lemmas establishes that, for σ sufficiently small, equilibrium is (essentially) unique.

Lemma 41 *Fix a $\theta \in \mathbb{R}$. For σ sufficiently small, the marginal cost $c'(x)$ intersects marginal revenue $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ exactly twice, first from above at $x_t < \theta$, then from below at $x_s > \theta$.*

Proof. When $\sigma \to 0$, density f converges to the Dirac measure with mass 1 at zero. Hence, for all $\delta > 0$, $\lim_{\sigma \to 0} \frac{v}{\sigma} f\left(\frac{\delta}{\sigma}\right) = 0 = \lim_{\sigma \to 0} \frac{v}{\sigma} f\left(\frac{-\delta}{\sigma}\right)$ and $\lim_{\sigma \to 0} = \frac{\frac{v}{\sigma} f(0) - \frac{v}{\sigma} f\left(\frac{\delta}{\sigma}\right)}{\delta} = -\infty = -\lim_{\sigma \to 0} = \frac{\frac{v}{\sigma} f(0) - \frac{v}{\sigma} f\left(\frac{-\delta}{\sigma}\right)}{\delta}$. At $x = \theta$, $c'(x)$ is finite and differentiable. Moreover, $\lim_{\sigma \to 0} \frac{v}{\sigma} f(0) = \infty$. Hence, for small σ, there must exist an $x_t < \theta$ where $c'(x)$ intersects $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ from above. Moreover, $x_t \uparrow \theta$ when $\sigma \to 0$. Similarly, there must exist an $x_s > \theta$ where $c'(x)$ intersects $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ from below, and $x_s \downarrow \theta$ for $\sigma \to 0$.

The previously mentioned properties of $c'(x)$ and $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ also imply that, for small σ, any additional intersection points must be in the tails, i.e., remain bounded away from θ when $\sigma \to 0$. It remains to show that, in fact, no such intersection points exist.

First we show that there exists no intersection point $x_{s_0} > x_s$ when $\sigma \to 0$. This follows from: 1) $\lim_{\sigma \to 0} \frac{v}{\sigma} f\left(\frac{\varepsilon}{\sigma}\right) = 0$ for all $\varepsilon > 0$; 2) $c'(x) = XC'(X) > 0$ for all $x = \ln X > -\infty$; 3) and $c'(x) = XC'(X)$ remains bounded away from 0 when $X \to \infty$.

For a hypothetical intersection point $x_{t_0} < x_t$, we will show that for σ sufficiently small,

$$\frac{d}{dx}\left[\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x)\right]\bigg|_{x = x_{t_0}} > 0 .$$

Clearly, the sign of this derivative is inconsistent with $c'(x)$ dropping below $\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right)$ to the left of x_t and, thus, with the existence of x_{t_0}.

Observe that $d\left[\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x)\right] / dx > 0$ iff

$$\frac{\frac{1}{\sigma} f'\left(\frac{\theta - x}{\sigma}\right)}{f\left(\frac{\theta - x}{\sigma}\right)} + \frac{c''(x)}{c'(x)} < 0 . \tag{52}$$

(Notice that (52) is equivalent to a violation of *soc* in (4). In other words, we will show that all extrema to the left of x_t are minima. Because x_t itself is a minimum, and minima and maxima must alternate, this implies there can be no other extrema to the left of x_t.)

Recall that f is strictly log-concave. Hence, f'/f is strictly decreasing in its argument $\frac{\theta - x}{\sigma}$. Because any x_{t_0} is bounded away from θ,

$$\lim_{\sigma \to 0} \frac{\frac{1}{\sigma} f'\left(\frac{\theta - x_{t_0}}{\sigma}\right)}{f\left(\frac{\theta - x_{t_0}}{\sigma}\right)} = -\infty .$$

Let $X_{t_0} \equiv e^{x_{t_0}}$. To prove that the inequality in (52) holds, we still need to show that $c''(x_{t_0}) / c'(x_{t_0}) = 1 + X_{t_0} C''(X_{t_0}) / C'(X_{t_0})$ remains finite as $\sigma \to 0$. Because $x_{t_0} < \theta$, it suffices to show that $XC''(X) / C'(X) < \infty$ for $X \in (0, e^\theta)$. Notice that the only way the last inequality could be violated is if $X \to 0$.

Define $k \equiv \min \{l \in \mathbb{N} : C^{(l)}(0) > 0\}$, where $C^{(l)}$ denotes the l-th derivative. Notice that k must exist and be finite. Otherwise, $C(X) = 0$ in a neighborhood of $X = 0$, which is inconsistent with our assumption that $C'(X) > 0$ for all $X > 0$. Applying L'Hopital's rule k times yields

$$\lim_{X \to 0} \frac{X C''(X)}{C'(X)} = k + \lim_{X \to 0} \frac{X C^{(k+1)}(X)}{C^{(k)}(X)} < \infty \,,$$

where the inequality follows from the fact that $C^{(k)}(0) > 0$.

We may conclude that the inequality in (52) indeed holds and, therefore,

$$\frac{\mathrm{d}}{\mathrm{d}x} \left[\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) / c'(x) \right] \Bigg|_{x = x_{t_0}} > 0 \,.$$

This completes the proof of the lemma. ■

Lemma 41 implies:

Corollary 2 *Fix a $\theta \in \mathbb{R}$. For σ sufficiently small, $\hat{x}(\theta) \subset \{-\infty, x_s(\theta)\}$.*

Next observe that Lemmas 13, 14, and 15 carry over unchanged. Hence, so does Lemma 17. This allows us to show:

Lemma 42 *For σ sufficiently small, equilibrium is unique and $\hat{x}(\theta^*) = \{-\infty, x_s(\theta^*)\}$. Both elements of $\hat{x}(\theta^*)$ are played with strictly positive probability.*

Proof. From Lemma 40 we already know that θ^* is unique, while from Lemma 17 we know that $\hat{x}(\theta^*_\sigma) = \{-\infty, x_s(\theta^*_\sigma)\}$ for small σ. Finally, the same argument as in the proof of Lemma 12 establishes that there exists a unique mixture $(1 - \hat{\gamma}^*, \hat{\gamma}^*)$, $\hat{\gamma}^* \in (0, 1)$ that clears the market. Hence, in the unique equilibrium, contestants mix between $x_s(\theta^*)$ and $-\infty$ with strictly positive probabilities $\hat{\gamma}^*$ and $1 - \hat{\gamma}^*$, respectively. ■

Together, Lemmas 40 and 42 prove Proposition 7 in the body of the Appendix.

Finally, we provide a proof of Proposition 8. It relies on the following lemma.

Lemma 43 *For $\sigma > 0$ sufficiently small,*

$$\frac{dO/O}{d\sigma/\sigma} = \frac{c'/c - 1}{\frac{1}{\sigma} f'/f + c''/c'} \,.$$

Proof. From Proposition 7 we know that, for small σ, $\hat{x}\left(\theta_\sigma^*; \sigma\right) = \{-\infty, x_s\left(\theta_\sigma^*; \sigma\right)\}$. The remainder of the proof is the same as the proof of the first part of Lemma 4. ∎

Proof of Proposition 8:

Together, Lemma 43 and the SOC imply that, for small σ,

$$\mathrm{d}O/\left(\mathrm{d}\sigma/\sigma\right) \gtrless 0 \iff c'/c \gtrless 1 \;.$$

In turn, we know from Lemma 37 that $c'/c \gtreqless 1 \iff C'' \gtreqless 0$. Hence,

$$\mathrm{d}O/\left(\mathrm{d}\sigma/\sigma\right) \gtrless 0 \iff C'' \gtrless 0 \;.$$

Finally, because $\hat{x}\left(\theta_\sigma^*; \sigma\right) = \{-\infty, x_s\left(\theta_\sigma^*; \sigma\right)\}$ for small σ, we know that $\pi^* = 0$.

Combining these observations proves Parts 1 and 3.

To prove Part 2, simply observe that for $C'' = 0$, the conditions of the baseline model are fully satisfied. Hence, the relevant result from Theorem 1 carries over unchanged.   ∎