Working Papers

www.cesifo.org/wp

The Ponds Dilemma

John Morgan
Dana Sisak
Felix Várdy

CESIFO WORKING PAPER NO. 5539
CATEGORY 11: INDUSTRIAL ORGANISATION
SEPTEMBER 2015

ISSN 2364-1428

CESifo *Working Paper No. 5539*

The Ponds Dilemma

Abstract

Is it better to be a big fish in a small pond or a small fish in a big pond? To find out, we study self-selection into contests among a large population of heterogeneous agents. Our simple and highly tractable model generates many testable and sometimes surprising predictions. For example: 1) Entry into the big pond–in terms of show-up fees, number or value of prizes–is non-monotonic in ability; 2) Entry into the more meritocratic (i.e., discriminatory) pond is likewise non-monotonic, exhibiting two interior extrema and disproportionately attracting contestants of very low ability; 3) Changes in reward structures can produce unexpected selection effects. For instance, offering higher show-up fees may lower entry, while raising the value of prizes or making a contest more meritocratic may lower the average ability of entrants.

JEL-Code: J240, D440.

John Morgan
Haas School of Business and
Department of Economics
University of California at Berkeley
Berkeley / CA / USA
morgan@haas.berkeley.edu

Dana Sisak
Department of Economics
Erasmus School of Economics
Erasmus University Rotterdam
Rotterdam / The Netherlands
sisak@ese.eur.nl

Felix Várdy
International Monetary Fund
USA - 20431 Washington DC
fvardy@imf.org

September 8, 2015
For valuable comments and suggestions we thank Thomas Chapman, Josse Delfgaauw, Paolo Dudine, Robert Dur, Bob Gibbons, Mitchell Hoffman, Sam Ouliaris, Sander Renes, and seminar participants at Tinbergen Amsterdam, Bielefeld, Cambridge, CEU Budapest, LSE, NOVA Lisbon, and Rotterdam.

1 Introduction

Contests form an integral part of modern life; sometimes explicitly, as in innovation tournaments like the X-prize, but more often implicitly, as when individuals compete for promotions in an organization. In a world where contests are ubiquitous, people often have to choose which contest to enter. This choice may confront them with the familiar ponds dilemma: is it better to be a big fish in a small pond or a small fish in a big pond? For instance, a golfer struggling on the PGA tour may well consider his options on the Asian tour. A freshly minted JD from a reputable law school may have a choice between a "white shoe" law firm in New York or a less competitive, and less lucrative, practice in Peoria. A biotech start-up may have to decide whether to focus on a risky "blockbuster" drug or on a less risky and less profitable extension of an existing patent. Economically, the important point is that the structural properties of contests (ponds)—such as the number and value of prizes—not only affect behavior *within* but also selection *across* contests.

In this paper we present an analysis of the ponds dilemma in a simple and highly tractable framework. The parsimony of our modelling approach allows us to derive many testable and sometimes surprising predictions. For example: 1) Entry into the big pond—i.e., the contest richer in show-up fees or in the number or value of prizes—is non-monotonic in ability; 2) When the ponds differ in terms of meritocracy (i.e., discriminatoriness), entry into the more meritocratic contest takes on *two* interior extrema, first reaching a minimum, then a maximum; 3) All else equal, agents of very low ability disproportionately enter the *more* meritocratic contest. Nonetheless, this contest is exclusive, i.e., it attracts only a minority of the population; 4) Changes in reward structures can produce unexpected selection effects. For instance, higher show-up fees may reduce entry, while higher prizes or more meritocracy may lower the average ability of entrants.

When studying self-selection and entry, the usual first step is to postulate a fixed outside option available to all agents. We examine this scenario and show that the probability of entering the contest is strictly increasing in ability. In terms of the ponds dilemma, big fish tend to choose the big pond (i.e., they compete), while small fish tend to choose the small pond (i.e., they bow out). In light of this result, which aligns nicely with most people's intuition, one may reinterpret the bulk of the literature on contests as pertaining to an imperfectly truncated distribution of ability types for whom participating in the contest is more profitable than a fixed outside option.

This simple model of selection proves misleading when, rather than a fixed outside option, the alternative to participating in one contest is participating in another contest. When choosing between contests, agents weigh the potential rewards against the ability-dependent chances of success. The key observation is that, even though success probabilities in both contests are increasing in ability, the likelihood ratio is not. For those of extreme ability (high or low), the likelihood ratio is approximately one, since chances of success (or lack thereof) are essentially the same regardless of the pond chosen. For middling sorts, by contrast, the likelihood ratio favors the small pond. This non-monotonicity carries over to selection and, as a consequence, the common intuition for the ponds dilemma fails.

Concretely, we study how a large, heterogenous population of risk-neutral agents self-select across two mutually exclusive contests. We consider four dimensions in which the contests may differ: 1) show-up fees; 2) the number of (equal) prizes; 3) the value of these prizes; and 4) discriminatoriness. The last aspect corresponds to noisiness in performance evaluation and can be interpreted as a measure of meritocracy. While the existing literature is cognizant of the importance of discriminatoriness in determining contestants' behavior, interpreting discriminatoriness in terms of meritocracy and connecting it to entry is a central contribution of our paper.

When agents self-select across contests, the endogenously determined ability distributions no longer correspond to simple, truncated versions of the population at large. Suppose, for instance, that the contests only differ in terms of show-up fees. Then the contest offering the higher show-up fee disproportionately attracts those of extreme ability (both high and low) and repels middling sorts. Thus, even if the underlying ability distribution is unimodal, the conditional distribution in the high show-up fee contest tends to be bimodal.

Selection incentives are best understood by distinguishing between a parameter's *direct* and its *indirect* effect on payoffs. When contests differ only in terms of show-up fees, the direct effect of a higher show-up fee is to make the contest more attractive to all contestants. However, owing to the scarcity of prizes, more entry raises the performance standard required to succeed. A higher standard is particularly costly for agents "on the bubble," i.e., for middling sorts most uncertain about winning or losing. For them, the ratio of success probabilities favors the low-fee contest. This is the indirect effect. Together, the direct and indirect effect cause middling sorts to be underrepresented in the high-fee contest, while extreme types are overrepresented.

When contests differ in the number of (equal) prizes on offer, only indirect effects are present. To see why, notice that the number of prizes has no direct effect on payoffs, but affects behavior only indirectly, through its effect on performance standards. Offering more prizes reduces a contest's standard, which is most valuable for agents on the bubble but less relevant for infra-marginal types. As a result, middling sorts disproportionately enter the prize-rich contest, making abilities there more homogeneous than in the population at large.

A difference in prize values has direct as well as indirect effects, both of which are type-dependent. While show-up fees are equally valuable to all, higher prizes are most attractive to those anticipating to win, i.e., agents of high ability. For them, the positive direct effect of a higher prize dominates the negative indirect effect of a higher standard. As a result, high types are overrepresented in the high-prize contest. For middling sorts, it is the negative indirect effect that dominates. So they are overrepresented in the low-prize contest. Since agents of very low ability stand almost no chance of winning in either contest, they enter each contest with virtually equal probability. In other words, selection effects vanish in the lower tail. As we show, one noteworthy implication is that a contest may well raise the value of its prizes, only to see a fall in the average ability of its contestants.

Perhaps the most subtle difference between contests lies in their degree of meritocracy. In our model, a decrease in meritocracy corresponds to a mean-preserving spread in the noisiness of performance evaluations. Since agents are risk-neutral, such a spread might seem immaterial. Indeed, meritocracy would not matter if measured performance and payoffs varied proportionately, as in a Roy (1951) model. In a contest, however, payoffs are a highly non-linear function of measured performance. To see the effect, notice that an increase in meritocracy reduces the chance that an agent's performance is mis-evaluated. This is beneficial for high types, who worry about an evaluation that does not reflect their true ability, and detrimental to low types, who actually require a mis-evaluation in order to succeed. Nonetheless, monotone selection still fails since, at the extremes, agents care little about meritocracy. For very high types, even an adverse performance evaluation suffices to win, while for very low types, even an advantageous evaluation will not save the day. Hence, meritocracy does produce positive sorting, but with waning power in the tails. We also show that, provided pecuniary motives dominate, the more meritocratic contest is exclusive; that is, it attracts only a minority of the population. Jointly, the loss of selection power in the tails and exclusivity have the counterintuitive implication that agents of very low ability disproportionately enter the more meritocratic contest. As with higher prizes, a rise in meritocracy may cause a drop in the average ability of contestants.

Strictly speaking, the results and intuitions discussed so far pertain to contests differing in one dimension only. We also characterize selection when contests differ in multiple—or even all—dimensions simultaneously. In that case, show-up fees alone determine selection of very low types, while the sum of show-up fees and prize values determine the selection of very high types. Meritocracy shapes behavior in between these extremes, producing two interior extrema. Finally, the number of prizes affects selection only indirectly, through its effect on standards.

Before proceeding, a comment on methodology is in order. In the extant literature on contests, the population of contestants is generally taken to be exogenous, while effort levels are endogenous. Initially, we focus on the polar opposite case, i.e., endogenous entry with exogenous effort. This allows us to derive crisp results with clear intuitions, for all possible parameter values of the model. Subsequently, we show that our results carry over to environments with endogenous populations *and* endogenous effort, provided that the differences in structural parameters across contests are not too large.

The paper proceeds as follows. In Section 2 we introduce the baseline model with exogenous effort and endogenous selection. In Section 3 we prove existence of equilibrium. Section 4 starts off by illustrating selection behavior by means of a numerical example. It then proceeds with a formal analysis of selection, both for univariate as well as for multivariate differences between contests. In Section 5 we extend the model to allow for endogenous effort and show that our previous results carry over, provided the contests' structural parameters are "close." Section 6 discusses the related literature. Finally, Section 7 concludes. While intuitions for our results are provided in the main text, formal proofs have been relegated to an appendix. Mathematica code implementing the numerical examples is available from the authors.

2 Model

Consider a unit mass of risk-neutral agents with heterogenous abilities $a \in \mathbb{R}$. Abilities are distributed according to an atomless cumulative distribution function (CDF) G with strictly positive probability density function (PDF) g. Each agent must choose between two contests, 1 and 2. An agent of ability a entering Contest $i \in \{1, 2\}$ has measured performance $y_i(a)$, where

$$y_i = a + \varepsilon_i \ .$$

The random variable (RV) ε_i represents noise in performance measurement. Its dispersion typically differs across contests but not across agents within a contest, while its realizations are independent across contests and agents.[1].

We assume that the distribution of ε_i belongs to a location-scale family with location parameter zero and scale parameter $\sigma_i > 0$. Noise ε_i admits a CDF $F\left(\frac{\varepsilon_i}{\sigma_i}\right)$ with associated PDF $\frac{1}{\sigma_i} f\left(\frac{\varepsilon_i}{\sigma_i}\right)$. Density $f(\cdot)$ is assumed to be single-peaked around zero and strictly positive on $\mathbb{R}$. Moreover, f is twice continuously differentiable and strictly log-concave. We interpret precision $1/\sigma_i$ as a measure of the meritocracy of the contest. Indeed, the greater σ_i, the less reliable the performance evaluation process, and, hence, the less an agent's measured performance y_i reflects his true ability a. Notice that this setup allows us to measure meritocracy by means of a single parameter, while it is still rich enough to encompass many, if not most, standard distributions.

Sometimes we will assume that f also satisfies the following technical condition.

Condition 1 $\frac{f''}{f'} / \frac{f'}{f}$ *is strictly increasing in* $|\varepsilon|$ *for* $\varepsilon \neq 0$*.*

Strict log-concavity of f is equivalent to $\frac{f''}{f'} / \frac{f'}{f} < 1$. Hence, Condition 1 does not imply log-concavity, nor is it implied by it. While we do not have an economic interpretation, to the best of our knowledge, virtually all commonly-used, strictly log-concave probability distributions satisfy this condition, including the Normal, Logistic, Extreme Value, and Gumbel distributions.[2] Condition 1 is only relevant for $\sigma_1 \neq \sigma_2$, when it guarantees single-peakedness of the likelihood ratio $\lambda(a) \equiv f\left(\frac{\theta_1 - a}{\sigma_1}\right) / f\left(\frac{\theta_2 - a}{\sigma_2}\right)$ for all $\theta_1, \theta_2 \in \mathbb{R}$. Even

[1] We have cast the model as one where performance is deterministic but noisily measured. Alternatively, one may suppose that performance itself is stochastic. The first interpretation is appropriate for settings where measurement is difficult or highly subjective. The second interpretation applies when actual performance is subject to random factors outside the control of contestants, such as in most sports. A combination of noisy performance and noisy measurement can also be accommodated.

[2] The only standard distributions we are aware of that do not satisfy Condition 1 are the Laplace, Pareto, and Lognormal distributions. None of these are admissible, however, because they violate strict log-concavity. One way to break Condition 1 while potentially still satisfying our other assumptions is to have a single-peaked density with multiple inflection points on each side of the peak.

then, its main role is to simplify exposition. For maximum clarity, we explicitly invoke the condition whenever we rely on it.

Regardless of performance, an agent entering Contest $i \in \{1, 2\}$ receives a show-up fee $w_i \geq 0$.[3] In addition, the agent earns a prize $v_i > 0$ iff he is among the winners of the contest. The set of winners in Contest i consists of the mass $m_i > 0$ of agents with the highest performance measures. Prizes are scarce overall, i.e., $m_1 + m_2 < 1$. We refer to show-up fees w_i, values of prizes v_i, number of prizes m_i, and measures of meritocracy σ_i as the structural parameters of the contests. Notice that the quantiles of measured performance among a continuum of agents are perfectly predictable. Hence, the condition for winning in Contest i is characterized by a *deterministic* performance threshold, or standard, which we denote by $\theta_i \in [-\infty, \infty)$.

To summarize, an agent of ability a choosing Contest i with standard θ_i enjoys an expected pecuniary payoff

$$\pi_i (a, \theta_i) = w_i + v_i \bar{F} \left(\frac{\theta_i - a}{\sigma_i} \right) .$$

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

In addition to valuing money, agents also derive (potentially small) non-pecuniary payoffs from participating in each contest. These payoffs are idiosyncratic and might derive from the nature of the task required, the physical location of the contest, the personalities of the organizers, and so on.[4] Let δ denote the *difference* in an agent's non-pecuniary payoffs from participating in Contest 2 versus Contest 1. Hence, an agent non-pecuniarily prefers Contest 1 iff $\delta \leq 0$. We assume that the realizations of δ have full support on $\mathbb{R}$ and are i.i.d. across agents. Notice that this rules out situations where the task in one contest but not in the other is satisfying for high types, say, but much less so for low types. The distribution of δ belongs to a location-scale family with location parameter and median $\tau \in \mathbb{R}$ and scale parameter $\rho > 0$. The CDF and associated PDF of δ are $\Gamma \left(\frac{\delta - \tau}{\rho} \right)$ and $\frac{1}{\rho} \gamma \left(\frac{\delta - \tau}{\rho} \right)$, respectively. Agents enter the contest that offers them the higher total expected payoff, which is equal to

[3] Technically, negative show-up fees (i.e., entry fees) would not pose a problem. However, they do imply that expected payoffs are negative for agents of sufficiently low ability. In that case, one may want to extend the model to allow agents to earn zero by opting out entirely. While this complicates the analysis, entry into both contests remains non-monotonic in ability, "trailing off" in the lower tail. (Cf. Proposition 9 in Appendix A.) Calculations are available from the authors upon request.

[4] In addition to realism, an advantage of including non-pecuniary payoffs is that they smooth out agents' selection behavior. That is, non-pecuniary payoffs make agents' entry probabilities a continuous function of pecuniary payoffs and ability, rather than a step function. As a consequence, entry probabilities not only indicate an agent's preference for one contest over the other, but also express the *intensity* of that preference.

the sum of pecuniary and non-pecuniary payoffs. In case of indifference, they flip a coin.

Let $H_i(a)$ denote the cumulative *mass* function (CMF) of endogenously determined abilities in Contest i. That is, $H_i(a)$ is the measure of entrants into i with ability a or lower. Provided it exists, the corresponding mass density function (MDF) is denoted by $h_i(a)$, i.e., $h_i(a) \equiv dH_i(a)/da$. Because we have normalized the population mass to 1, the CMFs in the two contests must add up to the CDF of abilities in the population as a whole. That is, $\forall a \in \mathbb{R}$, $H_1(a) + H_2(a) = G(a)$. Moreover, $\lim_{a\to\infty} H_i(a)$ must equal the fraction of the population entering Contest i, which we denote by $\Pr i$. Finally, let $G_i(a)$ denote the CDF of endogenously determined abilities in Contest i, i.e., $G_i(a) \equiv H_i(a)/\Pr i$. The corresponding PDF is $g_i(a)$.

To close the model, we offer a formal definition of market clearing and define equilibrium of the game as a whole. If fewer than a mass m_i of agents enter Contest i, then $\theta_i = -\infty$ and all entrants receive a prize v_i. In that case, we say that Contest i is *uncompetitive.* A contest is said to be *competitive* when strictly more than m_i have entered. In a competitive contest, standard θ_i adjusts such that the mass of winners W_i —i.e., contestants whose performance exceeds the standard—equals the mass m_i of prizes. That is, θ_i solves

$$W_i(\theta_i) \equiv \int_{-\infty}^{\infty} \bar{F}\left(\frac{\theta_i - a}{\sigma_i}\right) dH_i(a) = m_i \ . \tag{1}$$

Agents simultaneously and independently choose which contest to enter. A Bayesian Nash equilibrium of the game consists of a tuple $\{(H_1^*(a), H_2^*(a)), (\theta_1^*, \theta_2^*)\}$ of CMFs $H_i^*(a)$ and standards θ_i^* such that: 1) conditional on H_i^*, standard θ_i^* clears the market for prizes in Contest i; and 2) profit maximizing entry decisions induced by (θ_1^*, θ_2^*) give rise to CMFs $\{H_1^*(a), H_2^*(a)\}$.[5]

[5]Notice that the analysis remains unchanged if agents choose their contest sequentially. The reason is that, due to the atomicity of agents, 'unilateral' deviations do not affect the payoffs of other agents. Hence, any Bayesian Nash equilibrium of the simultaneous game corresponds to a perfect Bayesian equilibrium of the sequential game, and vice versa. For the same reason, we could allow agents to switch contests upon observing the contest choices—and even performance evaluations—of other agents. What we cannot allow for is switching upon observing one's *own* performance evaluation.

3 Equilibrium

We solve for equilibrium in three steps. First we show that, conditional on entry decisions characterized by (H_1, H_2), there exist unique performance standards (θ_1, θ_2) that clear the market for prizes in each contest. Second, we show that standards (θ_1, θ_2) induce a unique pair of CMFs (H_1, H_2). Together, these two steps define a mapping from the space of standards into itself. Finally, we show that there exists a pair (θ_1^*, θ_2^*) that constitutes a fixed point of the system. Notice that such a fixed point gives rise to an equilibrium $\{(H_1^*(a), H_2^*(a)), (\theta_1^*, \theta_2^*)\}$.

Standards Conditional on Entry

Using the market-clearing condition (1), our first lemma shows that, for a given CMF of abilities H_i, a contest's standard θ_i is uniquely determined.

Lemma 1 *For every H_i, $i \in \{1,2\}$, there exists a unique standard $\theta_i \in [-\infty, \infty)$ that clears the market for prizes in Contest i.*

From the market-clearing condition (1) it is immediate that standards are higher when prizes are scarcer, all else equal. By contrast, conditional on H_i, neither w_i nor v_i have any influence on θ_i. The reason is that, in this version of the model, "effort" is equal to ability and, hence, exogenous.

Entry Conditional on Standards

We now derive the unique pair of CMFs (H_1, H_2) that result from standards (θ_1, θ_2). For given (θ_1, θ_2), an agent of ability a enters Contest 1 iff

$$\delta \leq \pi_1(a, \theta_1) - \pi_2(a, \theta_2) \ .$$

Hence, the probability, $\Pr i(a)$, that the agent enters Contest $i \in \{1,2\}$ is

$$\Pr 1(a) = \Gamma\left(\frac{\pi_1 - \pi_2 - \tau}{\rho}\right) = 1 - \Pr 2(a) \ .$$

The uniquely determined MDF h_i is given by

$$h_i(a) = g(a) \Pr i(a) \ , \tag{2}$$

while the associated CMF $H_i(a)$ is found by integrating h_i up to a.

Fixed Point

The previous steps define a function, ξ, from the space of standards $[-\infty, \infty) \times [-\infty, \infty)$ into itself. Specifically, each pair of standards (θ_1, θ_2) gives rise to a unique pair of MDFs (h_1, h_2) according to equation (2). In turn, each pair of MDFs (h_1, h_2) with associated CMFs (H_1, H_2) give rise to a unique pair of standards (θ_1, θ_2) according to Lemma 1. Endowing the set of MDFs with the $\|\cdot\|_{L^1}$ norm—i.e., $\|h_i\|_{L^1} \equiv \int h_i(a)\, da$—it is easily verified that these mappings are continuous. Finally, notice that the function ξ is bounded from above. To see this, observe that θ_i takes on its largest and finite value when all agents enter Contest i. We may conclude that ξ is a continuous function on a compact space. Brouwer's fixed-point theorem then implies that ξ has a fixed point, which we denote by (θ_1^*, θ_2^*). In turn, standards (θ_1^*, θ_2^*) induce a pair of (internally consistent) CMFs (H_1^*, H_2^*). Hence, equilibrium exists.

A *symmetric baseline* refers to a situation where the values of the structural parameters w, m, v, and σ are the same in both contests and, on average, the two contests are equally attractive in non-pecuniary terms, i.e., $\tau = 0$. In that case, equilibrium is unique and takes on a particularly simple form:

Proposition 1 *For all structural parameters, equilibrium exists. In a symmetric baseline, equilibrium is unique, standards are the same in both contests, and 50% of every ability type enter each contest.*

4 Selection

In this section we study the selection effects of differences in structural parameters across contests. To motivate our analysis, we begin by presenting an example that illustrates the workings of the model and the selection patterns that may arise.

4.1 Example 1

Suppose that ability is Standard Normally distributed, differences in non-pecuniary payoffs are $\delta \sim N(\tau = .05, \rho = .05)$, and noise in performance measurement is $\varepsilon_i \sim Logistic(0, \sigma_i)$,

$i \in \{1,2\}$.[6] Let $(w_1, w_2) = (1.1, 1)$, $(m_1, m_2) = (.1, .2)$, $(v_1, v_2) = (1, 1.1)$, and $(\sigma_1, \sigma_2) = (.6, 1)$. That is, Contest 1 is more meritocratic than Contest 2 and pays a 10% higher show-up fee. However, Contest 2 offers twice as many prizes and their value is 10% higher. On average, Contest 2 is somewhat more attractive in non-pecuniary terms.

Case 1: For the parameter values above, equilibrium standards are $(\theta_1^*, \theta_2^*) = (1.04, 1.02)$, and the fraction of the population entering each contest is $(\Pr 1, \Pr 2) = (.24, .76)$. Figure 1(a) depicts $\Pr 1(a)$, the probability of entering Contest 1 as a function of ability. The figure shows that selection is highly non-monotonic, with two interior extrema. The resulting PDFs, $g_i(a)$, are given in Figure 1(b). Notice that the distribution of abilities in Contest 1 is bimodal. That is, Contest 1 attracts both the best and the worst. Average abilities are $(E_1[a], E_2[a]) = (.44, -.14)$. Hence, average ability is higher in the contest offering fewer and lower-value prizes.

Case 2: Now reduce ρ to .0005. This means that there is virtually no heterogeneity in how agents perceive the two contests in non-pecuniary terms. As a result, agents with the same ability enter the same contest, and selection is essentially deterministic. This is illustrated in Figure 1(c). Figure 1(d) depicts the resulting ability distributions in the two contests. Standards are $(\theta_1^*, \theta_2^*) = (1.06, 1.06)$, $(\Pr 1, \Pr 2) = (.18, .82)$ and $(E_1[a], E_2[a]) = (.85, -.19)$.

Case 3: Next, suppose ρ is very large; say, 10^4. In that case, the extreme dispersion of non-pecuniary payoffs dominates all other considerations. As a result, selection is virtually indistinguishable from that in a symmetric baseline, with approximately 50% of every ability type entering each contest.

Case 4: Finally, reset ρ to .05 and raise v_1 from 1 to 15 and w_1 from 1.1 to 2. In this case, $(\theta_1^*, \theta_2^*) = (1.85, -\infty)$ and $(\Pr 1, \Pr 2) = (.82, .18)$. This means that there are fewer entrants into Contest 2 than there are prizes. Hence, Contest 2 is uncompetitive and all entrants earn $w_2 + v_2 = 2.1$ in pecuniary payoffs. The resulting selection behavior is depicted in Figure 1(e). Notice that the probability of entering competitive Contest 1 is strictly increasing in ability. The resulting ability distributions are shown in Figure 1(f). Average abilities are $(E_1[a], E_2[a]) = (.31, -1.43)$.

[6] In anticipation of revisiting the current example in the model with endogenous effort, we assume that noise is Logistic rather than Normal. Normal noise does not materially change the example. However, in the endogenous-effort model, the second-order condition for optimal effort is more easily satisfied with Logistic than with Normal noise.

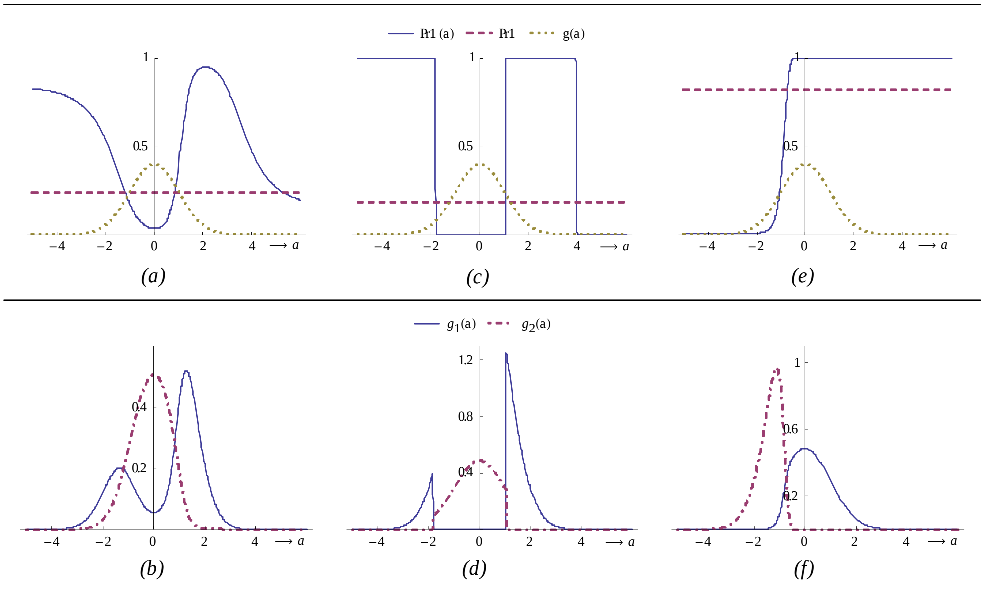

Figure 1: For Example 1 in Section 4.1, panels *(a)* and *(c)* depict the probability of entering Contest 1 as a function of ability when $\rho = .05$ and $\rho = .0005$, respectively. The resulting ability distributions are given in panels *(b)* and *(d)*. Panels *(e)* and *(f)* depict selection and ability distributions when Contest 1 is so much more attractive in pecuniary terms that Contest 2 is uncompetitive.

4.2 Analysis

In Example 1 multiple forces were at play simultaneously, resulting in the rather complex selection behavior of Figure 1. In this section we disentangle these forces and show that the selection properties of the example are in fact generic.

Before proceeding, we may dispense with one of the model parameters by observing that $\tau \neq 0$ is isomorphic to a difference in show-up fees, w. To see this, observe that

$$\Pr 1\,(a) = \Gamma \left\{ \frac{1}{\rho} \left[w_1 - w_2 - \tau + v_1 \bar{F}\left(\frac{\theta_1 - a}{\sigma_1}\right) - v_2 \bar{F}\left(\frac{\theta_2 - a}{\sigma_1}\right) \right] \right\} \, .$$

Since only the net of $w_1 - w_2 - \tau$ figures in this expression, in the remainder of the paper we normalize τ to zero and incorporate into the show-up fees any median difference in non-pecuniary payoffs across contests.

4.2.1 Uncompetitive Case

As illustrated in Case 4 of Example 1, when one contest is overwhelmingly more attractive than the other, the less attractive contest obtains so few entrants that it becomes uncompetitive. That is, the number of prizes, m_i, exceeds the number of entrants, $\Pr i$, and all entrants win a prize (see Appendix B for details, and notice that at most one contest can be uncompetitive, since the population has unit mass and $m_1 + m_2 < 1$). The next proposition shows that, in this case, selection into the "big pond" is always monotone. As a result, ability distributions in the two ponds are ordered by first-order stochastic dominance (FOSD).

Proposition 2 *When one contest is uncompetitive, the probability of selecting into the competitive contest is strictly increasing in ability. The ability distribution in the latter FOSDs the one in the former.*

When $\rho \to 0$, sorting becomes deterministic. Agents enter the competitive contest iff their ability exceeds some threshold $\grave{a} \in \mathbb{R}$.

To see why agents of higher ability increasingly favor the competitive contest, recall that $\Pr 1(a) = \Gamma\left[\frac{\pi_1^*(a)-\pi_2^*(a)}{\rho}\right]$. When Contest 2 (say) is uncompetitive, all entrants into this contest obtain $w_2 + v_2$ in pecuniary payoffs. So the payoff difference reduces to

$$\pi_1^*(a, \theta_1^*) - \pi_2^*(a, -\infty) = w_1 - w_2 - v_2 + v_1 \bar{F}\left(\frac{\theta_1^* - a}{\sigma_1}\right) .$$

Ability affects payoffs only through the chance of winning in Contest 1, $\bar{F}\left(\frac{\theta_1^*-a}{\sigma_1}\right)$, which is strictly increasing in a. Therefore, $\Pr 1(a)$ is also increasing. When $\rho \to 0$ the effect of non-pecuniary preferences vanishes, making the sign of $\pi_1^*(a) - \pi_2^*(a)$ the sole selection criterion. As a result, $\Pr 1(a)$ becomes a step function and selection deterministic.

An agent's pecuniary payoff in the uncompetitive contest is $w_i + v_i$, regardless of ability. Hence, the uncompetitive case is isomorphic to a model with a single contest and a fixed outside option. We may conclude:

Corollary 1 *In a selection model with a single competitive contest and a fixed outside option, the probability of entering the contest is strictly increasing in ability. When $\rho \to 0$, agents enter the contest iff their ability exceeds some threshold.*

4.2.2 Competitive Case

The competitive case, and the main focus of our analysis, occurs when the number of entrants into each contest exceeds the number of prizes on offer. The following proposition shows that this case pertains in and around symmetric baselines.

Proposition 3 *Both contests are competitive in a symmetric baseline and in a neighborhood of structural parameters around it. Moreover, this competitive region remains non-degenerate when* $\rho \to 0$.

In a symmetric baseline, the two contests are equally attractive in pecuniary terms. As a result, 50% of each ability type enter each contest and, since $m_1 = m_2 < \frac{1}{2}$, both contests are competitive. Proposition 3 shows that this situation extends beyond symmetric baselines, provided the structural parameters in the two contests are not too far apart.

The remainder of the paper focuses on the competitive case and, if necessary, constrains the parameter space accordingly. In some cases, as when contests only differ in meritocracy, no constraints are needed. (See Lemma 16 in Appendix B.) In other cases, as when contests differ in prize values or show-up fees, the structural parameters of the two contests cannot be too far apart (see, e.g., Example 1). While we do not repeat this condition at the beginning of each formal result, it should be understood that, from hereon, both contests are assumed to be competitive in equilibrium. Notice, however, that Proposition 2 applies whenever this assumption fails. Hence, we do provide a full characterization of selection behavior for all parameter values.

When both contests are competitive, $-\infty < \theta_1^*, \theta_2^* < \infty$. Hence, the payoff difference $\pi_1^* - \pi_2^*$ is

$$\pi_1^*(a, \theta_1^*) - \pi_2^*(a, \theta_2^*) = w_1 - w_2 + v_1 \bar{F}\left(\frac{\theta_1^* - a}{\sigma_1}\right) - v_2 \bar{F}\left(\frac{\theta_2^* - a}{\sigma_2}\right). \tag{3}$$

The following lemma characterizes selection in the tails.

Lemma 2 *In the tails of the ability distribution, selection is as follows:*

$$\lim_{a \to -\infty} \Pr 1(a) = \Gamma\left(\frac{w_1 - w_2}{\rho}\right)$$

$$\lim_{a \to \infty} \Pr 1(a) = \Gamma\left(\frac{w_1 + v_1 - (w_2 + v_2)}{\rho}\right)$$

The proof of Lemma 2, which is omitted, follows immediately from equation (3) and the fact that $\Pr 1(a) = \Gamma\left[\frac{\pi_1^*(a,\theta_1^*)-\pi_2^*(a,\theta_2^*)}{\rho}\right]$. The intuition is straight forward. In anticipation of losing in either contest, show-up fees are the sole pecuniary consideration for very low types. By contrast, in anticipation of winning, very high types only consider the sum of show-up fees and prize values. They pecuniarily prefer whichever contest offers the higher total.

The derivative $d(\pi_1^* - \pi_2^*)/da$ can be written as

$$\frac{d(\pi_1^* - \pi_2^*)}{da} = \frac{v_1}{\sigma_1} f\left(\frac{\theta_2^* - a}{\sigma_2}\right)\left[\lambda(a) - \frac{v_2/v_1}{\sigma_2/\sigma_1}\right]. \tag{4}$$

Recall that $\lambda(a) = f\left(\frac{\theta_1^*-a}{\sigma_1}\right)/f\left(\frac{\theta_2^*-a}{\sigma_2}\right)$ is the likelihood ratio of agent a just meeting the standard in each contest. Equation (4) reveals that the shape of $\pi_1^* - \pi_2^*$, and hence $\Pr 1(a)$, crucially depends on the properties of $\lambda(a)$.

Let $\underline{\lambda} \equiv \inf_{a\in\mathbb{R}} \lambda(a)$ and $\overline{\lambda} \equiv \sup_{a\in\mathbb{R}} \lambda(a)$. The following lemma shows that, for $\sigma_1 = \sigma_2$, $\lambda(a)$ is strictly monotone in ability and takes on values on either side of 1. This result relies on the log-concavity of f. For $\sigma_1 < \sigma_2$, log-concavity implies that $\lambda(a)$ goes to zero in the tails. It does not pin down the number of interior extrema, however. This is where Condition 1 comes into play. Condition 1 guarantees that $\lambda(a)$ is single-peaked.

Lemma 3 *Properties of $\lambda(a)$:*

1. *If $\sigma_1 = \sigma_2$ then: a) $\lambda'(a) \gtreqless 0$ iff $\theta_1 \gtreqless \theta_2$; b) for $\theta_1 \neq \theta_2$, $\underline{\lambda} < 1 < \overline{\lambda}$.*
2. *If $\sigma_1 < \sigma_2$ then: a) $\lim_{|a|\to\infty} \lambda(a) = 0 = \underline{\lambda}$; b) for $\theta_1 \neq \theta_2$, $\bar{\lambda} > \lambda(\theta_1) > 1$; c) if f satisfies Condition 1, then $\lambda(a)$ is single-peaked in $a \in (-\infty, \infty)$.*

Together, Lemmas 2 and 3 and equation (4) allow us to characterize selection behavior as a function of ability. Assume that Condition 1 holds and that $\sigma_1 < \sigma_2$. Other structural parameters can be arbitrary. From equation (4) and Lemma 3 it readily follows that $\Pr 1(a) = \Gamma\left(\frac{\pi_1^*-\pi_2^*}{\rho}\right)$ either takes on two extrema, first reaching a minimum and then a maximum, or is strictly decreasing in a. Selection in the tails is described by Lemma 2: show-up fees alone determine selection of very low types, while the sum of show-up fees and prize values determine the selection of very high types. The following proposition summarizes these observations.

Proposition 4 *Let $\sigma_1 < \sigma_2$ while the other structural parameters are arbitrary. Suppose Condition 1 holds. Then* $\Pr 1(a)$ *either takes on two extrema, first a minimum and then a maximum, or is strictly decreasing. Specifically:*

1. *Provided* $\frac{v_2/v_1}{\sigma_2/\sigma_1} < \overline{\lambda}$, $\Pr 1(a)$ *has two extrema; a minimum (maximum) at the smaller (larger) value of 'a' solving* $\lambda(a) = \frac{v_2/v_1}{\sigma_2/\sigma_1}$. *Otherwise,* $\Pr 1(a)$ *is strictly decreasing.*

2. $\lim_{a \to -\infty} \Pr 1(a) = \Gamma\left(\frac{w_1 - w_2}{\rho}\right)$ *and* $\lim_{a \to \infty} \Pr 1(a) = \Gamma\left(\frac{w_1 + v_1 - (w_2 + v_2)}{\rho}\right)$.

Proposition 4 assumes that Condition 1 holds and limits attention to the generic case where $\sigma_1 \neq \sigma_2$. When Condition 1 fails, $\lambda(a)$ may take on multiple interior extrema. In turn, this implies that $\Pr 1(a)$ may take on up to twice the number of extrema of $\lambda(a)$. The knife-edge case where contests differ in multiple dimensions but are equally meritocratic is relegated to the appendix. (See Proposition 12 in Appendix A.)

At the outset we presented Example 1. It illustrated that, despite the relative simplicity of our model, selection behavior in the ponds dilemma can be quite complex. Proposition 4 establishes that the "bimodal" selection patterns of Cases 1 and 2 are, in fact, generic. Case 3's selection profile for $\rho \to 0$ also generalizes: indeed, it is easily verified that the ability space $\mathbb{R}$ can be partitioned into at most four intervals such that, in the limit, types belonging to the same interval enter the same contest, while types belonging to adjacent intervals enter different contests.

4.2.3 Isolating Selection Effects

While Proposition 4 describes agents' selection behavior, in terms of intuition it has little to offer. To better understand why agents behave the way they do, we now study the selection effects of each structural parameter in isolation. That is, we analyze selection across contests that are identical in all respects save one.

Remark 1 and Figure 2 summarize our findings. Remark 1 is a corollary of a sequence of more detailed propositions that have been relegated to the appendix. (See Propositions 9 to 13 in Appendix A.)

Remark 1 *Suppose contests are identical in all dimensions save one. Then selection is as follows:*

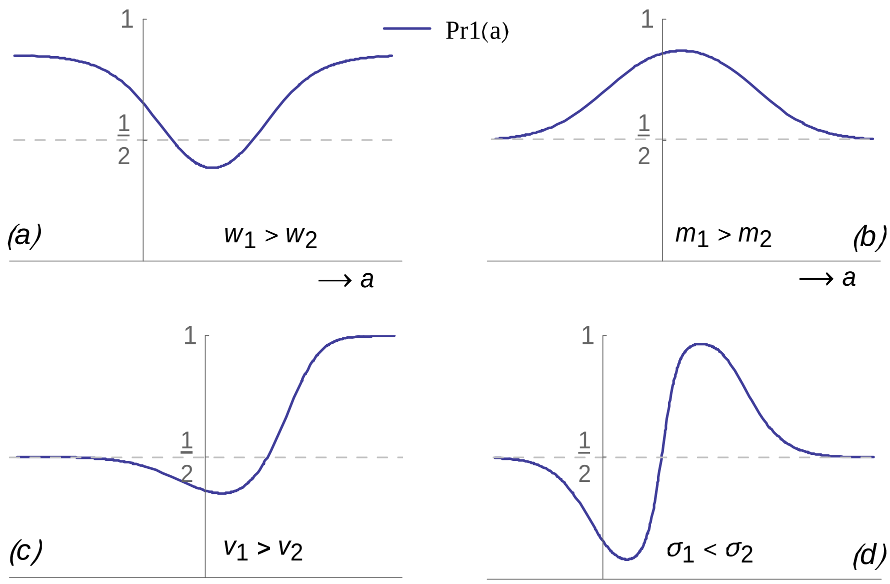

Figure 2: Probability of entering Contest 1 when the contests differ in one dimension only.

1. *w: Higher show-up fees attract the best and the worst, while repelling the middle.*

2. *m: More prizes attract middling abilities, while not affecting selection of the best and the worst.*

3. *v: Higher prizes attract the best, while not affecting selection of the worst. Middling abilities tend to be repelled.*

4. *σ: Greater meritocracy attracts high types and repels low types. However, these selection effects dissipate toward the tails.*

One by one we now discuss the intuitions behind these selection effects.

Show-Up Fees

Suppose that the two contests are identical except for their show-up fees. From Proposition 2, above, we already know that Contest 1 benefits from positive selection when w_1 is so much larger than w_2 that Contest 2 is uncompetitive. Selection is more nuanced in the competitive case. To see why, start from a situation where the two contests are identical and suppose that Contest 1 raises its show-up fee. This makes Contest 1 more attractive to agents of all abilities, who now enter in larger numbers. For both markets to clear, Contest 1's equilibrium standard must rise and Contest 2's must fall. Since standards have risen in tandem with show-up fees, agents now face a clear trade-off: the higher show-up fee in Contest 1 (the big pond) must be weighed against the lower standard—and hence better chance of winning—in Contest 2 (the small pond).

A common intuition for the ponds dilemma is that only the most able should enter the big pond: "if you can't stand the heat, stay out of the kitchen!" While it is true that the most able suffer little from heightened competition, so do the least able, however. For both types, a difference in standards is of little import because only extremely unlikely realizations of ε_i affect their almost pre-ordained success or failure. Hence, agents of extreme ability (both high *and* low) tend to opt for the contest with the higher show-up fee—i.e., the big pond. Not so for agents of intermediate ability, whose chances of success are noticeably hurt by a higher standard. They tend to opt for the contest with the lower show-up fee—i.e., the small pond. As a result, $\Pr 1\,(a)$ is U-shaped. This is illustrated in Figure 2(a), which depicts the propensity to enter the big pond as a function of ability.

Because of its higher standard and greater show-up fee, we have referred to Contest 1 as the big pond. However, the following example shows that the "big pond" may in fact be smaller than the "small pond." That is, despite its higher show-up fee, Contest 1 may attract only a minority of agents. The driving factor is the mass of middling sorts, who are repelled by the big pond's higher standard. The identity and size of this group depends on the number of prizes on offer and the shape of the ability distribution in the population as a whole. The upshot is that a contest may well raise its show-up fee, only to see participation decline.

Example 2 *Let* $a \sim N(0,1)$, $\delta \sim N(0,.05)$, *and* $\varepsilon_i \sim \text{Logistic}(0,1)$. *If* $w_1 = 1.1 > 1 = w_2$, $m_i = .4$ *and* $v_i = 4$, $i \in \{1,2\}$, *then* $\Pr 1 = .44 < .56 = \Pr 2$.

Number of Prizes

While it is easy to see that offering more prizes increases entry, the selection effects are less clear. Who are these new entrants? Because contestants do not care about the number of prizes *per se*, offering more prizes only has an indirect effect, namely, a reduction in the performance standard. This is valuable regardless of ability, though more so for intermediate types, whose chances of winning improve the most. Hence, an increase in the number of prizes, m_i, unambiguously raises entry of *all* ability types into Contest i, but especially of middling sorts.

More formally, suppose that $m_1 > m_2$ while the contests are otherwise identical. It is easily verified that $\lim_{|a|\to\infty} \Pr 1(a) = \Gamma(0) = 1/2$. This reflects that agents of extreme ability do not care about standards. The derivative $d(\pi_1^* - \pi_2^*)/da$ is the same as in the show-up fee case. However, because the order of standards is reversed (i.e., $\theta_1^* < \theta_2^*$), $\pi_1^* - \pi_2^*$ and $\Pr 1(a)$ are *inverse*-U-shaped rather than U-shaped. The resulting selection pattern is depicted in Figure 2(b).

Owing to its higher standard, one might consider Contest 2 to be the big pond. Notice, however, that it does not offer higher rewards as compensation. Therefore, the "big pond" repels all abilities, but to differing degrees, depending on their relative chances of success in the two contests.

Value of Prizes

The canonical ponds dilemma arises when contests differ in prize values. Naturally, higher prizes lead to an inflow of contestants and, hence, to a higher performance standard. Thus, as was the case for show-up fees, contestants face a trade-off between payoffs and standards. However, in this case, both the costs and the benefits of entering the high-prize contest are ability-dependent. While show-up fees are equally valuable to all, the expected benefit of a higher prize is proportional to an agent's probability of winning. Therefore, all else equal, a higher prize in Contest 1 makes $\Pr 1(a)$ strictly increasing in a. The cost of a higher standard continues to be greatest for intermediate types. Together, the two effects make $\Pr 1(a)$ U-shaped, with the right asymptote exceeding the left, i.e., $\lim_{a\to\infty} \Pr 1(a) > \lim_{a\to-\infty} \Pr 1(a)$.[7] The resulting selection pattern is depicted in Figure 2(c).

As always, those of extreme ability are unaffected by the difference in standards. Yet, entry decisions differ markedly between the top and the bottom. For those at the bottom, prize differences are irrelevant because prizes are unattainable. Therefore, they perceive the two contests as equally attractive, leading to a 50-50 split. Those at the top are virtually guaranteed to win a prize in either contest. Therefore, they are much more likely to opt for high-prize Contest 1, i.e., the big pond. Still, in general, selection into the big pond fails to be monotone.[8] As with show-up fees, the ratio of success probabilities favors the small pond for middling sorts, and this consideration tends to dominate their entry decisions. The key

[7] To be precise: when v_2/v_1 is sufficiently lopsided *and* $\lambda(a)$ is bounded, it may happen that $\Pr 1(a)$ is strictly increasing rather than U-shaped. See Proposition 11 in Appendix A for details.

[8] See Proposition 11 in Appendix A for details.

insight is that, regardless of whether the riches in the big pond come in the form of contingent prizes or non-contingent show-up fees, selection tends to non-monotonic in ability.

Meritocracy

We now examine how meritocracy drives selection. For risk-neutral agents, noise in performance evaluation might seem irrelevant, because it does not affect expected performance. The flaw in this reasoning is that measurement errors have asymmetric effects, which depend on an agent's ability relative to the standard. When an agent's ability falls below the standard, he can only succeed if he gets a "lucky break," i.e., a positive realization of ε_i. When his ability exceeds the standard, he can only fail if he suffers an "unlucky break," i.e., a negative realization of ε_i. In the former situation the agent seeks out noisy measurement, since therein lies his only path to success. In the latter, he avoids noisy measurement, since it constitutes his only possible undoing.

Even in this case, selection is not monotone, however. To see why, recall that individuals of extreme ability—both high and low—are essentially unaffected by measurement noise, since only extremely unlikely realizations of ε_i can alter their almost pre-ordained success or failure. As the contests are identical in all other respects, these types enter each contest with almost equal probability. Hence, meritocracy does produce favorable selection, but with waning power in the tails. This is illustrated in Figure 2(d).[9]

As it is more harshly revealing of true performance, we may consider the more meritocratic contest to be the big pond. Still, rewards are the same in both contests, while standards cannot be ranked. To see why either contest can have the higher standard, notice that most agents need a lucky break when prizes are scarce. This induces the bulk of the population to opt for the noisy contest and, as a result, the less meritocratic contest has the higher standard. On the other hand, when prizes are plentiful, most agents need to avoid an unlucky break. This induces them to opt for the more meritocratic contest, resulting in the opposite ranking of standards.

Deciding which contest to enter is most complicated for types who need a lucky break in one contest, but need to avoid an unlucky break in the other contest, i.e., types $a \in [\min\{\theta_1^*, \theta_2^*\}, \max\{\theta_1^*, \theta_2^*\}]$. Their predicament blunts payoff differences and makes selection

[9] In the figure we have assumed that Condition 1 holds, such that $\lambda(a)$ is single-peaked. Otherwise, $\Pr 1(a)$ may exhibit multiple extrema on either side of $\tilde{a}$, the point where $\Pr 1(a)$ single-crosses $1/2$. (See Proposition 13 in Appendix A.)

less pronounced. To see why, suppose the more meritocratic contest also has the higher standard. In that case, the agent needs a lucky break in the more meritocratic contest, while he needs to avoid an unlucky break in the less meritocratic contest. Since neither contest is very likely to produce the desired result, there is little to distinguish between them. Alternatively, when the more meritocratic contest has the lower standard, the agent needs a lucky break in the less meritocratic contest, while he needs to avoid an unlucky break in the more meritocratic contest. Since both contests are quite likely to produce the desired result, again, there is little to distinguish between them. Thus, selection is weak in this region.

In Proposition 13 in Appendix A, we prove that the more meritocratic contest is exclusive. That is, provided pecuniary motives dominate, the more meritocratic contest attracts only a minority of the population. The intuition is as follows. As illustrated in Figure 2(d), agents are more likely to enter the more meritocratic contest than the less meritocratic contest iff their ability exceeds some threshold, $\tilde{a}$. The threshold is such that an agent of ability $\tilde{a}$ is equally likely to win in either contest. Within each contest, the probability of winning is strictly increasing in ability. Hence, agents who choose to enter the more meritocratic contest (tend to) have a better chance of winning than those entering the less meritocratic contest. This implies that, for a given mass of entrants, the more meritocratic contest produces more winners than the less meritocratic contest. Because the number of prizes is the same in both contests, the more meritocratic contest must attract fewer entrants. In other words, it is exclusive.

Together, exclusivity and dissipation of selection power in the tails—i.e., $\lim_{a\to-\infty} \Pr 1(a) = 1/2$—imply that very low types disproportionately enter the *more* meritocratic contest. Perversely, this means that the average ability in the more meritocratic contest may be lower than in the less meritocratic contest.[10]

Combining the uncovered selection effects provides an intuitive understanding of the behavior in Proposition 4: Show-up fees alone determine selection of very low types, while the sum of show-up fees and prize values determine the selection of very high types. Meritocracy shapes behavior in between these extremes, producing two interior extrema. Finally, the number of prizes affects selection only indirectly, through its effect on standards.

[10] Let $a \sim N(0,1)$, $\delta \sim N(0,.05)$, and $\varepsilon_i \sim \mathit{Logistic}(0,\sigma_i)$. If $w_i = 1$, $m_i = .01$, $v_i = 1$, $i \in \{1,2\}$, and $(\sigma_1,\sigma_2) = (.3,1)$, then $E_1[a] = -0.023 < 0.017 = E_2[a]$.

4.2.4 Stochastic Ordering of Abilities and Limit Behavior for $\rho \to 0$

Finally, we consider the implications of selection for the distribution of abilities across contests and study selection behavior when non-pecuniary preferences vanish, i.e., $\rho \to 0$.

From Proposition 2 we already know that, in the uncompetitive case, abilities in the competitive contest FOSD abilities in the uncompetitive contest. Hence, in line with common intuition, the "big pond" attracts the best-and-the-brightest, while the "small pond" is the refuge of low ability types.

For the competitive case with $w_1 > w_2$, we have seen that $\Pr 1(a)$ is U-shaped. This suggests that abilities in Contest 1 are more dispersed than in Contest 2. Similarly, the inverse-U-shapedness of $\Pr 1(a)$ when $m_1 > m_2$ suggests that, in this case, abilities in Contest 2 are more dispersed than in Contest 1. Notice, however, that the contests' ability distributions cannot be ranked by second-order stochastic dominance (SOSD), since neither case allows for a consistent ranking of average abilities. To remedy the situation, we introduce the concept of *single-crossing dispersion* (Ganuza and Penalva, 2006).

Definition 1 *A RV with CDF $J_1(a)$ is more single-crossing (SC) dispersed than a RV with CDF $J_2(a)$ iff there exists a unique $a' \in \mathbb{R}$ such that $J_1(a') = J_2(a')$ and, $\forall a \overset{(>)}{<} a'$, $J_1(a) \overset{(<)}{>} J_2(a)$.*

For example, consider two RVs drawn from Normal distributions with different means and different variances. It may be verified that, regardless of the means, the distribution with the higher variance is more SC dispersed than the distribution with the lower variance.

In the next proposition we show that abilities in the two contests can indeed be ranked according SC dispersion.[11]

Proposition 5 *Let $w_1 > w_2$ while the contests are otherwise identical. Provided ρ is not too large, abilities in Contest 1 are more SC dispersed than in Contest 2.*

For $m_1 > m_2$ and ρ not too large, abilities in Contest 2 are more SC dispersed than in Contest 1.

[11] For arbitrary sets of probability distributions, the concept of SC dispersion has the serious drawback that it may violate transitivity. This problem does not arise in our setting, however. To see this, fix a set of contests whose structural parameters are identical save for their show-up fees or number of prizes. Proposition 5 implies that the set can be completely ordered on the basis of SC dispersion ($\geq_{SC}$). Specifically, under endogenous sorting between contests (i, j), $G_i(\cdot) \geq_{SC} G_j(\cdot)$ iff $w_i \geq w_j$ or $m_i \leq m_j$, respectively.

By attracting extreme types and repelling middling sorts, the contest with the higher show-up fee or fewer prizes attracts the more diverse talent pool. For the case of show-up fees, this phenomenon is most cleanly captured in the limit when non-pecuniary considerations vanish. Such an analysis is also of independent interest, since non-pecuniary payoffs are mostly absent from the extant literature. When $\rho \to 0$, we find that selection becomes deterministic, leaving a "hole" in the ability distribution of the contest offering the higher fee. That is, for $w_1 > w_2$, extreme types enter Contest 1 while middling sorts enter Contest 2.

By contrast, when contests differ in the number of prizes, selection remains strictly stochastic even in the limit. Moreover, $\lim_{\rho \to 0} \Pr 1(a)$ remains inverse-U-shaped. This result is driven by a no-arbitrage condition which implies that, in the limit, the contests' standards must be the same. Otherwise, all agents would choose the contest with the lower standard, which is inconsistent with it having the lower standard in the first place. When both contests have the same standard, there is no particular reason for agents with the same ability to enter the same contest, even when $\rho \to 0$. However, equal standards and $m_1 > m_2$ do imply that more contestants enter Contest 1 than Contest 2. For $\rho \to 0$, a particular "mixed" entry pattern is selected among a continuum of patterns consistent with the requirements of equal standards and market clearing. Formal statements and proofs of these results have been relegated to the appendix. (See Propositions 14 and 15 in Appendix A.)

Next we turn to differences in prize values. Agents are more likely to enter the high-prize contest than the low-prize contest iff their ability exceeds some threshold, $\hat{a}$. Hence, it would seem that high types are overrepresented in the high-prize contest and low types in the low-prize contest. However, this ignores the base rate of selection into the two contests. Relative to its population share, a type is overrepresented in Contest i iff its propensity to enter, $\Pr i(a)$, is greater than the average propensity, $\Pr i$. Therefore, if $v_1 > v_2$, high types are indeed overrepresented in high-prize Contest 1. However, if $\Pr 1 < 1/2$, so are very low types, since they enter both contests with equal probability. The potential overrepresentation of low types in the high-prize contest implies that ability distributions in the two contests cannot be ranked by FOSD. In fact, average ability in the high-prize contest may well be lower than in the low-prize contest, as the following example shows.

Example 3 *Suppose* $a \sim N(0,1)$, $\varepsilon_i \sim Logistic(0,.3)$, $\delta \sim N(0,.05)$. *Let* $v_1 = 1.1 > 1 = v_2$, $w_i = 1$, *and* $m_i = .05$, $i \in \{1,2\}$. *Then* $E_1[a] = -.023 < .020 = E_2[a]$. *Hence, the* *high-prize contest attracts individuals of lower average ability.*

Nonetheless, for small ρ, it is still true that a random individual with ability greater than $\hat{a}$ is much more likely to enter the high-prize contest, while a random individual with ability smaller than $\hat{a}$ is much more likely to enter the low-prize contest. One formalization of this idea is to compare ability quantiles across contests. For example, we can ask how the ability of the (lowest) 1st percentile in the high-v contest compares to the ability of the (highest) 99th percentile in the low-v contest. As we show, for small ρ, the former exceeds the latter with probability one. In fact, Proposition 6 generalizes this idea to arbitrary quantiles.

Proposition 6 *Let $v_1 > v_2$ while the contests are otherwise identical. For any $0 < p_1, p_2 < 1$, there exists a $\bar{\rho} > 0$ such that for all $0 \leq \rho < \bar{\rho}$ the following holds: with probability 1, an agent at the p_1-th ability-quantile in Contest 1 has strictly greater ability than an agent at the p_2-th ability-quantile in Contest 2.*

When contests differ in terms of meritocracy, $\Pr 1(a)$ single-crosses $\frac{1}{2}$ from below. Hence, we can once again rank arbitrary ability quantiles across contests for ρ sufficiently small. We omit a formal statement and proof of this result, since it is analogous to Proposition 6.

5 Endogenous Effort

In the model of Section 2, "effort" was exogenous and equal to ability. Yet, in practice, agents' effort levels may vary with the structural parameters of the contest. Moreover, anticipated effort may play a role in deciding which contest to enter. Therefore, we now add endogenous effort back into the model.

As before, a unit mass of agents choose between two contests, 1 and 2. The determination of success and failure in each contest is analogous to the earlier model. However, measured performance now depends on endogenous effort rather than exogenous ability. Specifically, the performance of an agent who exerts effort $X \in [0, \infty)$ in Contest i is $Y_i = X \cdot E_i$. Here, $E_i \in (0, \infty)$ represents noise in performance measurement. Taking logs we get

$$y_i = x + \varepsilon_i \ .$$

Our assumptions on ε_i are the same as before.

For an agent of ability $a \in \mathbb{R}$, the cost of exerting (log of) effort $x \in [-\infty, \infty)$ is given by $c(x, a)$. In addition to continuity and differentiability (twice), we impose the following, fairly standard properties on the cost function: for all $a \in \mathbb{R}$: 1) $c(-\infty, a) = 0$; 2) $\left.\frac{\partial c(x,a)}{\partial x}\right|_{x=-\infty} = 0$ and $\frac{\partial c(x,a)}{\partial x} > 0$ for $x \in \mathbb{R}$; 3) $\frac{\partial^2 c(x,a)}{(\partial x)^2}$ is strictly positive and bounded away from zero; 4) $\frac{\partial c(x,a)}{\partial a} < 0$ and $\frac{\partial^2 c(x,a)}{\partial a \partial x} < 0$ for $x \in \mathbb{R}$.

An agent of ability a who exerts effort x in Contest i with standard θ_i enjoys an expected pecuniary payoff

$$\pi_i(x, a, \theta_i) = w_i + v_i \bar{F}\left(\frac{\theta_i - x}{\sigma_i}\right) - c(x, a) \ . \tag{5}$$

Non-pecuniary payoffs, δ, are the same as before, and total payoffs continue to be the sum of pecuniary and non-pecuniary payoffs.

Agents simultaneously and independently choose which contest to enter and how much effort to exert.[12] If fewer than m_i enter Contest i, then $\theta_i = -\infty = x_i(a, -\infty)$ and all entrants receive a prize v_i. Otherwise, for a given CMF $H_i(a)$, effort schedule $x_i(a, \theta_i)$ and performance standard θ_i constitute an equilibrium of Contest i if: 1) $x_i(a, \theta_i)$ is optimal for every $a \in \mathbb{R}$; 2) θ_i is such that the mass of winners, W_i, equals the mass of prizes, m_i. Hence, an equilibrium $(x_i^*(a, \theta_i^*), \theta_i^*)$ of Contest i satisfies

$$\begin{aligned} x_i^*(a, \theta_i^*) &\in \arg\sup_x \pi_i(x, a, \theta_i^*), \text{ and} \\ W_i(\theta_i^*) &= \int_{-\infty}^{\infty} \bar{F}\left[\frac{\theta_i^* - x_i^*(a, \theta_i^*)}{\sigma_i}\right] dH_i(a) = m_i \ . \end{aligned}$$

A Bayesian Nash equilibrium of the full game consists of a tuple $\{(H_1^*(a), H_2^*(a)), (x_1^*(a, \theta_1^*), \theta_1^*), (x_2^*(a, \theta_2^*), \theta_2^*)\}$ of CMFs $H_i^*(a)$ and equilibria $(x_i^*(a, \theta_i), \theta_i^*)$, $i \in \{1, 2\}$, such that if H_i^* assigns positive mass density to type a in Contest i, then this type cannot gain by switching contests.

5.1 Equilibrium

We solve for equilibrium as before, save for the additional consideration of effort optimization. First, we characterize the optimal-effort schedule $x_i(a, \theta_i)$ conditional on standard θ_i. Second, for each contest we determine the market-clearing standard θ_i conditional on

[12] Again, the analysis remains unchanged if agents move sequentially or if they can switch contests and adjust their effort upon observing others' entry and effort choices. As before, the argument relies on the atomicity of individuals and the absence of aggregate uncertainty.

CMF H_i. Third, we derive agents' entry decisions and resulting CMFs (H_1, H_2) conditional on standards (θ_1, θ_2). Together, these three steps define a mapping from the space of performance standards into itself. Finally, we show that there exist standards (θ_1^*, θ_2^*) that constitute a fixed point of the system. These standards gives rise to an equilibrium $\{(H_1^*(a), H_2^*(a)), (x_1^*(a, \theta_1^*), \theta_1^*), (x_2^*(a, \theta_2^*), \theta_2^*)\}$.

We begin by characterizing the optimal effort profile, $x^*(a, \theta)$, conditional on standard θ. (We suppress subscript i in the remainder of this section because it plays no role.) Differentiating equation (5) with respect to x yields the following first-order condition (FOC) for optimal effort:

$$\frac{v}{\sigma} f\left(\frac{\theta - x}{\sigma}\right) - \frac{\partial c(x, a)}{\partial x} = 0 \ .$$

The second-order condition (SOC) for the FOC to characterize a maximum is

$$f'\left(\frac{\theta - x}{\sigma}\right) / f\left(\frac{\theta - x}{\sigma}\right) + \sigma \frac{\partial^2 c(x, a)}{(\partial x)^2} / \frac{\partial c(x, a)}{\partial x} > 0$$

Notice that the SOC is always satisfied when $x \geq \theta$. When $x < \theta$, performance measurement must be sufficiently noisy or the log of marginal costs must increase sufficiently fast. For the remainder of the analysis we assume that the SOC is satisfied.

Because the marginal cost of effort is strictly decreasing in a, the optimal-effort schedule is strictly increasing. All else equal, effort is increasing in v as well, because a higher prize raises the marginal benefit of effort. The effect of an exogenous rise in standards critically depends on whether an agent needs a lucky break or needs to avoid an unlucky break. When the agent needs to avoid an unlucky break, increasing the standard raises his effort. To see this, notice that a higher θ narrows the "gap" $|x - \theta|$. Since the density of ε is single-peaked around zero, this narrowing raises the marginal benefit of effort. Hence, optimal effort increases. By contrast, when an agent needs a lucky break, a higher standard widens the gap between effort and standard. Again owing to the single-peakedness of f, the marginal benefit of effort falls, and so does optimal effort. Interestingly, effort is not uniformly increasing in meritocracy either. To see why, notice that a fall in σ lifts the peak of f and thins the tails. This raises the marginal benefit of effort for agents operating close to the standard but reduces it for those operating farther away. Naturally, optimal effort follows suit. Put differently, a rise in meritocracy discourages low types, encourages intermediate types, and makes high types complacent.

Returning to the two-contest environment, the remainder of the equilibrium derivation proceeds along the same lines as in the exogenous-effort model. This yields:

Proposition 7 *An equilibrium exists in the selection model with endogenous effort.*

In a symmetric baseline, the equilibrium is unique. Both contests are competitive and have the same standard. 50% of every ability type enter each contest.

When $\rho \to 0$, both contests remain competitive in a neighborhood of a symmetric baseline.

5.2 Selection Around a Symmetric Baseline

We now revisit the sorting effects of cross-contest differences in structural parameters. For tractability reasons, we focus on a neighborhood of structural parameters around a symmetric baseline. That is, the two contests cannot be "too different." In that case, all our previous findings carry over. Formally,

Proposition 8 *In a neighborhood of structural parameters around a symmetric baseline, selection patterns in the endogenous-effort model are the same as in the exogenous-effort model of Section 2. That is, mutatis mutandis, Propositions 1 to 4 continue to hold.*

The proof of Proposition 8 relies on the envelope theorem. In the model of Section 2, "effort" was fixed at a regardless of the values of structural parameters. By contrast, effort is now parameter-dependent. Indeed, even a marginal change in one or more parameters has a first-order effect on optimal effort. However, by the envelope theorem, this effort adjustment only has a second-order effect on payoffs. Hence, when studying $\Pr 1(a)$ in a neighborhood of a symmetric baseline, we can ignore changes in $x_i^*(a)$ and pretend that effort is *exogenous*. In fact, this argument holds around any parameter point—not merely around a symmetric baseline. However, in a symmetric baseline, every type's effort level is the same across contests. Therefore, the cost of effort differences out of $\pi_1^* - \pi_2^*$, which makes the endogenous-effort model locally isomorphic to the exogenous-effort model. This allows us to reinterpret an individual's effort at a symmetric baseline as his type and apply all the arguments and machinery of the exogenous-effort model.

Proposition 8 raises the question how "close" the two contests have to be for the selection pattern in the endogenous-effort model to be similar to that in the exogenous-effort model. To get a feel, we re-analyze Example 1 with endogenous effort.

Example 4 *Suppose* $a \sim N(0,1)$, $\delta \sim N(\tau = .05, \rho = .05)$, *and* $\varepsilon_i \sim Logistic(0, \sigma_i)$, $i \in \{1,2\}$. *Let* $(w_1, w_2) = (1.1, 1)$, $(m_1, m_2) = (.1, .2)$, $(v_1, v_2) = (1, 1.1)$, $(\sigma_1, \sigma_2) = (.6, 1)$, *and* $c(x,a) = \left(e^{e^x} - e^x - 1\right) / \left(e^{e^a} - 1\right)$. *Hence, probability distributions and parameter values are as in Example 1, while the cost function is chosen such that it satisfies our assumptions. Specifically, for* σ *not too small, the SOC is satisfied.*[13]

1. *The resulting effort schedules are shown in Figure 3(a), while* $\Pr 1(a)$ *is shown in Figure 3(b). The PDFs of abilities in the two contests are given in Figure 3(c). Standards are* $(\theta_1^*, \theta_2^*) = (.40, .36)$. *The fraction of the population entering each contest is* $(\Pr 1, \Pr 2) = (.25, .75)$, *while average abilities are* $(E_1[a], E_2[a]) = (.32, -.11)$.

2. *When* ρ *is reduced to* $.0005$, *effort, selection, and ability densities are as in Figures 1(d), (e), and (f), respectively. Standards are* $(\theta_1^*, \theta_2^*) = (.43, .42)$, *while* $(\Pr 1, \Pr 2) = (.20, .80)$ *and* $(E_1[a], E_2[a]) = (.49, -.13)$.

The similarities between Figures 1 and 3 are quite striking. Indeed, the selection function and the resulting endogenous ability distributions are almost indistinguishable.[14] Further simulations with different sets of parameters suggest that, even when the visual likeness with the exogenous-effort model is lost, $\Pr 1(a)$'s characteristic "bimodal" shape is maintained. This provides some comfort that the results and intuitions of the exogenous-effort model are reasonably robust and, for most intents and purposes, carry over to the endogenous-effort model.

5.3 Effort Across Contests

Agents of the same ability who enter the same contest exert the same effort. However, across contests, effort levels will generally differ, i.e., $x_1^*(a, \theta_1^*) \neq x_2^*(a, \theta_2^*)$. This raises the question

[13] Our cost function is rather complicated and ugly. A simpler cost function, such as $c(x,a) = e^{\beta x}/e^a$, $\alpha > 0, \beta > 1$, would not materially affect our findings. However, it would complicate the exposition. The reason is that, for sufficiently low ability types, $c(x,a) = e^{\beta x}/e^a$ and other functions like it violate the SOC. (Alternatively, the FOC may not even have a solution.) As a result, $x_i(a)$ becomes discontinuous. Specifically, below a certain ability, contestants "drop out" and choose $x = -\infty$. This discontinuity is more or less immaterial, however, because all the relevant integrals such as W_i and H_i remain continuous. Therefore, equilibrium continues to exist, and the discontinuity of $x_i(a)$ has only minor effects on standards and selection. Details are available from the authors upon request.

[14] It is easy to destroy any visual likeness by changing the cost function's dependence on a. For example, if $c(x,a) = \left(e^{e^x} - e^x - 1\right)/a$, then the horizontal axes in Figure 4 are stretched out by a factor $\exp(\exp(a))$. However, provided multiplicative separability between effort and ability is maintained, any change in the cost function's dependence on a simply corresponds to a relabeling of types. In that sense, our initial choice was without loss of generality.

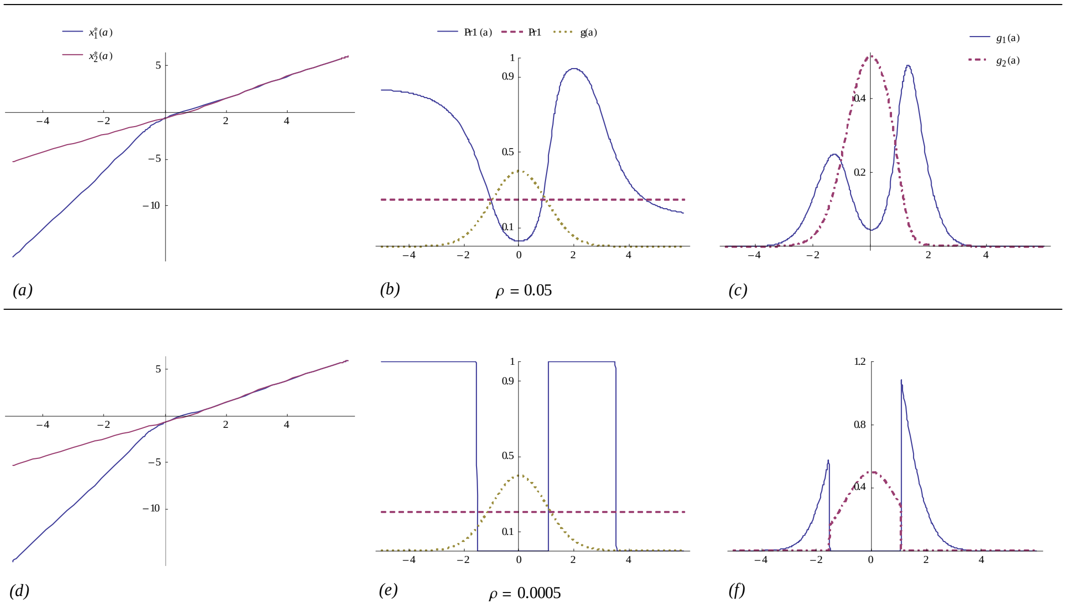

Figure 3: Optimal effort, selection, and the resulting ability distributions in the two contests of Example 4, for $\rho = .05$ (top) and $\rho = .0005$ (bottom).

which contest induces agents to work harder and how the prospect of hard work is related to entry. Intuitively, the relationship between effort and entry could go either way. Since effort is costly, one might argue that agents should avoid contests that make them work hard. Conversely, if agents work hard, the rewards must be great, which should be attractive.

Our next lemma shows that effort and entry are, in fact, unrelated. Rather, it is the *change* in the probability of entry in a neighborhood around an agent—i.e., the slope of $\Pr 1(a)$—that determines which contest elicits the greater effort. Specifically,

Lemma 4

$$x_1^*(a) \gtreqqless x_2^*(a) \Longleftrightarrow \frac{d \Pr 1(a)}{da} \gtreqqless 0 \ .$$

Notice that Lemma 4 holds globally, i.e., not just in a neighborhood of a symmetric baseline. The intuition for the lemma is as follows. The envelope theorem implies that the rise in payoffs associated with a small rise in ability is equal to the cost savings from exerting the same level of effort at higher ability. Due to the sub-modularity of costs in x and a, i.e., $\partial^2 c(x,a) / (\partial a \partial x) < 0$, these cost savings are increasing in the initial level of effort. Hence, when $\pi_1^*(a)$ increases faster in a than $\pi_2^*(a)$, it must be that $x_1^*(a) > x_2^*(a)$, and vice versa.

Finally, recall that $d\Pr 1(a)/da$ takes on the same sign as $d(\pi_1^* - \pi_2^*)/da$. The sign of the slope of $\Pr 1(a)$ is therefore a sufficient statistic for ranking an agent's effort across contests.

From Propositions 4 and 8 we know that, in a neighborhood of a symmetric baseline, the slope of $\Pr 1(a)$ is determined by meritocracy alone. In combination with Lemma 4 this allows us to rank agents' efforts across contests. Let a_0 denote the lower, and a_1 the higher value of a where $\lambda(a) = \frac{v_2/v_1}{\sigma_2/\sigma_1}$.[15] Then:

Corollary 2 *Agents of intermediate ability work harder in the more meritocratic contest, while agents of extreme ability work harder in the less meritocratic contest.*

Formally, suppose Condition 1 holds and $\sigma_1 < \sigma_2$. In a neighborhood of a symmetric baseline,

$$x_1^*(a) > x_2^*(a) \text{ iff } a_0 < a < a_1 .$$

In Section 5.1 we observed that, in a single-contest, an increase in meritocracy discourages low types, encourages intermediate types, and makes high types complacent. Corollary 2 can be viewed as an across-contest analogue. However, in one important respect, the result in Corollary 2 is stronger: it compares contests with (somewhat) different w, m, v, and σ, and shows that, qualitatively, effort comparisons *only* depend on the difference in σ. On the other hand, the single-contest result is more robust, as it extends to global comparisons of σ.

To conclude, a basic insight from the industrial organization literature is what we call the *mitigation principle.* It refers to the fact that, in two-stage games of positioning and competition, equilibrium often entails positioning strategies in the first stage that mitigate the intensity and cost of competition in the second stage. For instance, in a Hotelling "linear city" game with quadratic transportation costs, firms reduce competition by locating at the endpoints. Applied to our model, the mitigation principle suggests that agents of similar ability split up across contests, so as to minimize the cost of head-on competition. Results from small contests confirm this intuition. For example, suppose that two agents of similar ability must choose between two winner-take-all contests, one of which offers a somewhat higher prize. Even though both agents are tempted by the higher prize, they will be wary of the harsh competition that ensues if they enter the same contest. As a result, in all pure-strategy equilibria, the agents split up.

[15] Because $\bar{\lambda} > 1$, a_0 and a_1 always exist in a neighborhood of a symmetric baseline.

Interestingly, we have observed the exact opposite phenomenon in the ponds dilemma: individuals of similar ability enter the *same* pond, especially when $\rho \to 0$. The reason the mitigation principle breaks down is that our large-population assumption precludes "market-impact" effects. That is, the presence or absence of a single agent has no measurable effect on the competitiveness of a contest. As a consequence, the dyadic nature of competition in small contests, which is most intense between agents of the same ability, is lost and replaced by an anonymous battle against a seemingly fixed standard. The result is homophily rather than mitigation because, if one agent pecuniarily strictly prefers a particular pond, so too do all other agents of similar ability. Thus, self-selection in the ponds dilemma induces sorting rather than splitting.

6 Related Literature

The ponds dilemma has stimulated interest since antiquity and, as attested by the many books and articles written about it, continues to do so. Caesar's claim that he would "rather be first in this village than second in Rome" (Plutarchus), is perhaps the earliest literary expression of the dilemma. It is also the title of a recent paper on the topic by Damiano *et al.* (2010). The actual phrase "big fish in a small pond" is of American origin. It appears to have been coined by the Galveston Daily News in 1881. The popular book by Frank (1985), entitled "Choosing the Right Pond," builds on and expands the metaphor. And in his recent best-seller "David and Goliath," Gladwell (2013) continues the discussion, arguing in favor of being a big fish in a small pond, rather than a small fish in a big pond.

Generally, the technical literature on the ponds dilemma is of recent vintage. It includes papers by Leuven *et al.* (2011, 2010), Azmat and Möller (2012, 2009), Konrad and Kovenock (2012), and Damiano *et al.* (2010, 2012). An important exception is the seminal paper by Lazear and Rosen (1981) on rank-order tournaments. Even though their main focus is on competition in a single contest, Lazear and Rosen do show that self-selection sorts workers inefficiently across contests.

In Leuven *et al.* (2010), abilities are binary and success is determined by a parametric contest success function (Tullock, 1980). Their main finding is that high-ability individuals are not necessarily attracted by higher prizes. This is consistent with our findings in so far as, also in our model, the attractiveness of higher prizes is non-monotonic in ability.

Meritocracy, show-up fees, and number of prizes do not feature in their analysis. Leuven *et al.* (2011) conduct a field experiment to disentangle the selection and incentive effects of contests. They find that selection effects dominate.

Azmat and Möller (2009) study how competing contests should be structured in order to maximize participation. Their main finding—for which they find empirical support in professional road running—is that the more discriminatory the contest, the more prizes should be offered. Unlike in our model, contestants in Azmat and Möller (2009) are identical in ability and the contest success function is parametric. In Azmat and Möller (2012), abilities are binary. The authors show that the fraction of high-ability agents choosing the more competitive, high-prize contest is a decreasing function of their population share. Data on entry into marathons support their finding. Also studying entry into contests, Konrad and Kovenock (2012) show that mixing in the entry stage can lead to coordination failure in entry decisions. This coordination failure shelters rents, even among homogenous contestants.

Damiano *et al.* (2010, 2012) study sorting across organizations, focusing on pecking order and peer effects. In their 2010 paper, individuals only care about the average ability of their peers and their own place in the pecking order. The authors show that high and low types self-segregate, while middling sorts are present in both organizations. In their 2012 paper, individuals still care about the average ability of their peers, but money is now a consideration as well. The competing organizations try to maximize the average ability of their workforce. The authors show that, while both organizations attract some high-ability types, equilibrium is asymmetric. Moreover, the 'low-ability' organization offers a steeper wage schedule than the 'high-ability' organization.

Compared to the extant literature, our modeling innovations allow us to analyze the ponds dilemma in considerable generality. We do not restrict the distribution of abilities and allow for a broad class of contest success functions. Simultaneous differences in discriminatoriness, show-up fees, the number and value of prizes are also an original contribution of the current paper. Our analysis uncovers an interplay between direct and indirect effects of differences in structural parameters. Jointly, these effects explain the selection behavior observed in our model.

A small but growing literature considers self-selection into alternative remuneration schemes. Lazear (2000) studies output per worker in a firm that changes from fixed wages to piece rates. He finds that as much as fifty percent of the resulting increase in productivity comes from positive selection, while the other half can be attributed to an increase in the productivity of existing workers. In the auction literature, Moldovanu *et al.* (2008) consider quantity competition between two auction sites, while McAfee (1993), Peters and Severinov (1997), and Burguet and Sakovics (1999) study competition by means of reserve prices.

One of the workhorses of the labor literature is the Roy model (see Roy, 1951, as well as Borjas, 1987, Heckman and Honoré, 1990, and Heckman and Taber, 2008.) As in our model, agents in the Roy model self-select into the sector that provides them with the highest expected payoff. An important difference is that ability in the Roy model is sector-specific. Multidimensionality of ability implies that entry decisions are driven by comparative advantage. Depending on the variances and correlation of an agent's abilities in the two sectors, either sector may benefit from positive selection. In our model, comparative advantage plays no role because an agent's ability is the same in both contests. On the other hand, we allow for general-equilibrium effects that are absent from the Roy model. Specifically, additional entry into a contest negatively affects the expected payoffs of agents already there. In turn, this may induce these agents to reconsider their own choice of contest. Similarly, changes in effort in one contest affect equilibrium entry and effort in both contests.

Finally, our paper is also related to the literature on Hotelling's "linear city" model and its many variants. (See d'Aspremont *et al.*, 1979, for a correction to the original analysis by Hotelling, 1929.) In these games of positioning and competition, equilibrium entails positioning strategies that mitigate competition. We have referred to this phenomenon as the "mitigation principle." Applied to the ponds dilemma, the mitigation principle suggests that individuals of similar ability split up across contests in order to soften competition down the line. Perhaps surprisingly, we have shown that the exact opposite occurs in our model: individuals of similar ability enter the same contest. As we have argued, this is a consequence of our focusing on large contests.

7 Conclusion

In this paper we have analyzed various versions of the ponds dilemma, the question whether it is better to be a big fish in a small pond or a small fish in a big pond. A common intuition is that bigger fish (i.e., those of higher ability) are more likely to choose the big pond (i.e., the contest with the greater rewards and more intense competition). Unless one of the ponds is entirely uncompetitive, we have shown that this intuition is incorrect. The key insight is that the likelihood ratio of success across contests varies non-monotonically with ability, because extreme types can safely ignore differences in competitiveness while middling sorts cannot. This non-monotonicity carries over to selection such that, in large regions of the ability distribution, bigger fish are *less* likely to choose the big pond.

Changes in reward structures can have unexpected selection effects. For instance, offering a higher show-up fee makes the distribution of contestants bimodal, since such a policy attracts the extremes while repelling the middle. Even a seemingly straightforward increase in prize values yields non-obvious selection patterns, owing to the competitive changes wrought. While higher prizes do attract high types, they also drive out middling sorts and have little effect on the selection of low types. As a result, a contest may well raise the value of its prizes, only to see the average ability of contestants fall.

A different kind of trade-off arises when contests differ in meritocracy (i.e., discriminatoriness). Here, agents must compare the benefit of a "lucky break" in measured performance against the cost of an "unlucky break." We obtain the intuitive result that higher types are overrepresented in the more meritocratic contest, while lower types are underrepresented. However, selection effects attenuate toward the tails, because extreme types find meritocracy almost irrelevant to their choice of contest. This has the striking implication that very low types disproportionately enter the more meritocratic contest.

Our model is quite general in a number of respects. We impose essentially no restrictions on the distribution of abilities. We allow for both pecuniary and non-pecuniary preferences, encompassing the "neoclassical" case where the latter are vanishingly small. And, apart from log-concavity and the restriction to location-scale families, we make few assumptions about the distribution of noise in performance measurement, i.e., the "contest success function". Probably the most important limitation of our model is the restriction to prizes of equal value within each contest. Unfortunately, we do not see an easy way to relax this assumption in a tractable manner.

In his lecture notes on the Roy model, Autor (2003) observes that "self-selection points to the existence of equilibrium relationships that should be observed in ecological data, and these can be tested without an instrument. In fact, there are some natural sciences that proceed almost entirely without experimentation—for example, astrophysics. How do they do it? Models predict non-obvious relationships in data. These implications can be verified or refuted by data, and this evidence strengthens or overturns the hypotheses. Many economists seem to have forgotten this methodology." We believe that some of the predictions of our model constitute such non-obvious relationships. We look forward to them being verified—or perhaps refuted—by the data.

References

[1] Autor, David, 2003, "Lecture Note: Self-Selection — The Roy Model," available at *http://economics.mit.edu/files/551.*

[2] d'Aspremont, C., J. Jaskold Gabszewicz, and J.-F. Thisse, 1979, "On Hotelling's "Stability in Competition"," *Econometrica* 47: 1145-1150.

[3] Azmat, Ghazala and Marc Möller, 2009, "Competition Among Contests," *RAND Journal of Economics* 40: 743–768.

[4] Azmat, Ghazala and Marc Möller, 2012, "The Distribution of Talent across Contests," Barcelona GSE Working Paper, January 2012.

[5] Borjas, George, 1987, "Self-Selection and the Earnings of Immigrants" *American Economic Review.*

[6] Burguet, R. and J. Sakovics, 1999, "Imperfect Competition in Auction Designs," *International Economic Review* 40: 231–247.

[7] Damiano, Ettore, Li Hao and Wing Suen, 2010, "First in Village or Second in Rome," *International Economic Review* 51: 263-288.

[8] Damiano, Ettore, Li Hao and Wing Suen, 2012, "Competing for Talents," *Journal of Economic Theory* 147: 2190-2219.

[9] Frank, Robert, 1985, *Choosing the Right Pond: Human Behavior and the Quest for Status.* New York: Oxford University Press.

[10] Ganuza, Juan-José and José S. Penalva, 2006, "On Information and Competition in Private Value Auctions," mimeo, available at: *htpp://papers.ssrn.com/sol3/papers.cfm?abstract_id=1001885.*

[11] Gladwell, Malcolm, 2013, *David and Goliath: Underdogs, Misfits, and the Art of Battling Giants.* New York: Little, Brown, and Company.

[12] Heckman, James and Bo Honoré, 1990, "The Empirical Content of the Roy Model," *Econometrica* 58: 1121–49.

[13] Heckman, James and Christopher Taber, 2008, "Roy model," *The New Palgrave Dictionary of Economics.* 2nd edition. Eds. Steven Durlauf and Lawrence Blume. Palgrave Macmillan.

[14] Hotelling, Harold, 1929, "Stability in Competition," *Economic Journal* 39: 41–57.

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

[16] Konrad, Kai and Dan Kovenock, 2012, "The Lifeboat Problem," *European Economic Review* 56: 552-559.

[17] Lazear, Edward, 2000, "Performance Pay and Productivity," *American Economic Review* 90: 1346–61.

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

[19] Leuven, Edwin, Hessel Oosterbeek and Bas van der Klaauw, 2010, "Splitting Tournaments," IZA Discussion Paper 5186.

[20] Leuven, Edwin, Hessel Oosterbeek, Joep Sonnemans, and Bas van der Klaauw, 2011, "Incentives versus Sorting in Tournaments: Evidence from a Field Experiment," *Journal of Labor Economics* 29: 637-658.

[21] McAfee, R., 1993, "Mechanism Design by Competing Sellers," *Econometrica* 61: 1281–1312.

[22] Moldovanu, Benny, Aner Sela, and X. Shi, 2008, "Competing Auctions with Endogenous Quantities," *Journal of Economic Theory* 141: 1-27.

[23] Peters, M. and S. Severinov, 1997, "Competition among Sellers Who Offer Auctions Instead of Prices," *Journal of Economic Theory* 75: 114–179.

[24] Plutarchus, Lucius Mestrius, *Life of Caesar*, available at: *http://classics.mit.edu/Plutarch/caesar.html*.

[25] Roy, A., 1951, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers* 3: 135–46.

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

A Proofs

A.1 Equilibrium

Proof of Lemma 1: If H_i is such that $\Pr i \leq m_i$, then all individuals win a prize and $\theta_i = -\infty$.

If $\Pr i > m_i$, standard θ_i solves

$$W_i(\theta_i) = \int_{-\infty}^{\infty} \bar{F}\left(\frac{\theta_i - a}{\sigma_i}\right) dH_i(a) = m_i \ .$$

I.e., it equalizes the mass of individuals achieving or exceeding the standard, $W_i(\theta_i)$, to the mass m_i of promotion opportunities. To see that θ_i exists and is unique, notice that: 1) $W_i(\theta_i)$ is continuous and strictly decreasing in θ_i; 2) $W_i(\theta_i) \to \Pr i > m_i$ when $\theta_i \to -\infty$; 3) $W_i(\theta_i) \to 0 < m_i$ when $\theta_i \to \infty$. ■

Proof of Proposition 1: For arbitrary structural parameters, existence of equilibrium was proved in the main text.

For the case of symmetric baselines, we first show that equilibrium standards must be the same across contests. Suppose not. Because $\tau = 0$, strictly more than 50% of each ability type enter the contest with the lower standard; say, Contest 1. This implies that the mass of winners in Contest 1 is strictly greater than in Contest 2. Since $m_1 = m_2 = m$, this is inconsistent with equilibrium.

Finally, as standards are identical across contests and $\tau = 0$, 50% of each ability type enter each contest. Applying Lemma 1 we know that this selection pattern uniquely determines the (identical) standards in the two contests. Hence, equilibrium is unique. ■

A.2 Uncompetitive Case

Proof of Proposition 2: Assume, without loss of generality, that Contest 1 is competitive and Contest 2 is uncompetitive. Then $\theta_1^* > \theta_2^* = -\infty$, such that

$$\pi_1^*(a) - \pi_2^*(a) = w_1 + \bar{F}\left(\frac{\theta_1^* - a}{\sigma_1}\right) v_1 - w_2 - v_2 \ .$$

This payoff difference is strictly increasing in a. By monotonicity of Γ, the same is true for $\Pr 1(a) = \Gamma\left[\frac{\pi_1^*(a) - \pi_2^*(a)}{\rho}\right]$.

To prove FOSD, observe that $G_1(a) < G_2(a)$ iff

$$\frac{\int_{-\infty}^{a} \Pr 1(\alpha)\, g(\alpha)\, d\alpha}{\Pr 1} < \frac{\int_{-\infty}^{a} (1 - \Pr 1(\alpha))\, g(\alpha)\, d\alpha}{1 - \Pr 1} .$$

This is equivalent to

$$\int_{-\infty}^{a} (\Pr 1(\alpha) - \Pr 1)\, g(\alpha)\, d\alpha < 0 .$$

Finally, notice that the last inequality indeed holds for all a, because $\Pr 1(a)$ is strictly increasing in a and the LHS converges to zero when $a \to \infty$.

We now turn to the second part of the proposition. Let $\tilde{\theta}_1^*$ denote the limit value of the standard in Contest 1 as $\rho \to 0$. First, we prove that this limit indeed exists. Suppose to the contrary that, as $\rho \to 0$, there exists a sequence of equilibrium thresholds that does not converge. Then there are at least two convergent subsequences, A, B, with differing limit values $\tilde{\theta}_1^A$, $\tilde{\theta}_1^B$, such that, wlog, $\tilde{\theta}_1^A > \tilde{\theta}_1^B$. The resulting entry pattern in the limit of subsequence A is:

$$\Pr 1(a) \to \begin{cases} 0 & \text{if} \quad a < \tilde{\theta}_1^A - \sigma_1 \bar{F}^{-1}\left(\frac{w_2 - w_1 + v_2}{v_1}\right) \\ 1/2 & \text{if} \quad a = \tilde{\theta}_1^A - \sigma_1 \bar{F}^{-1}\left(\frac{w_2 - w_1 + v_2}{v_1}\right) \\ 1 & \text{otherwise} \end{cases} , \qquad (6)$$

and similarly for sequence B. Since $\tilde{\theta}_1^A > \tilde{\theta}_1^B$, the set of entering types in the limit of sequence A is a strict subset of the set of entering types in the limit of sequence B. Because of this subset property and the fact that $\tilde{\theta}_1^A > \tilde{\theta}_1^B$, the mass of winners under A is strictly smaller than under B. Hence, market clearing must be violated in at least one of these cases. Therefore, all subsequences converge to the same standard, $\tilde{\theta}_1^*$, and entry is as in equation (6) with $\tilde{\theta}_1^A = \tilde{\theta}_1^*$. An analogous argument shows that $\tilde{\theta}_1^*$ is, in fact, unique. ■

A.3 Competitive Case

Recall that interior equilibrium standards (θ_1^*, θ_2^*) are characterized by the market clearing conditions

$$\begin{aligned} \int_{-\infty}^{\infty} \bar{F}\left(\frac{\theta_1 - a}{\sigma_1}\right) dH_1(a) &= m_1 \\ \int_{-\infty}^{\infty} \bar{F}\left(\frac{\theta_2 - a}{\sigma_2}\right) dH_2(a) &= m_2 . \end{aligned}$$

Denote the left-hand side of this system by $S(\theta_1, \theta_2)$ and denote the first and second component of $S(\theta_1, \theta_2)$ by S_1 and S_2, respectively.

The following two lemmas are used in the proof of Proposition 3 below.

Lemma 5 *In a symmetric baseline, the Jacobian of $S(\theta_1, \theta_2)$ is non-singular for generic values of ρ.*

Proof When evaluated at a symmetric baseline, we have to show that

$$\det \left[\begin{array}{cc} \frac{\partial S_1}{\partial \theta_1} & \frac{\partial S_1}{\partial \theta_2} \\ \frac{\partial S_2}{\partial \theta_1} & \frac{\partial S_2}{\partial \theta_2} \end{array} \right] \Bigg|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)} \neq 0 \ .$$

Notice that

$$\frac{\partial S_1}{\partial \theta_1} = \int_{-\infty}^{\infty} \bar{F}_1 \frac{\partial h_1(a, \theta_1, \theta_2)}{\partial \theta_1} da + \int_{-\infty}^{\infty} \frac{1}{\sigma_1} f_1 h_1(a, \theta_1, \theta_2)\, da \ ,$$

where $\bar{F}_i$ and f_i are short for $\bar{F}\left(\frac{\theta_i - a}{\sigma_i}\right)$ and $f\left(\frac{\theta_i - a}{\sigma_i}\right)$. Similarly,

$$\frac{\partial S_1}{\partial \theta_2} = \int_{-\infty}^{\infty} \bar{F}_1 \frac{\partial h_1(a, \theta_1, \theta_2)}{\partial \theta_2} da \ .$$

Recall that $h_1(a, \theta_1, \theta_2) = g(a)\,\Gamma\left(\frac{\pi_1 - \pi_2}{\rho}\right)$. Differentiating h_1 with respect to θ_1, we find

$$\frac{\partial h_1(a, \theta_1, \theta_2)}{\partial \theta_1} = g(a)\frac{1}{\rho}\gamma \frac{\partial \pi_1(a, \theta_1)}{\partial \theta_1} = -g(a)\,\gamma f_1 \frac{v_1}{\rho \sigma_1} \ .$$

Here, γ is short for $\gamma\left[\frac{\pi_1 - \pi_2}{\rho}\right]$.

At a symmetric baseline this reduces to

$$\left.\frac{\partial h_1(a, \theta_1, \theta_2)}{\partial \theta_1}\right|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)} = -g(a)\,\gamma(0)\, f^* \frac{v}{\rho \sigma} = \left.\frac{\partial h_2(a, \theta_1, \theta_2)}{\partial \theta_2}\right|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)} ,$$

where $f^* \equiv f\left(\frac{\theta^* - a}{\sigma}\right)$. Similarly,

$$\left.\frac{\partial h_1(a, \theta_1, \theta_2)}{\partial \theta_2}\right|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)} = g(a)\,\gamma(0)\, f^* \frac{v}{\rho \sigma} = \left.\frac{\partial h_2(a, \theta_1, \theta_2)}{\partial \theta_1}\right|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)} .$$

Let $\frac{\partial S_1}{\partial \theta_i^*} \equiv \left.\frac{\partial S_1}{\partial \theta_i}\right|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)}$. Substituting the expressions for $\frac{\partial h_1}{\partial \theta_1}$ and $\frac{\partial h_1}{\partial \theta_2}$ into $\frac{\partial S_1}{\partial \theta_1}$ and $\frac{\partial S_2}{\partial \theta_1}$ we find

$$\frac{\partial S_1}{\partial \theta_1^*} = \int_{-\infty}^{\infty} \left[\Gamma(0) - \bar{F}^* \gamma(0) \frac{v}{\rho}\right] \frac{1}{\sigma} f^* g(a)\, da = \frac{\partial S_2}{\partial \theta_2^*}$$

and

$$\frac{\partial S_2}{\partial \theta_1^*} = \int_{-\infty}^{\infty} \bar{F}^* \gamma(0) \frac{v}{\rho} \frac{1}{\sigma} f^* g(a)\, da = \frac{\partial S_1}{\partial \theta_2^*} \, .$$

Therefore,

$$\frac{\partial S_1}{\partial \theta_1^*} = \Gamma(0) \int_{-\infty}^{\infty} \frac{1}{\sigma} f^* g(a)\, da - \frac{\partial S_1}{\partial \theta_2^*} = \Xi - \frac{\partial S_1}{\partial \theta_2^*} \, ,$$

where $\Xi > 0$.

Due to symmetry, the determinant of the Jacobian of S then simplifies to

$$\det \left[\begin{array}{cc} \frac{\partial S_1}{\partial \theta_1} & \frac{\partial S_1}{\partial \theta_2} \\ \frac{\partial S_2}{\partial \theta_1} & \frac{\partial S_2}{\partial \theta_2} \end{array} \right] \Bigg|_{(\theta_1,\theta_2)=(\theta^*,\theta^*)} = \left(\frac{\partial S_1}{\partial \theta_1^*} + \frac{\partial S_1}{\partial \theta_2^*} \right) \left(\frac{\partial S_1}{\partial \theta_1^*} - \frac{\partial S_1}{\partial \theta_2^*} \right) \, . \tag{7}$$

Since $\frac{\partial S_1}{\partial \theta_1^*} = \Xi - \frac{\partial S_1}{\partial \theta_2^*}$ and $\Xi > 0$, the first factor in (7) is strictly positive at the baseline. Thus, it remains to show that the second factor, $\frac{\partial S_1}{\partial \theta_1^*} - \frac{\partial S_1}{\partial \theta_2^*}$, is nonzero. Substituting and simplifying yields the required condition

$$\int_{-\infty}^{\infty} \left[\Gamma(0) - 2\bar{F}^* \gamma(0) \frac{v}{\rho} \right] \frac{1}{\sigma} f^* g(a)\, da \neq 0 \, . \tag{8}$$

Solving for ρ we get

$$\rho \neq 2v \frac{\gamma(0)}{\Gamma(0)} \frac{\int_{-\infty}^{\infty} \bar{F}^* f^* g(a)\, da}{\int_{-\infty}^{\infty} f^* g(a)\, da} \, .$$

Notice that θ^* does not depend on ρ since, in a symmetric baseline, selection is 50-50 irrespective of ρ. Hence, the RHS is a strictly positive constant. Therefore, generically, the Jacobian is non-singular at the baseline. ■

Lemma 6 *Fix some $\sigma_1, \sigma_2 > 0$, $m_1, m_2 > 0$, and $m_1 + m_2 < 1$. If the contests' show-up fees w and prizes v are sufficiently close together, a measure of individuals strictly greater than m_i enter Contest $i \in \{1, 2\}$ when $\rho \to 0$.*

Proof Suppose to the contrary that, no matter how small $|w_1 - w_2|$ and $|v_1 - v_2|$, when $\rho \to 0$, fewer than m_i individuals enter Contest i. In that case, $\Pr j > 1 - m_i$ individuals enter Contest $j \neq i$. Moreover, since $1 - m_i > m_j$, it follows that Contest j is competitive. This configuration yields standards $\theta_i^* = -\infty$ and $\infty > \theta_j^* \geq \underline{\theta}_j > -\infty$. Here, $\underline{\theta}_j$ is a lower bound on θ_j^* that is reached when only the lowest-ability $(1 - m_i)$-quantile of individuals enter Contest j.

Because $\theta_i^* = -\infty$, all individuals entering Contest i win a prize with certainty. The pecuniary payoff of entering this contest is therefore $w_i + v_i$. The pecuniary payoff of entering

Contest j is $w_j + \bar{F}\left(\frac{\theta_j^* - a}{\sigma_j}\right) v_j$. Hence, when $\rho \to 0$, an ability type a enters Contest j iff

$$w_j + \bar{F}\left(\frac{\theta_j^* - a}{\sigma_j}\right) v_j \geq w_i + v_i \ .$$

This is equivalent to

$$a \geq \theta_j^* - \sigma_j F^{-1}\left(\frac{w_i - w_j + v_i - v_j}{v_j}\right) \ .$$

It follows that, for $\rho \to 0$, $\Pr j$ equals

$$1 - G\left[\theta_j^* - \sigma_j F^{-1}\left(\frac{w_i - w_j + v_i - v_j}{v_j}\right)\right] < 1 - G\left[\underline{\theta}_1 - \sigma_j F^{-1}\left(\frac{w_i - w_j + v_i - v_j}{v_j}\right)\right] \ . \quad (9)$$

Finally, notice that when $w_i - w_j + v_i - v_j < v_j F\left\{\left[\underline{\theta}_1 - G^{-1}(m_i)\right]/\sigma_j\right\}$, the RHS of (9) is strictly smaller than $1 - m_i$. This contradicts the notion that, no matter how small $|w_1 - w_2|$ and $|v_1 - v_2|$, $\Pr j > 1 - m_i$. ■

Proof of Proposition 3: From Proposition 1 we know that, in a symmetric baseline, 50% of every ability type enter each contest. Hence, $\Pr 1 = \Pr 2 = 1/2$. Because $m_1 = m_2 = m < 1/2$, we may conclude that both contests are competitive.

From Lemma 5 we know that the Jacobian of S is non-singular at a symmetric baseline. Hence, at such a point, we may apply the implicit function theorem (IFT) to the market-clearing conditions $S(\theta_1^*, \theta_2^*) = [m, m]^T$. The IFT implies that equilibrium standards (θ_1^*, θ_2^*) remain finite in a neighborhood of structural parameters around a symmetric baseline. Thus, both contests are competitive.

Finally, Lemma 6 implies that the competitive region remains non-degenerate when $\rho \to 0$. ■

Let $f_i \equiv f\left(\frac{\theta_i - a}{\sigma_i}\right)$ and let f_i' denote the derivative of f_i with respect to its argument $\eta_i(a) \equiv \frac{\theta_i - a}{\sigma_i}$.

Proof of Lemma 3:

1(a). Differentiating $\lambda(a)$ with respect to a we obtain

$$\lambda'(a) = \frac{-f_1' f_2 + f_2' f_1}{\sigma (f_2)^2} \ ,$$

which takes the sign of the numerator.

From log-concavity of f we know that $\frac{f'(\cdot)}{f(\cdot)}$ is strictly decreasing. Hence, for $\theta_1 > \theta_2$,

$$f_1'/f_1 < f_2'/f_2 \ .$$

This implies that $\lambda'(a) > 0$.

For $\theta_1 < \theta_2$ the argument is analogous. The result for $\theta_1 = \theta_2$ is trivial.

1(b). If $\theta_1 > \theta_2$, then we know from Part 1 that $\lambda'(a) > 0$. Hence,

$$\underline{\lambda} = \lim_{a \to -\infty} \lambda(a) < \lambda(\theta_2) < 1 < \lambda(\theta_1) < \lim_{a \to \infty} \lambda(a) = \overline{\lambda} \ ,$$

where the second and third inequalities follow from single-peakedness of f. For $\theta_1 < \theta_2$, the argument is analogous.

2(a). Notice that $\lim_{|a| \to \infty} \lambda(a) = 0$ is equivalent to

$$\lim_{|a| \to \infty} \log \frac{f[\eta_1(a)]}{f[\eta_2(a)]} = \lim_{|a| \to \infty} \log f[\eta_1(a)] - \log f[\eta_2(a)] = -\infty \ .$$

Now consider the two cases:

(i) $\eta_1(a) \geq \eta_2(a)$: Then, by concavity of $\log f$,

$$\log f_1 - \log f_2 \leq f_2'/f_2 \cdot [\eta_1(a) - \eta_2(a)] \ . \tag{10}$$

Hence, it suffices to show that the RHS goes to $-\infty$ when $|a| \to \infty$.

First notice that $\lim_{a \to \infty} \eta_1(a) - \eta_2(a) = -\infty$ and $\lim_{a \to -\infty} \eta_1(a) - \eta_2(a) = \infty$, because $\sigma_1 < \sigma_2$ by assumption. Next notice that, by continuity of f, $\lim_{|a| \to \infty} f = 0$, such that $\lim_{|a| \to \infty} \log f = -\infty$. Hence, there must be an argument η_0 and a $\beta > 0$ such that $d \log f / d\eta|_{\eta = \eta_0} < -\beta$. By strict concavity of $\log f$, it follows that $d \log f / d\eta < d \log f / d\eta|_{\eta = \eta_0} < -\beta$ for all $\eta > \eta_0$. Hence, $d \log f / d\eta = f'/f$ is strictly negative and bounded away from zero for $\eta \to \infty$. Similarly, there must be an argument η_0' and a $\beta > 0$ such that $d \log f / d\eta|_{\eta = \eta_0'} > \beta$. By strict concavity of $\log f$, $d \log f / d\eta > \beta$ for all $\eta < \eta_0'$. Hence, $d \log f / d\eta = f'/f$ is strictly positive and bounded away from zero when $\eta \to -\infty$.

It then follows that the RHS of equation (10) must go to minus infinity when $|a| \to \infty$.

(ii) $\eta_1(a) \leq \eta_2(a)$: Then, by concavity of $\log f$,

$$\log f_2 - \log f_1 \geq f_2'/f_2 \cdot [\eta_2(a) - \eta_1(a)] \ ,$$

This is equivalent to

$$\log f_1 - \log f_2 \leq f_2'/f_2 \cdot [\eta_1(a) - \eta_2(a)] \ .$$

An analogous argument as in (i) now establishes the required limit inequality.

2(b). For $\theta_1 \neq \theta_2$, $\bar{\lambda} > \lambda(\theta_1) > 1$ follows immediately from single-peakedness of f around zero and $\sigma_1 < \sigma_2$.

2(c). Single-peakedness of $\lambda(a)$ follows from the sequence of lemmas given below. ■

Proof of Single-peakedness of $\lambda(a)$:

First, we derive the FOC for an extremum.

Lemma 7 *Ability $a' \in (-\infty, \infty)$ is an extremum of $\lambda(a)$ only if*

$$\sigma_2 f_1'/f_1 = \sigma_1 f_2'/f_2 \ . \tag{11}$$

Proof Differentiating $\lambda(a)$ with respect to a reveals

$$\frac{\partial \lambda(a)}{\partial a} = \frac{-f_1' f_2/\sigma_1 + f_2' f_1/\sigma_2}{(f_2)^2} \ .$$

At an interior extremum, $\partial \lambda(a)/\partial a = 0$ and, hence,

$$f_2' f_1 \sigma_1 = f_1' f_2 \sigma_2 \ .$$

Rearranging yields equation (11). ■

Next, we establish properties that must hold at an extremum.

Lemma 8 *Let $\sigma_1 < \sigma_2$. At an interior extremum of $\lambda(a)$:*

1. *$sign(f_1'/f_1) = sign(f_2'/f_2)$.*
2. *If $f_i'/f_i > 0$ then $f_1'/f_1 < f_2'/f_2$. Else if $f_i'/f_i < 0$ then $f_1'/f_1 > f_2'/f_2$.*

Proof Both assertions follow immediately from the FOC (11). ■

The second part of Lemma 8 and strict log-concavity of f now imply:

Lemma 9 *Let $\sigma_1 < \sigma_2$. At an interior extremum of $\lambda(a)$, $f_i'/f_i \lesseqgtr 0$ iff $\eta_1(a) \lesseqgtr \eta_2(a)$.*

Next we derive a necessary and sufficient condition for an interior extremum to be a maximum.

Lemma 10 *Let $\sigma_1 < \sigma_2$. An interior extremum of $\lambda(a)$ is a maximum iff*

$$\frac{f_1''}{f_1'} / \frac{f_1'}{f_1} < \frac{f_2''}{f_2'} / \frac{f_2'}{f_2} \ .$$

Proof Let a' be an interior extremum. Then

$$\lambda''(a)|_{a=a'} = \frac{f_2 f_1''/\sigma_1^2 - f_1 f_2''/\sigma_2^2}{(f_2)^2} ,$$

where we have used the FOC (11) to simplify the numerator. Hence, a' is a maximum iff

$$f_2 f_1''/\sigma_1^2 - f_1 f_2''/\sigma_2^2 < 0 .$$

Since $f_i > 0$, this condition is equivalent to

$$\sigma_2^2 \frac{f_1''}{f_1'} \frac{f_1'}{f_1} < \sigma_1^2 \frac{f_2''}{f_2'} \frac{f_2'}{f_2} \tag{12}$$

From Lemma 7 we know that, at a', f_1'/f_1 takes the same sign as f_2'/f_2. Hence, there are only two cases to consider, depending upon the sign of f_i'/f_i at a'.

Case 1: $f_i'/f_i > 0$

We may then rewrite the inequality in (12) as

$$\frac{\sigma_2^2}{\sigma_1^2} \frac{f_1''}{f_1'} \frac{f_1'}{f_1} / \frac{f_2'}{f_2} < \frac{f_2''}{f_2'} ,$$

which is equivalent to

$$\frac{\sigma_2^2}{\sigma_1^2} \frac{f_1''}{f_1'} \left(\frac{f_1'}{f_1} / \frac{f_2'}{f_2} \right)^2 \frac{f_2'}{f_2} / \frac{f_1'}{f_1} < \frac{f_2''}{f_2'} .$$

Since a' is an extremum, we may substitute for $\frac{f_1'}{f_1}$ using equation (11) to obtain

$$\frac{f_1''}{f_1'} \frac{f_2'}{f_2} / \frac{f_1'}{f_1} < \frac{f_2''}{f_2'} .$$

Once more using that $f_i'/f_i > 0$, the required inequality becomes

$$\frac{f_1''}{f_1'} / \frac{f_1'}{f_1} < \frac{f_2''}{f_2'} / \frac{f_2'}{f_2} .$$

Case 2: $f_i'/f_i < 0$

The argument is analogous to Case 1. ■

Lemma 11 *Let $\sigma_1 < \sigma_2$ and suppose Condition 1 holds. Then an extremum of $\lambda(a)$ is a maximum.*

Proof Recall from Lemma 7 that, at an extremum, f_1'/f_1 takes the same sign as f_2'/f_2. Hence, there are only two cases to consider, depending upon the sign of f_i'/f_i at a'.

Case 1: $f_i'/f_i > 0$

In that case, it follows from Lemma 9 that $0 > \eta_1(a) > \eta_2(a)$. Because, by Condition 1, $\frac{f_i''}{f_i'}/\frac{f_i'}{f_i}$ is strictly decreasing in that range, it then follows that

$$\frac{f_1''}{f_1'}/\frac{f_1'}{f_1} < \frac{f_2''}{f_2'}/\frac{f_2'}{f_2} .$$

The result now follows from Lemma 10.

Case 2: $f_i'/f_i < 0$

The argument is analogous to Case 1. ■

Finally, it remains to show that

Lemma 12 *An interior extremum of $\lambda(a)$ always exists.*

Proof Wlog, assume that $\sigma_1 < \sigma_2$.

Fact 1: The unique value of a for which $\eta_1(a) = \eta_2(a)$ is $\tilde{a} = \frac{\sigma_2\theta_1 - \sigma_1\theta_2}{\sigma_2 - \sigma_1}$.

This follows directly from solving $\eta_1(a) = \eta_2(a)$ for a.

Fact 2: $\eta_1(a) - \eta_2(a)$ exhibits decreasing differences in a.

Observe that

$$\eta_1(a) - \eta_2(a) = \frac{\theta_1 - a}{\sigma_1} - \frac{\theta_2 - a}{\sigma_2} = \frac{\theta_1\sigma_2 - \theta_2\sigma_1 - (\sigma_2 - \sigma_1)a}{\sigma_1\sigma_2} . \tag{13}$$

Because $\sigma_1 < \sigma_2$, we have $\frac{d}{da}[\eta_1(a) - \eta_2(a)] < 0$.

Fact 3: $\eta_1(a) < \eta_2(a)$ iff $a > \tilde{a}$

This follows from Facts 1 and 2.

Fact 4. $\frac{f_i'}{f_i} > 0$ iff $a > \theta_i$

Follows from the fact that f is single-peaked and achieves a maximum at zero.

From these facts we may deduce:

Claim 1: For all $a > \max(\theta_1, \theta_2, \tilde{a})$,

$$\sigma_1 f_2'/f_2 - \sigma_2 f_1'/f_1 < 0 .$$

Proof: Fact 3 implies that for all $a > \max(\theta_1, \theta_2, \tilde{a})$, we have $\eta_1(a) < \eta_2(a)$. From the log-concavity of f we may then conclude that, for these values of a, $f_1'/f_1 > f_2'/f_2$. Fact 4 implies that for all $a > \max(\theta_1, \theta_2, \tilde{a})$, we also have $f_i'/f_i > 0$, $i = 1, 2$. Therefore, $f_1'/f_1 > f_2'/f_2 > 0$. Because $\sigma_1 < \sigma_2$ by assumption, we may conclude that $\sigma_2 f_1'/f_1 > \sigma_1 f_2'/f_2$. QED.

Claim 2: For a sufficiently small

$$\sigma_1 f_2'/f_2 - \sigma_2 f_1'/f_1 > 0$$

Proof: Recall from the proof of part 2(a) of Lemma 3 that $\lim_{a\to-\infty} f_i'/f_i$ stays bounded away from zero.

i) If $\lim_{a\to-\infty} f_i'/f_i \to -K$ for some finite constant $K > 0$, then it may be readily shown that

$$\lim_{a\to-\infty} \sigma_1 f_2'/f_2 - \sigma_2 f_1'/f_1 = (\sigma_2 - \sigma_1) K > 0 \ .$$

ii) If $\lim_{a\to-\infty} f_i'/f_i \to -\infty$, then we claim that, for a sufficiently small,

$$\sigma_1 f_2'/f_2 - \sigma_2 f_1'/f_1 > 0 \ .$$

To see why, notice that $\lim_{a\to-\infty} \eta_1(a) - \eta_2(a) = \infty$ by equation (13), while f_i'/f_i is strictly decreasing in η_i by strict log-concavity of f. Furthermore, by Fact 4, $f_i'/f_i < 0$ for a sufficiently small. Together with $\sigma_1 < \sigma_2$ this implies that the second (negative) term always dominates the first (negative) term. Hence, $\sigma_1 f_2'/f_2 - \sigma_2 f_1'/f_1 > 0$.

iii) Even if f_i'/f_i does not converge when $a \to -\infty$, it must have a convergent subsequence. Repeating the above argument for every convergent subsequence guarantees an interior extremum. QED.

Existence of an extremum now follows from Claims 1 and 2, and the Intermediate Value Theorem. ■

Together, Lemmas 11 and 12 imply single-peakedness of $\lambda(a)$. ■

Proof of Proposition 4:

1) Recall that $\frac{d(\pi_1^* - \pi_2^*)}{da} = f\left(\frac{\theta_2^* - a}{\sigma_2}\right)\left[\lambda(a) - \frac{v_2/v_1}{\sigma_2/\sigma_1}\right]\frac{v_1}{\sigma_1}$. From Lemma 3 part 2) we know that $\lambda(a)$ is single-peaked and converges to zero in the tails. Hence, if $\frac{v_2/v_1}{\sigma_2/\sigma_1} < \overline{\lambda}$, $d(\pi_1^* - \pi_2^*)/da$ is U-shaped, crossing the x-axis twice, first from below and then from above. In turn, this implies that $\pi_1^*(a) - \pi_2^*(a)$ and $\Pr 1(a)$ have the shape claimed in the proposition. If $\frac{v_2/v_1}{\sigma_2/\sigma_1} > \overline{\lambda}$, then $d(\pi_1^* - \pi_2^*)/da < 0$. Hence, $\pi_1^*(a) - \pi_2^*(a)$ and $\Pr 1(a)$ are strictly decreasing in ability.

2) Trivial. ■

A.4 Isolating the Selection Effects

Proof of Remark 1:

The remark is a corollary of Propositions 9, 10, 11, 13 below. ■

Proposition 9 *A higher show-up fee disproportionately attracts the best and the worst, while repelling the middle.*

Formally, let $w_1 > w_2$ *while the contests are otherwise identical. In equilibrium: 1)* $\Pr 1(a)$ *is U-shaped in ability; 2)* $1/2 < \lim_{|a|\to\infty} \Pr 1(a) = \Gamma\left(\frac{w_1 - w_2}{\rho}\right) < 1$*; 3)* $\theta_1^* > \theta_2^*$*; and 4)* $\Pr 1$ *can be greater or smaller than* $1/2$.

Proof First we prove part **3)**, namely, that $\theta_1^* > \theta_2^*$. Suppose, by contradiction, that $\theta_1^* \leq \theta_2^*$. In that case, Contest 1 is pecuniarily strictly more attractive to all agents. Hence, strictly more than 50% of every ability type enter this contest. In combination with $\theta_1^* \leq \theta_2^*$, this means that there are strictly more winners in Contest 1 than in Contest 2. However, this is inconsistent with equilibrium because $m_1 = m_2$.

For $\theta_1^* > \theta_2^*$, Lemma 3 part 1) implies that $\lambda(a)$ is strictly increasing in a and $\underline{\lambda} < 1 < \overline{\lambda}$. From equation (4) it then follows that $d(\pi_1^* - \pi_2^*)/da$ single-crosses zero from below. Hence, $\pi_1^* - \pi_2^*$ and $\Pr 1(a)$ are U-shaped in a. This proves part **1)**.

Finally, part **2)** is proved in the text, while part **4)** is proved by example: Suppose $a \sim N(0,1)$, $\delta \sim N(0, .05)$, and $\varepsilon_i \sim Logistic(0,1)$. Let $w_1 = 1.1 > 1 = w_2$. If $m_i = .1$ and $v_i = 1$, $i \in \{1,2\}$, then $\Pr 1 = .71 > .29 = \Pr 2$. However, if $m_i = .4$ and $v_i = 4$, $i \in \{1,2\}$, then $\Pr 1 = .44 < .56 = \Pr 2$. Hence, $\Pr 1$ may take on values on either side of $1/2$. ■

Proposition 10 *Offering more prizes attracts all types, but disproportionately those of middling ability.*

Formally, let $m_1 > m_2$ *while the contests are otherwise identical. In equilibrium: 1)* $\Pr 1(a)$ *is inverse-U-shaped; 2)* $\lim_{|a|\to\infty} \Pr 1(a) = \frac{1}{2}$*; 3)* $\theta_1^* < \theta_2^*$*; and 4)* $\forall a$, $\Pr 1(a) > \frac{1}{2}$.

Proof The pecuniary payoff difference is

$$\pi_1^* - \pi_2^* = \left[\bar{F}\left(\frac{\theta_1^* - a}{\sigma}\right) - \bar{F}\left(\frac{\theta_2^* - a}{\sigma}\right)\right] v\ .$$

Hence, $\lim_{|a|\to\infty} \Pr 1(a) = \Gamma(0) = \frac{1}{2}$.

To prove that $\theta_1^* < \theta_2^*$, suppose by contradiction that $\theta_1^* \geq \theta_2^*$. In that case, at least 50% of every ability type enter Contest 2. As a result, the number of winners in Contest 2 is greater than the number of winners in Contest 1. This is inconsistent with equilibrium because, by assumption, $m_1 > m_2$.

Because $\theta_1^* < \theta_2^*$ while the contests are otherwise identical in all payoff relevant dimensions, we have that $\Pr 1(a) > 1/2$ for all a.

For $\theta_1^* < \theta_2^*$ we know from Lemma 3 part 1) that $\lambda(a)$ is strictly decreasing, taking on values on either side of 1. Equation (4) then implies that $\pi_1^* - \pi_2^*$ is inverse-U-shaped in a. By monotonicity of Γ, the same holds for $\Pr 1(a)$. ■

The next proposition deals with differences in price values, v. For $\sigma_1 = \sigma_2$, we know from Lemma 3 that $\lambda(a)$ is strictly monotone and takes on values on either side of 1. When contests only differed in show-up fees or number of prizes, this was enough to establish single-peakedness of the payoff difference. (See Propositions 9 and 10, above.) Here, this is no longer the case. Inspection of (4) reveals that the sign of $d(\pi_1^* - \pi_2^*)/da$ also depends on the prize ratio v_2/v_1 and whether $\lambda(a)$ is bounded.

For $\sigma_1 = \sigma_2$, we say that $\lambda(a)$ is bounded if $\underline{\lambda} > 0$ and $\overline{\lambda} < \infty$. For example, the Logistic distribution falls into this category, since its likelihood ratio runs from $e^{-\frac{1}{\sigma}}$ to $e^{\frac{1}{\sigma}}$. We say that $\lambda(a)$ is unbounded if $\underline{\lambda} = 0$ and $\overline{\lambda} = \infty$ for $\sigma_1 = \sigma_2$. The Normal distribution is a case in point.[16] As we now show, monotone selection requires that $v_2/v_1 \notin (\underline{\lambda}, \overline{\lambda})$—i.e., $\lambda(a)$ is bounded *and* the prize ratio is sufficiently lopsided. Alternatively, when $v_2/v_1 \in (\underline{\lambda}, \overline{\lambda})$—i.e., $\lambda(a)$ is unbounded *or* the prize ratio is sufficiently close to 1—then $d(\pi_1^* - \pi_2^*)/da$ changes sign exactly once, which makes $\Pr 1(a)$ is single-peaked.

Proposition 11 *Higher prizes most strongly attract the best-and-the-brightest while not affecting entry decisions of the worst.*

Formally, let $v_i > v_j$ while the contests are otherwise identical. In equilibrium:

1. $\theta_i^* > \theta_j^*$; *2.* $\lim_{a\to-\infty} \Pr 1(a) = 1/2$ *and* $\lim_{a\to\infty} \Pr 1(a) = \Gamma\left(\frac{v_1 - v_2}{\rho}\right)$.

3. If $v_2/v_1 \in (\underline{\lambda}, \overline{\lambda})$ *then: i)* $\exists \hat{a} \in \mathbb{R}$ *such that* $\Pr i(a) \overset{(<)}{>} 1/2$ *iff* $a \overset{(<)}{>} \hat{a}$*; ii)* $\Pr i(a)$ *is U-shaped on* $(-\infty, \hat{a})$*; iii)* $\Pr i(a)$ *is strictly increasing on* $[\hat{a}, \infty)$*; and iv)* $\Pr i$ *can be greater or smaller than* $1/2$.

4. If $v_2/v_1 \notin (\underline{\lambda}, \overline{\lambda})$ *then,* $\forall a \in \mathbb{R}$, $\Pr i(a) > 1/2$ *and strictly increasing.*

Proof

1) The proof is analogous to that of Proposition 9 part 3).

2) Trivial.

[16] For ease of exposition, our definitions ignore the "semi-bounded" cases, where $\underline{\lambda} > 0$ and $\overline{\lambda} = \infty$ or vice versa. For example, the Extreme Value distribution fall into this category. These cases are handled like the bounded or the unbounded case, depending on whether v_2/v_1 is smaller or greater than 1.

3) Suppose, without loss of generality, that $v_1 > v_2$. Part 1) then implies that $\theta_1^* > \theta_2^*$. In turn, we may apply Lemma 3 part 1) to conclude that $\lambda'(a) > 0$. Next, recall that $d(\pi_1^* - \pi_2^*)/da = \frac{1}{\sigma} f\left(\frac{\theta_2^* - a}{\sigma}\right)[\lambda(a) v_1 - v_2]$. Hence, for $v_2/v_1 \in (\underline{\lambda}, \overline{\lambda})$, $d(\pi_1^* - \pi_2^*)/da$ single-crosses zero from below, which makes $\pi_1^* - \pi_2^*$ and $\Pr 1(a)$ U-shaped in a. Now recall from part 2) that $\lim_{a\to-\infty} \Pr 1(a) = 1/2$ and $\lim_{a\to\infty} \Pr 1(a) = \Gamma\left(\frac{v_1 - v_2}{\rho}\right) > 1/2$, where the inequality follows from $v_1 > v_2$. Combining the U-shapedness of $\Pr 1(a)$ with these limit values implies parts *i)* and *ii)*. Part *iii)* is proved by example. Let $a \sim N(0,1)$, $\delta \sim N(0, .05)$, $\varepsilon_i \sim \mathit{Logistic}(0, .6)$, $m_i = .1$, $w_i = 1$, $i \in \{1,2\}$ and $v_2 = 1$. If $v_1 = 2$, then $(\theta_1^*, \theta_2^*) = (1.54, .91)$ and $\Pr 1 = .44 < .56 = \Pr 2$. If $v_1 = 5$, then $(\theta_1^*, \theta_2^*) = (1.73, .47)$ and $\Pr 1 = .51 > .41 = \Pr 2$.

4) From Lemma 3 part 1) we know that $\underline{\lambda} < 1$. Hence, if $v_2/v_1 < \underline{\lambda}$, then $v_1 > v_2$. By part 1), $\theta_1^* > \theta_2^*$. By the expression for $d(\pi_1^* - \pi_2^*)/da$ above, if $v_2/v_1 < \underline{\lambda}$, then $\pi_1^* - \pi_2^*$ and $\Pr 1(a)$ are strictly increasing in a. An analogous proof holds for $v_2/v_1 > \overline{\lambda} > 1$. Finally, to see that $\Pr i(a) > 1/2$, combine part 2) with the observation that $\Pr 1(a)$ is strictly increasing in a. ■

The next proposition, which is similar to Proposition 11 above, deals with the case where contests differ in multiple dimensions but are equally meritocratic.

Proposition 12 *Let $\sigma_1 = \sigma_2$ while other structural parameters are arbitrary. Then* $\Pr 1(a)$ *is either single-peaked or monotone in ability. Specifically:*

1. *If $v_2/v_1 \in (\underline{\lambda}, \overline{\lambda})$, then* $\Pr 1(a)$ *is single-peaked, taking on a minimum iff $\theta_1^* > \theta_2^*$.*
2. *If $v_2/v_1 \notin (\underline{\lambda}, \overline{\lambda})$, then* $\Pr 1(a)$ *is strictly monotone; it is increasing iff $v_2/v_1 < \underline{\lambda}$.*
3. $\lim_{a\to-\infty} \Pr 1(a) = \Gamma\left(\frac{w_1 - w_2}{\rho}\right)$ *and* $\lim_{a\to\infty} \Pr 1(a) = \Gamma\left(\frac{w_1 + v_1 - (w_2 + v_2)}{\rho}\right)$.

Proof 1) Recall that $d(\pi_1^* - \pi_2^*)/da = \frac{1}{\sigma} f\left(\frac{\theta_2^* - a}{\sigma}\right)[\lambda(a) v_1 - v_2]$ while, from part 1 of Lemma 3, we know that $\lambda'(a) \gtreqless 0$ iff $\theta_1^* \gtreqless \theta_2^*$. Hence, if $v_2/v_1 \in (\underline{\lambda}, \overline{\lambda})$, then $\pi_1^*(a) - \pi_2^*(a)$ is single-peaked, taking on a minimum iff $\theta_1^* > \theta_2^*$. Finally, by monotonicity of $\Pr 1(a)$ in $\pi_1^*(a) - \pi_2^*(a)$, $\Pr 1(a)$ inherits these properties.

2) If $v_2/v_1 < \underline{\lambda}$, then $\frac{d(\pi_1^* - \pi_2^*)}{da} = \frac{1}{\sigma} f\left(\frac{\theta_2^* - a}{\sigma}\right)[\lambda(a) v_1 - v_2] > 0$ for all a. If $v_2/v_1 > \overline{\lambda}$, then $d(\pi_1^* - \pi_2^*)/da < 0$ for all a. By monotonicity of Γ, the same holds for $\Pr 1(a)$.

3) Lemma 2. ■

Proposition 13 *Meritocracy attracts high types and repels low types. However, these selection effects dissipate toward the tails. The majority of the population enters the less meritocratic contest.*

Formally, let $\sigma_1 < \sigma_2$ while the contests are otherwise identical. In equilibrium: 1) $\Pr 1(a) \overset{(<)}{>} 1/2$ *iff* $a \overset{(<)}{>} \tilde{a} \equiv \frac{\sigma_2\theta_1^* - \sigma_1\theta_2^*}{\sigma_2 - \sigma_1}$*; 2)* $\lim_{|a|\to\infty} \Pr 1(a) = 1/2$*; 3) For small ρ,* $\Pr 1 < 1/2$*; 4) Either contest may have the higher standard; 5) If Condition 1 holds,* $\Pr 1(a)$ *is single-peaked on either side of $\tilde{a}$.*

Proof 1) From equation (3) it follows that $\pi_1^* - \pi_2^*$ single-crosses zero from below at $\tilde{a} \equiv \frac{\sigma_2\theta_1^* - \sigma_1\theta_2^*}{\sigma_2 - \sigma_1}$. This implies the claim.

2) From equation (3) it also follows that $\lim_{|a|\to\infty} \pi_1^* - \pi_2^* = 0$. This implies the claim.

3) From part 1) we know that $\Pr 1(a)$ single-crosses $1/2$ from below at $\tilde{a}$. This implies that, in the limit for $\rho \to 0$, agents enter Contest 1 iff $a > \tilde{a}$. (A more careful proof of this claim is analogous to the proof of the limit result in Proposition 6, below, and hence omitted.) Therefore, $\lim_{\rho\to 0} \Pr 1 = 1 - G(\tilde{a})$ and $\lim_{\rho\to 0} \Pr 2 = G(\tilde{a})$.

Now suppose by contradiction that $\Pr 1 = 1 - G(\tilde{a}) \geq 1/2$ for $\rho \to 0$. Then,

$$\begin{aligned} m &= \int_{\tilde{a}}^{\infty} \bar{F}\left(\frac{\theta_1^* - a}{\sigma}\right) g(a)\, da > \bar{F}\left(\frac{\theta_1^* - \tilde{a}}{\sigma}\right) \Pr 1 = \bar{F}\left(\frac{\theta_2^* - \tilde{a}}{\sigma}\right) \Pr 1 \\ &\geq \bar{F}\left(\frac{\theta_2^* - \tilde{a}}{\sigma}\right) \Pr 2 > \int_{-\infty}^{\tilde{a}} \bar{F}\left(\frac{\theta_2^* - a}{\sigma}\right) g(a)\, da = m\ , \end{aligned}$$

where we have used that the chance of winning is strictly increasing in a and, at $\tilde{a}$, the same across contests. Contradiction. Hence, in the limit for $\rho \to 0$, $\Pr 1 < 1/2$. By continuity of $\Pr i$ in ρ, the exclusivity of Contest 1 extends to a neighborhood of ρ around zero.

4) We establish the result by example. Let $a \sim N(0,1)$, $\varepsilon_i \sim N(0, \sigma_i^2)$, $\delta \sim N(0, \rho^2)$, $v = w = 1$, $(\sigma_1, \sigma_2) = (0.5, 1)$, and $\rho = 0.1$. If $m = 0.2$, then $\theta_1^* = 0.51 > 0.40 = \theta_2^*$. If $m = 0.1$, then $\theta_1^* = 1.15 < 1.21 = \theta_2^*$. Hence, θ_1^* and θ_2^* cannot be ranked.

5) This claim follows from the expression for $d(\pi_1^* - \pi_2^*)/da$ in equation (4), single-peakedness of $\lambda(a)$ and $\lim_{|a|\to\infty} \lambda(a) = 0$ proved in Lemma 3 part 2), and the fact that $\bar{\lambda} > 1 > \frac{\sigma_1}{\sigma_2}$. ■

A.5 Stochastic Ordering and Limit Behavior

Proof of Proposition 5: We prove the proposition for the case of show-up fees, i.e., $w_1 > w_2$. The proof for $m_1 > m_2$ is analogous.

First we show that $G_1(a)$ starts out above $G_2(a)$. Notice that

$$G_1(a) - G_2(a) = \frac{\Pr 2 \cdot \int_{-\infty}^{a} \Gamma g(\alpha)\, d\alpha - \Pr 1 \cdot \int_{-\infty}^{a} (1-\Gamma) g(\alpha)\, d\alpha}{\Pr 1 \Pr 2}, \tag{14}$$

where Γ is short for $\Gamma\left[\frac{\pi_1^*(\alpha)-\pi_2^*(\alpha)}{\rho}\right]$. Fix α such that $\pi_1^*(\alpha) - \pi_2^*(\alpha) \neq 0$. Then,

$$\lim_{\rho\to 0} \Gamma\left[\frac{\pi_1^*(\alpha)-\pi_2^*(\alpha)}{\rho}\right] = \begin{cases} 0 & \text{if} \quad \pi_1^*(\alpha)-\pi_2^*(\alpha) < 0 \\ 1 & \text{if} \quad \pi_1^*(\alpha)-\pi_2^*(\alpha) > 0 \end{cases}.$$

For $|\alpha|$ sufficiently large, $\pi_1^*(\alpha) - \pi_2^*(\alpha) \approx \frac{w_1 - w_2}{\rho}$. Since $w_1 > w_2$, it follows that for small ρ and $|\alpha|$ large, $\Gamma \approx 1$. Hence, by inspection of equation (14), it may be seen that $G_1(a) > G_2(a)$ when a and ρ are sufficiently small.

Next, it is easily verified that $\frac{d}{da}[G_1(a) - G_2(a)]$ takes the same sign as $\Gamma - \Pr 1$. (For future reference, also note that $\Gamma - \Pr 1$ is U-shaped, because $\pi_1^*(a) - \pi_2^*(a)$ is.) Since $\Gamma \approx 1$ for ρ small and $|a|$ sufficiently large, we may conclude that, in that case, $\Gamma - \Pr 1 > 0$, as is $\frac{d}{da}[G_1(a) - G_2(a)]$. However, because $G_1(a)$ starts out strictly above $G_2(a)$ while $G_1(\infty) = G_2(\infty) = 1$, $\frac{d}{da}[G_1(a) - G_2(a)]$ and, by extension $\Gamma - \Pr 1$, must be strictly negative somewhere. As the latter is U-shaped in a and takes on the same sign as $\frac{d}{da}[G_1(a) - G_2(a)]$, it must be $\frac{d}{da}[G_1(a) - G_2(a)]$ changes signs exactly twice. Taken together, this implies that $G_1(a)$ single-crosses $G_2(a)$ from above. ■

Recall that we have constrained the parameter space such that both contests are competitive in equilibrium. For limit results like the ones in Propositions 14, 15 and 6 below, we require that the entire converging (sub)sequence of equilibria is competitive for $\rho \to 0$. As before, this amounts to the assumption that structural parameters are not too far apart.

Proposition 14 *Let $w_1 > w_2$ while the contests are otherwise identical. When $\rho \to 0$, selection becomes deterministic. Extreme ability types enter Contest 1, while middling sorts enter Contest 2.*

Formally, for any convergent (sub)sequence of equilibria, there exist a pair $\{\underline{a}, \bar{a}\} \in \mathbb{R}^2$, $\underline{a} < \bar{a}$, such that

$$\lim_{\rho\to 0} \Pr 1(a) = \begin{cases} 0 & \textit{if} \quad \underline{a} < a < \bar{a} \\ 1 & \textit{otherwise} \end{cases}.$$

Proof For $\rho \to 0$, consider a convergent (sub)sequence of equilibria with limit standards (θ_1^*, θ_2^*). First, we show that when $\rho \to 0$, $\pi_1^*(a) - \pi_2^*(a)$ remains U-shaped. To see this, notice that the arguments in the proof of Proposition 9 establishing this result for fixed $\rho > 0$ continue to hold without modification when $\rho \to 0$.

U-shapedness of $\pi_2^*(a) - \pi_1^*(a)$ implies that, in pecuniary terms, almost all ability types have strict preferences over contests. Because pecuniary payoffs determine entry decisions when $\rho \to 0$, this means that individuals of the same ability choose the same contest—namely, the one that strictly maximizes their pecuniary payoffs.

Because $\lim_{|a|\to\infty} \pi_1^*(a) - \pi_2^*(a) = w_1 - w_2 > 0$, when $\rho \to 0$ extreme ability types enter Contest 1. Let a' denote the point where $\pi_2^*(a) - \pi_1^*(a)$ takes on its minimum. By assumption, w_1 and w_2 are sufficiently close together such that both contests are competitive. Therefore, it must be that $\pi_1^*(a') - \pi_2^*(a') < 0$ for $\rho \to 0$. Once more using the single-peakedness of $\pi_2^*(a) - \pi_1^*(a)$ we may conclude that, for a convergent (sub)sequence of equilibria, there exist $-\infty < \underline{a} < \bar{a} < \infty$ such that $\lim_{\rho\to 0} \Pr 1 = \begin{cases} 0 & \text{if} \quad \underline{a} < a < \bar{a} \\ 1 & \text{otherwise} \end{cases}$. ■

Proposition 15 *Let $m_1 > m_2$ while the contests are otherwise identical. When $\rho \to 0$, selection remains strictly stochastic. Standards in the two contests converge, while* $\Pr 1(a)$ *remains inverse-U-shaped.*

Formally, for any convergent (sub)sequence of equilibria, $\lim_{\rho\to 0} \theta_1^* = \lim_{\rho\to 0} \theta_2^* = \theta^*$, and

$$\frac{1}{2} < \lim_{\rho\to 0} \Pr 1(a) = \Gamma\left[c\frac{v}{\sigma} f\left(\frac{\theta^* - a}{\sigma}\right)\right] < 1 ,$$

where $c > 0$ is a constant.

Proof For $\rho \to 0$, consider a convergent (sub)sequence of equilibria with limit standards (θ_1^*, θ_2^*). First notice that, in the limit, $\theta_1^* = \theta_2^* = \theta^*$. Otherwise, when $\rho \to 0$, all individuals would enter the contest with the lower performance standard. This is inconsistent with both contests being competitive.

Assuming that $\Pr 1(a)$ converges when $\rho \to 0$, the propensity to enter Contest 1 in the limit is

$$\lim_{\rho\to 0} \Pr 1(a) = \Gamma\left[v \lim_{\rho\to 0} \frac{\bar{F}\left(\frac{\theta_1^*(\rho) - a}{\sigma}\right) - \bar{F}\left(\frac{\theta_2^*(\rho) - a}{\sigma}\right)}{\rho}\right] .$$

Applying l'Hôpital's rule we get

$$\lim_{\rho\to 0} \Pr 1(a) = \Gamma\left[\frac{v}{\sigma} f\left(\frac{\theta^* - a}{\sigma}\right) \lim_{\rho\to 0}\left(\frac{d\theta_2^*}{d\rho} - \frac{d\theta_1^*}{d\rho}\right)\right] . \tag{15}$$

It remains to prove that $\frac{d\theta_2^*}{d\rho} - \frac{d\theta_1^*}{d\rho} \to c$, $0 < c < \infty$, when $\rho \to 0$. First, by contradiction, suppose that there exist multiple convergence points c_1, c_2, where, wlog, $c_1 < c_2$. From equation (15) it then follows that, in the limit with convergence point c_2, a larger fraction of every ability type enters Contest 1 than in the limit with convergence point c_1. In turn, this means that there are strictly more winner under c_2 than under c_1. Hence, market clearing is violated either for c_1 or c_2. We may conclude that, in fact, $c_1 = c_2$.

Next, suppose that $c \leq 0$ or $c = \infty$. If $c \leq 0$ then, in the limit, less than 50% of each ability type enter Contest 1. Because $m_1 > m_2$, this would imply that $\theta_1^* < \theta_2^*$, contradicting our conclusion above that $\theta_1^* = \theta_2^* = \theta^*$. If $c = \infty$ then, in the limit, almost everybody enters Contest 1, contradicting our result above that both contests are competitive. ■

Proof of Proposition 6: First we prove that selection becomes deterministic in the limit for $\rho \to 0$. That is, agents enter high-v Contest 1 iff their ability exceeds a threshold level, $\hat{a}$.

Suppose that $v_2/v_1 \in \left(\underline{\lambda}, \overline{\lambda}\right)$. For v_i and v_j sufficiently close such that both contests are competitive in the limit, consider a converging (sub)sequence of equilibria with limit standards $(\theta_1^*, \theta_2^*) \in (-\infty, \infty)^2$. Because $\theta_1^*, \theta_2^* > -\infty$, the same arguments as in the proof of part 3) of Proposition 11 imply that, for $\rho \to 0$, there continues to exist an $\hat{a} \in \mathbb{R}$ such that $\pi_1^*(a) - \pi_2^*(a) \overset{(<)}{>} 0$ iff $a \overset{(<)}{>} \hat{a}$. Hence, in the limit, (almost) all $a \overset{(<)}{>} \hat{a}$ enter contest $\overset{(2)}{1}$.

When $v_2/v_1 \notin \left[\underline{\lambda}, \overline{\lambda}\right]$, it is never the case that both contests are competitive in the limit. To see this, suppose by contradiction that both contests remain competitive. In that case, part 4) of Proposition 11 continues to hold and, therefore, $\pi_i^*(a) - \pi_j^*(a) > 0$ for all a. Hence, everybody enters the high-prize contest when $\rho \to 0$. Contradiction.

Next we prove the ranking of quantiles for small but positive ρ.

For $\rho \to 0$, consider a converging (sub)sequence of equilibria with limit standards $(\theta_1^*, \theta_2^*) \in [-\infty, \infty)^2$. For a given value of ρ, let $a_i(p_i; \rho)$ denote the ability of an individual in the p_i-th percentile of Contest i. (Formally, $a_i(p_i; \rho) \equiv G_i^{-1}(p_i; \rho)$.) Above, we have shown that sorting is "perfect" in the limit, i.e. individuals choose to enter the contest with the higher prize iff their ability exceeds $\hat{a}$. Hence, $\lim_{\rho \to 0} a_1(p_1; \rho) > \hat{a} > \lim_{\rho \to 0} a_2(p_2; \rho)$. By continuity, $a_1(p_1; \rho) > \hat{a} > a_2(p_2; \rho)$ continues to hold for ρ sufficiently small. This proves the claim. ■

A.6 Endogenous Effort

The following lemma is a useful building block in proving that, in the model with endogenous effort, each H_i induces a uniquely determined θ_i.

Lemma 13 *Properties of $x(a, \theta)$:*

1. *$\frac{dx(a,\theta)}{d\theta}$ is bounded strictly below 1. Formally, there exists a $\zeta > 0$ such that $\frac{dx(a,\theta)}{d\theta} < 1 - \zeta$ for all $x \in \mathbb{R}$.*

2. *For all 'a', $\lim_{\theta \to \infty} \theta - x(a, \theta) = \infty$ and $\lim_{\theta \to -\infty} \theta - x(a, \theta) = -\infty$.*

Proof 1) Implicitly differentiating the FOC for optimal effort we get

$$\frac{dx\,(a\,,\theta)}{d\theta} = \frac{vf'}{vf' + \frac{\partial^2 c(x,a)}{(\partial x)^2}} \ .$$

The SOC for a maximum guarantees that the denominator of this expression is positive. The result then follows from the f' being bounded and $\frac{\partial^2 c(x,a)}{(\partial x)^2}$ being bounded away from zero.

2) The FOC and single-peakedness of f around zero imply that

$$\frac{\partial c}{\partial x} \leq f\,(0)\,v$$

and, therefore,

$$x \leq \left(\frac{\partial c}{\partial x}\right)^{-1} [f\,(0)\,v] \ .$$

where $\left(\frac{\partial c}{\partial x}\right)^{-1}[\cdot]$ denotes the inverse of $\partial c/\partial x$ with respect to x. Hence, x is bounded and $\lim_{\theta \to \infty} \theta - x\,(a\,,\theta) = \infty$.

Next notice that

$$\frac{d}{d\theta}\,[\theta - x\,(a\,,\theta)] = 1 - \frac{dx\,(a\,,\theta)}{d\theta} \ .$$

Part 1) implies that $\frac{d}{d\theta}\,[\theta - x\,(a\,,\theta)]$ is strictly positive and bounded away from zero. Hence, $\lim_{\theta \to -\infty} \theta - x\,(a\,,\theta) = -\infty$. ■

Lemma 13 allows us to show that standards are unique. Formally,

Lemma 14 *In a contest with endogenous effort, there exists a unique equilibrium standard* θ_i *for every* H_i*.*

Proof If H_i is such that $H_i\,(\infty) \leq m_i$, then all individual win a prize and $\theta_i = -\infty$.

If $H_i\,(\infty) > m_i$, then the equilibrium standard θ_i solves

$$W_i\,(\theta_i) = \int_{-\infty}^{\infty} \bar{F}\left[\frac{\theta_i - x_i\,(a, \theta_i)}{\sigma_i}\right] dH_i\,(a) = m_i \ . \tag{16}$$

An implication of Lemma 13 part 2) is that $W_i\,(\theta_i) \to H_i\,(\infty) > m_i$ when $\theta_i \to -\infty$, and $W_i\,(\theta_i) \to 0 < m_i$ when $\theta_i \to \infty$. Continuity of $W_i\,(\theta)$ in θ_i and the intermediate value theorem then imply that there exists a θ_i such that equation (16) holds.

To prove uniqueness it suffices to show that $W_i\,(\theta_i)$ is strictly decreasing in θ_i. By 13 part 1), $dx_i/d\theta_i < 1$. Hence,

$$\frac{d}{d\theta_i}\int_{-\infty}^{\infty} \bar{F}\left[\frac{\theta_i - x_i\,(a, \theta_i)}{\sigma_i}\right] dH_i\,(a) = -\int_{-\infty}^{\infty} \frac{1}{\sigma_i} f\left[\frac{\theta_i - x_i\,(a, \theta_i)}{\sigma_i}\right]\left[1 - \frac{dx_i\,(a, \theta_i)}{d\theta_i}\right] dH_i\,(a) < 0 \ .$$

■

We are now in a position to prove Proposition 7.

Proof of Proposition 7: From Lemma 14 we know that there exists a unique equilibrium standard θ_i for every H_i. All the other steps in the proof of Proposition 7 are identical to those in the exogenous-effort model. ■

Proof of Proposition 8: The result follows from the fact that, in a neighborhood of a symmetric baseline, we may reinterpret an individual's endogenous equilibrium effort as his exogenous ability type, and treat the problem as one of pure selection without effort. This transformation is justified as follows:

(1) In a symmetric baseline, equilibrium effort of each ability type is the same across contests. Since effort is strictly increasing in ability, this implies that there exists a unique mapping from effort to ability, and vice versa.

(2) Since effort is the same in both contests, the cost of effort differences out when calculating the payoff difference across contests in a symmetric baseline.

(3) The envelope theorem implies that, in a neighborhood of structural parameters around a symmetric baseline, we may ignore changes in equilibrium effort when calculating payoff differences across contests. Hence, in such a neighborhood, we can reinterpret (endogenous) equilibrium effort in the symmetric baseline as a (new, but still exogenous) ability type.

Together, these observations imply that, in a neighborhood of a symmetric baseline, the model with endogenous effort is isomorphic to one with exogenous ability/effort types. Hence, all selection results carry over. ■

Proof of Lemma 4: First notice that $d \Pr 1(a)/da$ takes on the same sign as $d(\pi_1^* - \pi_2^*)/da$. Next observe that, by the envelope theorem,

$$\frac{d(\pi_1^* - \pi_2^*)}{da} = \frac{\partial c[x_1^*(a), a]}{\partial a} - \frac{\partial c[x_2^*(a), a]}{\partial a} = \int_{x_2^*(a)}^{x_1^*(a)} \frac{d^2 c(x,a)}{dxda} dx \; .$$

Finally, recall that $d^2 c(x,a)/(dxda)$ is strictly negative by assumption. Hence, $d(\pi_1^* - \pi_2^*)/da$ takes on the same sign as $x_1^*(a) - x_2^*(a)$. ■

B Competitive versus Uncompetitive Contests

In this appendix we derive some results regarding the (un)competitiveness of contests. We begin by identifying situations where only one of the contests is competitive. For w and v, competitiveness turns on the intuitive condition that the difference between contests should not be too large. For example, if Contest 1 offers an enormous show-up fee relative to Contest 2, Contest 1 will attract so many entrants that Contest 2 becomes uncompetitive.

The other two parameters, m and σ, do not (necessarily) have this property. In the case of discriminatoriness, this is intuitive: all else equal, there is little reason to expect one or the other contest to become uncompetitive due to a difference in discriminatoriness. For a difference in the number of prizes, we show that both contests remain competitive provided pecuniary factors dominate. The intuition is as follows. If contestants mainly care about money, small differences in standards are sufficient to induce large differences in entry across contests. Therefore, the endogenous adjustment of standards in response to a difference in the number of prizes suffices to maintain competitiveness of both contests.

The following lemmas formalize these intuitions by stipulating conditions such that one of the contests becomes *uncompetitive* in the case of w and v, or both contests remain *competitive* in the case of σ and m. For the result for w and v, the "all else equal" condition is not needed, i.e., parameters in the two contests can be arbitrary. Formally,

Lemma 15 *Fix all structural parameters save w_1 (v_1). For w_1 (v_1) sufficiently large, Contest 2 is uncompetitive.*

Proof Notice that

$$\begin{aligned}\lim_{v_1\to\infty}\Pr 2 &= 1-\int_{-\infty}^{\infty}\lim_{v_1\to\infty}\Gamma\left\{\frac{1}{\rho}\left[w_1-w_2+v_1\bar{F}\left(\frac{\theta_1^*-a}{\sigma_1}\right)-v_2\bar{F}\left(\frac{\theta_2^*-a}{\sigma_2}\right)\right]\right\}g(a)\,da\\ &\le 1-\int_{-\infty}^{\infty}\lim_{v_1\to\infty}\Gamma\left\{\frac{1}{\rho}\left[w_1-w_2+v_1\bar{F}\left(\frac{\hat{\theta}_1-a}{\sigma_1}\right)-v_2\right]\right\}g(a)\,da=1-1=0\ ,\end{aligned}$$

where $\hat{\theta}_1$ denotes the (finite) market clearing threshold if everybody entered Contest 1. Hence, for v_1 sufficiently large, $\Pr 2 < m_2$, such that Contest 2 is uncompetitive.

The argument for w_1 is analogous. ■

For differences in σ and m, the next two results provide conditions that ensure that both contests remain competitive. Obviously, such results cannot be obtained for arbitrary parameter values since, for sufficiently large differences in w or v, one contest is uncompetitive. Accordingly, we restrict attention to the "all else equal" case. Formally,

Lemma 16 *Suppose the two contests are identical save for their level of discriminatoriness. For all $(\sigma_1,\sigma_2)\in(0,\infty)^2$, both contests are competitive.*

Proof Suppose not. Then the standard in the uncompetitive contest is $-\infty$. Hence, all agents monetarily prefer this contest. As a result, more than 50% of each ability type enter.

However, this is inconsistent with this contest being uncompetitive because $2m < 1 \Leftrightarrow m < 1/2$. Contradiction. ■

If there is an imbalance in the number of prizes across contests, pecuniary considerations must come sufficiently to the fore for both contests to be competitive. The reason is that, when non-pecuniary motives dominate, each contest attracts roughly half of most ability types. Formally,

Lemma 17 *Suppose the two contests are identical save for the number of prizes. For any (m_1, m_2) such that $m_1 + m_2 < 1$, both contests are competitive if ρ is sufficiently small.*

Proof Suppose by contradiction that Contest 2, say, is uncompetitive for ρ sufficiently small. Then $\theta_2^* = -\infty$, while $\theta_1^* > -\infty$ because prizes are scarce in the aggregate. Since $w_2 = w_1$ and $v_2 = v_1$, $\Pr 2$ becomes

$$\Pr 2 = 1 - \int_{-\infty}^{\infty} \Gamma \left\{ \frac{v}{\rho} \left[\bar{F}\left(\frac{\theta_1^* - a}{\sigma_1} \right) - 1 \right] \right\} g(a)\, da \tag{17}$$

Provided that $\lim_{\rho \to 0} \theta_1^*(\rho) > -\infty$, equation (17) implies that $\lim_{\rho \to 0} \Pr 2 \to 1 > m_2$. Notice, however, that this contradicts Contest 2 being uncompetitive for small ρ.

To finish the proof, it remains to show that $\lim_{\rho \to 0} \theta_1^*(\rho) > -\infty$. Suppose to the contrary that $\lim_{\rho \to 0} \theta_1^*(\rho) = -\infty$. Let $W(\rho) = W_1(\rho) + W_2(\rho)$, i.e., $W(\rho)$ is the sum total of winners in both contests for a given value of ρ. Since $\lim_{\rho \to 0} \theta_1^*(\rho) = -\infty$ and Contest 2 is uncompetitive by assumption, we have $\lim_{\rho \to 0} W(\rho) = 1 > m_1 + m_2$. But this contradicts the notion that $\theta_1^*(\rho)$ is an equilibrium standard. Hence, $\lim_{\rho \to 0} \theta_1^*(\rho) > -\infty$. ■