

Lowering the bar and limiting the field: The effect of strategic risk-taking on selection contests*

Dawei Fang[†]
University of Gothenburg

Thomas Noe[‡]
University of Oxford

30th July, 2018

Abstract

Contests for grades, promotions, and job assignments which feature lax standards or consider only limited talent pools are often criticised for being unmeritocratic. We show that low standards and exclusivity can further meritocratic selection when contestants are strategic risk takers. Strategic risk taking bounds the gains to meritocratic contest designers from expanding the contestant pool while low standards mollify risk-taking incentives and thereby increase the correlation between selection and ability. The introduction of strategic risk taking thus results in “clubby” contest designs, featuring less inclusive contestant pools and larger expected rewards to pool members.

*We thank Ron Siegel, Philipp Strack, and seminar audiences at UC Berkeley, Chapman University, and SAET 2017 for helpful comments and suggestions. Any remaining errors are our responsibility. An earlier version of the paper was circulated under the title “Selection Contests Are Always Clubby.” Financial support by Browaldh Stiftelsen and VINNOVA (grant 2010-02449 Fang) is gratefully acknowledged.

[†]Centre for Finance and Department of Economics, University of Gothenburg. Email: dawei.fang@economics.gu.se.

[‡]Saïd Business School and Balliol College, University of Oxford. Email: thomas.noe@sbs.ox.ac.uk.

1 Introduction

Few would doubt that globalization, automation, and the information revolution have significantly increased the gap between the productivity of the talented and the mediocre and thereby increased the role of schemes to select and identify “the best and the brightest”—e.g., educational tests, worker performance evaluations, performance based league-tables for mutual funds—in allocating wealth and status. In fact, many authors assert that this process has led to the emergence of a meritocratic society dominated by a “meritocratic elite,” the “working rich,” whose *entré* into the higher rungs of the income distribution was produced by success in meritocratic contests for selection (Piketty, 2005).

Open competition and high standards are typically viewed as essential characteristics of meritocracy or even as the essence of meritocracy (e.g., Frost, 2017). One prominent survey of social values in fact treats “meritocratic society” as a synonym for “competitive society” (Ekins, 2014). The association between competition and meritocracy is understandable—increasing competition, by raising the bar for selection or increasing the number of candidates considered, can filter out mediocre candidates. However, the lavish rewards and daunting odds associated with highly competitive contests also increase the incentive to take risks, particularly for weaker competitors.

Risk-taking not only introduces noise into the selection process and thus impedes meritocratic selection, but, in some contexts, also brings obvious direct economic costs. For example, fund managers, struggling to win return-based competitions for inclusion in league tables, may increase portfolio risk (Chevalier and Ellison, 1997; Khorana, 2001; Kaniel and Parham, 2017). An incumbent president who fears losing reelection may risk a major regime change by starting a war (Hess and Orphanides, 1995). Local officials, competing for promotion, may manipulate reports of economic output and social stability, providing misleading data for economic diagnosis and plans (Ghanem and Zhang, 2014; Serrato et al., 2016). Firms, under the pressure of being taken over by competitors, may manipulate accounting earnings to sustain a bubbly share price, driving financial instability (Shleifer, 2004). Researchers, competing for a pool of grant awards, may falsify their data or exaggerate its significance (Weinstein, 1979).

More generally, the “perpetual competition” for rewards in meritocracies is believed by many to impose significant psychological costs (e.g., Wooldridge, 2005; Costin, 2014). If contest designers internalize some of these costs, one might expect that they would “dial down” competitive pressure even at the cost of making selection less meritocratic. However, if “competition” is synonymous with “meritocracy,” meritocratic designers should be willing to bear any costs attendant to selecting the best.

This paper shows that strategic risk-taking can break the link between competition and meritocracy. Thus, even meritocratic contest designers, who aim only to further merit-based selection, can adopt policies that dial down competition. Consider first contests in which both the number of contestants and selection places are beyond the designer’s control, e.g., an uni-

iversity admissions contest in which admission places are fixed by the university's capacity and the number of contestants by the number of applicants. Our analysis shows that, if the number of applicants relative to the number of places is sufficiently large, the designer can reduce risk-taking without reducing winner quality by adopting a "relaxed" selection policy which "approves" more applicants than can be admitted given the number of places, and then allocates places over approved applicants using a random lottery.

This mechanism coincides with the elite-university admission scheme proposed in Schwartz (2007). In our framework, the psychological and social benefits of relaxing competitive pressure asserted in Schwartz (2007) can be attained without sacrificing student quality.¹

Next, consider situations in which the number of selection places, the *selection quota*, is fixed but the number of contestants is under the designer's control, e.g., a competition for a firm's CEO position. The firm can either exclude external candidates and run an "in-house" competition or run an "open competition" that considers external as well as internal candidates. Our analysis shows that, when competition between internal candidates is sufficiently fierce, considering external candidates does not improve expected winner quality but does increase risk-taking. If the external candidates are less likely to exhibit ability than internal candidates, an in-house competition can yield strictly higher expected CEO quality, even though it is possible that an external candidate is better than any internal candidate.

Finally, consider situations in which the number of contestants is fixed but the designer controls the selection quota, e.g., a promotion contest within a firm in which the firm can choose the promotion rate. In such situations, selection contests are often used to assign tasks. The designer prefers assigning, i.e., promoting, strong agents to the desirable "selection task" and assigning weak agents to the less desirable "deselection task." We find that, in these situations, risk-taking leads to either full deselection (in which case there is effectively no contest) or is biased toward "Peter Principle" promotion policies: Some performance ranks are deemed sufficient for promotion even though the firm does not expect the promoted candidates in these ranks to be worthy of promotion (Peter and Hull, 1969).²

Social promotions and lax grading by schools can be interpreted as awards to sub-marginal performance ranks. In our analysis, such policies are rational and consistent with meritocratic objectives. For the same reasons, motivational promotions by firms can be rationalized even when such "motivational promotions" have no motivational effect.³ In retention contests, where

¹Using a large all-pay contest setting, which focuses on contestants' effort-bidding strategies while abstracts from contestants' risk-taking strategies, Olszewski and Siegel (2018) find that reducing competition in college admissions by pooling intervals of performance rankings, a policy that generalizes the one proposed in Schwartz (2007), can improve students' welfare in a Pareto sense by reducing student effort, even though pooling reduces the assortativity of the resulting assignment. Our result complements theirs. Our result implies that, if students can take risks, reducing competition need not reduce assortativity.

²The distortions produced by contest selection have been investigated empirically. Barmby et al. (2012), using a structural econometric model, decompose decreases in worker performance post promotion into mean-reversion, incentive change, and mismatching effects. Dillon and Smith (2013) document pervasive mismatching in university admissions.

³For a discussion of social promotion in schools, see Jimerson et al. (2006). For a discussion of motivational

the contest reward is not being dismissed, our results predict that dismissal rates will be lower under strategic risk-taking than when the dismissal rate is fixed purely on the basis of the distribution of contestant ability. This conclusion appears to be consistent with empirical studies of dismissals in mutual funds. For example, Khorana (Tab. 4 1996) finds that only 14% of managers in the lowest performance decile are replaced despite the fact that, as Khorana (2001) documents, replacing low performing managers improves mutual fund returns.

These implications follow from a parsimonious model of contest design in the shadow of contestant risk taking. In the model, n contestants compete, based on contest performance, for m quota places. The m quota places are filled by the contestants whose performance equals or exceeds the m^{th} highest performance. Contestants prefer selection to deselection. A meritocratic contest designer, who rationally anticipates contestant risk taking, sets, depending on case being considered, either the number of contestants or the selection quota.

Each contestant is endowed with *contest ability*, i.e., ability to perform in the contest. Contest ability determines a contestant's expected contest performance. Contest ability is positively related to ability. Contestants know their own contest ability but not the ability of their rivals. There are two types of contestants, strong and weak, with strong contestants having higher contest ability. Contestants can choose any "fair gamble" performance distribution, i.e., any distribution of (nonnegative) contest performance whose mean equals contest ability.

The fair-gambles framework has been adopted in many studies of contests, including political campaigns (Myerson, 1993; Lizzeri, 1999), status contests (Robson, 1992; Becker et al., 2005; Ray and Robson, 2012), and fund manager competitions (Cover, 1974; Seel and Strack, 2013; Fang and Noe, 2016; Strack, 2016). The fair-gambles assumption implies that contestants incur no direct cost of risk-taking; the only (implicit) cost of risk-taking is strategic—playing performance levels in excess of contest ability, through the mean performance constraint, implies also playing performance levels below contest ability, and such low performance levels imply a low probability of selection.

After developing the baseline model, we extend our analysis to show that its key implications—that risk-taking caps the gains from increasing contest participation and biases meritocratic contest designers toward inflating selection quotas—are robust to various modifications of the baseline model, which include but are not limited to (1) endogenous contest ability acquired through costly effort, (2) ex post discretionary filling of the selection quota, and (3) scoring caps that bound contestant performance.

Related literature

This paper models the design of selection contests and thus is generally part of a very broad stream of economics research on selection based on relative rather than absolute performance. This literature argues that selection based on contest performance has a number of advantages promotion in the workplace, see Deeprouse (2006).

over contracted selection based on absolute performance. First, in some contests, e.g., R&D contests or artistic competitions, performance depends on rather complex evaluations which are difficult for outsiders to verify (Che and Gale, 2003). Second, performance might be affected by common time-varying shocks, in which case, optimal contracting would require continuous adjustment of contract terms to reflect the shocks (Lazear and Rosen, 1981; Knoeber and Thurman, 1994). Third, the rank order of performance might be much easier to measure and observe than the absolute differences between contestant performance levels (Lazear and Rosen, 1981). Finally, the designer may have a strong preference for fixing the “wage bill” associated with the contest selection, i.e., the number of contestants selected; under a contracting regime this bill would vary with the realized performance of the contestants (Gürtler and Kräkel, 2010).

Within the contest literature, our model can be categorized as an incomplete-information *unrestricted risk-taking contest* model. Complete-information unrestricted (i.e., fair-gamble) risk-taking contests have been extensively modeled in the literature (Myerson, 1993; Lizzeri, 1999; Becker et al., 2005; Ray and Robson, 2012; Seel and Strack, 2013; Hart, 2015) and are closely related to an even more extensively analyzed mechanism, all-pay auctions (e.g., Barut and Kovenock, 1998; Fang et al., 2018).⁴ In contrast to most other risk-taking models, the main focus of our analysis is not on the welfare consequences of risk-taking *per se* but rather on the welfare consequences of risk-taking’s effect on selection efficiency.⁵

The focus of analysis in risk-taking contest models is intrinsic, i.e., on the strategies played by contestants in the contest itself. The skill required for the contest, contest ability, is fixed and exogenous. In contrast, as pointed out by Lazear and Rosen (1979), in tournament models, the focus of analysis is extrinsic to the contest—the contest’s effect on the *ex ante* effort required to acquire contest ability. Contest strategies are exogenous and fixed to equal to contest ability plus noise.⁶ In fact, both Lazear–Rosen tournament models and Tullock contest models can be interpreted as noisy-ranking contests with exogenous noise.⁷ Ryvkin and Ortmann (2008) and

⁴In addition, some restricted risk-taking models have also been developed. These models, like unrestricted risk-taking frameworks, fix the mean level of performance but, in contrast to unrestricted risk-taking models, only permit contestants to choose risk profiles from a parametric family of distributions (e.g., Gaba and Kalra, 1999; Hvide, 2002; Taylor, 2003; Gaba et al., 2004; Gilpatric, 2009).

⁵Some existing risk-taking models assume that contestants can choose both mean (through effort) and variance of performance (through risk-taking) (Hvide, 2002; Gilpatric, 2009). These models focus on the effect of risk-taking on effort choice. In contrast, while we allow contestants to choose both mean and riskiness of performance in one of our model extensions, our analysis still focuses on strategic risk-taking and selection efficiency.

⁶E.g., student effort when preparing for an examination will depend on the rewards consequent to passing. However, in almost all cases, students “try their best” when actually sitting the exam. In complete-information all-pay auction models of contests, the dichotomy between *ex ante* investment and contest strategy breaks down. In these models, the investment in the contest and performance in the contest are both represented by a monetary payment whose mean level and risk characteristics are simultaneously determined (e.g., Baye et al., 1993). Applying this setting to the two-contestant case, Kawamura and Moreno de Barreda (2014) find that selection efficiency can be increased by biasing the contest in favor of one of the two contestants, when contestants have complete information about rival ability but the contestants are *ex ante* identical to the contest designer.

⁷The Lazear–Rosen tournament framework uses an additive-noise ranking model and frequently assumes that the noise term has a normal distribution. Exceptions include Kalra and Shi (2001), which uses a Lazear–Rosen tournament model with a logistically distributed noise, and Drugov and Ryvkin (2018), which studies the optimal design of tournaments with general noise distributions. The nested Tullock contest framework is equivalent to a

Ryvkin (2010) investigate the selection properties of noisy-ranking contests with non-strategic contestants.⁸ In contrast, we focus on the effect of strategic risk-taking on contest selection. In a restricted risk-taking contest setting, Hvide and Kristiansen (2003) derive a model that produces examples of contests in which increasing competition reduces the expected ability of the best performer. In contrast, we use an unrestricted risk-taking contest setting and derive general conditions under which increasing competition does not favor meritocracy.

Our result—that optimal selection quotas when agents are strategic lead to over-assignment to desirable tasks and over-promotion relative to first best—is consistent with results in the task-assignment literature. In contrast to our result, over-assignment in this literature is driven by extrinsic factors rather than by the quota’s effect on contest performance strategies: ex ante effort choice in Lazear and Rosen (1981) and Gürtler and Kräkel (2010), contestant risk preferences in Prendergast (1992) and Fairburn and Malcomson (2001), and post-promotion motivational effects on self-esteem in Nafziger (2011).⁹

2 Risk taking

2.1 Risk-taking selection contests

Consider a contest in which a *pool* of $n \geq 2$ contestants compete for one of m identical prizes, where $0 < m < n$. Each prize represents a selection place. The number of selection places, m , which we call the *selection quota*, and the number of contestants, n , are fixed in advance and are common knowledge.

There are two types, t , of contestants: strong, S , and weak, W . Whether a contestant is strong or weak is determined by an independent draw from a Bernoulli distribution which assigns probability θ to S , and probability $1 - \theta$ to W . A contestant’s type is the contestant’s private information.

Contestants simultaneously and costlessly choose *risk-taking strategies*. Risk-taking strategies are modeled as the choice of a *performance distribution*, a distribution of the contest performance supported by the non-negative real line. Risk-taking strategies are restricted by the *contest-ability constraint*, i.e., the expected performance of a type- t contestant must equal the type- t contestant’s contest ability, μ_t , $t \in \{S, W\}$. Contest ability is positively related to the contestant’s type (or intrinsic ability), i.e., $0 < \mu_W < \mu_S$. In the baseline model, we fix contest ability. In Section 6.2, we endogenize contest ability by allowing each contestant to acquire

multiplicative-noise ranking model when the noise term has a Weibull (maximum) distribution (Fu and Lu, 2012).

⁸Assuming that performance is a noisy signal of ability due to exogenous noise, Morgan et al. (2015) analyze self-selection into contests with different prizes and noisiness, and Morgan and Várdy (2009) study an employer’s optimal selection strategy when minority job candidates have noisier ability signals than majority job candidates.

⁹Motivational effects are also considered by Kamphorst and Swank (2016) in the context of selecting *between* contestants to fill a fixed quota. Kamphorst and Swank show that these effects may lead firms to promote candidates known to be weaker over more qualified candidates. In contrast to our analysis, the motivational arguments in both Nafziger (2011) and Kamphorst and Swank (2016) require the firm to have information about employee ability.

contest ability through exerting costly effort. As we will show, the qualitative conclusions of our analysis are fairly robust to this extension.

Each contestant's realized performance is independently drawn from his performance distribution. The m contestants with the highest realized performances are selected and the remaining contestants are deselected, with ties broken randomly. Contestants are expected utility maximizers who strictly prefer selection to deselection. Thus, given that risk-taking is costless, each contestant chooses his performance distribution to maximize his *probability of winning* a selection place given his rivals' strategies and the contest's parameters.

2.2 Best reply and the probability of winning function

To determine the effect of contest design on selection efficiency, we first need to characterize equilibrium contestant behavior. We focus on symmetric equilibria in which contestants of the same type all play the same strategy, i.e., each type- t contestant chooses performance distribution F_t with support Supp_t .

A contestant's probability of winning function maps the contestant's realized performance, x , to his probability of being selected and is thus determined endogenously by his rivals' strategies. Because, each contestant faces the same distribution of rivals, and strategies are symmetric, all of the contestants face the same *probability of winning function*, $P : \mathbb{R}_+ \rightarrow [0, 1]$. Note that P is a cumulative distribution function (CDF) and that P must satisfy the following conditions in any symmetric equilibrium:

Continuity and $P(0) = 0$: First, note that P must be continuous. If P had a discontinuity point, say x^o , then, in a symmetric equilibrium, at least one type of contestant, say type- t^o , would be placing point mass on x^o . Because, contestant type is randomly assigned, it is possible that all of the contestants are type t^o . Thus, there would be a positive probability that all the contestants tie at x^o . If this event occurs with positive probability, a type- t^o contestant would be strictly better off transferring mass away from x^o to $x^o + \varepsilon$, for $\varepsilon > 0$ sufficiently small. To see this, note that the transfer's effect on satisfying the contest ability constraint could be made arbitrarily small by shrinking ε to zero while, for all positive ε , no matter how small, the transfer would generate a gain that is bounded below by a strictly positive number. Because no contestant places point mass on 0, continuity implies that $P(0) = 0$.

Concavity: Second, P must be concave. Otherwise, given the continuity of P , there would exist a non-degenerate closed interval over which P was convex and nonlinear. Jensen's inequality would then imply that a gamble between the extreme points of the interval with the same mean as the expected performance over the interval would produce a higher probability of winning than placing weight on the interior of the interval. Thus, no contestant would place any probability weight in the interior of this interval. Hence, P could not increase over the interior of this interval. Given that P is continuous, P would have to be flat over this interval, a contradiction.

Piecewise linearity: Third, P must be piecewise linear. To see this, note that P must not be strictly concave over any non-degenerate interval. Otherwise, by analogous reasons to those offered in the previous paragraph, spreading probability weight over the interval would be strictly dominated by a mean-preserving contraction that transferred all the probability weight to a single interior point of the interval. Given that there are only two types of contestants, at most two interior points of this interval could capture the probability weight placed by contestants. Thus, P could not be strictly concave over the interval, a contradiction. Given that P is continuous, concave, and there is no interval over which P is strictly concave, P must be piecewise linear.

Connected support with zero lower bound: Fourth, the support of P must be connected and the lower bound of the support of P must be zero. Otherwise, there would exist an open interval, say (x', x'') , and $\varepsilon > 0$ such that $[x'', x'' + \varepsilon]$ is in the support of P but (x', x'') does not meet the support of P . Then P would be flat over (x', x'') . However, because P is continuous, non-decreasing, and concave, P would have to be flat for all $x \geq x''$, which would contradict $[x'', x'' + \varepsilon]$ being in the support of P .

No more than one kink over its support: Finally, P can have at most one kink over its support. Otherwise, P would have at least three non-flat linear segments with different slopes. Because there are only two types of contestants, this would require at least one type of contestant to place probability weight over at least two intervals over which the slope of P differs. However, because P is concave, it would imply that this type was placing probability weight over a region in which P was concave and nonlinear. Such a strategy is suboptimal, because by Jensen's inequality, this type would be strictly better off by transferring all the probability weight to a single performance level equal to the type's contest ability, a contradiction.

Our first lemma summarizes the above observations. All of the formal proofs of our results are relegated to the Appendix.

Lemma 1. *In any symmetric equilibrium, the support of the probability of winning function, P , is a closed interval, with zero as the lower bound. Over its support, P is continuous, increasing, concave, with $P(0) = 0$, and is either linear or piecewise linear with one kink.*

Lemma 1 implies that there are only two possible configurations for the probability of winning function, P , in a symmetric equilibrium: P is either linear over its support or piecewise linear with one kink over its support. Figure 1.A illustrates the configuration in which P is linear over its support. In Figure 1.A, the union of Supp_S and Supp_W equals the interval between 0 and the upper endpoint of the support of P . In this configuration, except for a “boundary” case in which the upper bound of Supp_W equals the lower bound of Supp_S , the upper bound of Supp_W lies strictly above the lower bound of Supp_S , which implies that a weak contestant's performance sometimes tops a strong contestant's. Thus, except for the boundary case, we call equilibria with this configuration, *challenge equilibria*.

The linearity of P implies that, for each contestant, any performance distribution over the support of P that satisfies the contest ability constraint is optimal. Thus, in challenge config-

urations, each type’s risk-taking strategy cannot be uniquely identified. Any combination of F_S and F_W that satisfies the contest ability constraint for each type, and produces a probability of winning function that is linear over the union of the types’ supports, sustains a challenge equilibrium. However, we will show that all of these challenge equilibria produce the same probability of winning for each type and thus have the same implications for selection efficiency. We defer to the Appendix a discussion of the construction of performance distributions that sustain the equilibrium configurations (cf. Lemmas A-3 and A-4 in the Appendix).

Figure 1.B illustrates the contrasting configuration in which P has one kink over its support and consists of two linear segments with different positive slopes. The equilibrium strategy of each type of contestant places all probability weight over the interval associated with one of the two segments (the argument is the same as the argument used to show that P cannot have more than one kink over its support). Because strong contestants have higher contest ability than weak contestants, and because the union of Supp_S and Supp_W equals the support of P , the support of a weak type’s performance distribution, Supp_W , is the interval between 0 and the kink, and the support of a strong type’s performance distribution, Supp_S , is the interval between the kink and the upper endpoint of the support of P . Thus, Supp_S and Supp_W are *adjacent* intervals, i.e., the upper endpoint of Supp_W coincides with the lower endpoint of Supp_S . In this configuration, weak contestants concede to strong contestants and concentrate their contest ability on beating other weak contestants. We call equilibria with this configuration, *concession equilibria*. There is also a non-generic boundary case in which P is linear over its support but the supports of the performance distributions are adjacent. We also categorize this case as a concession configuration.

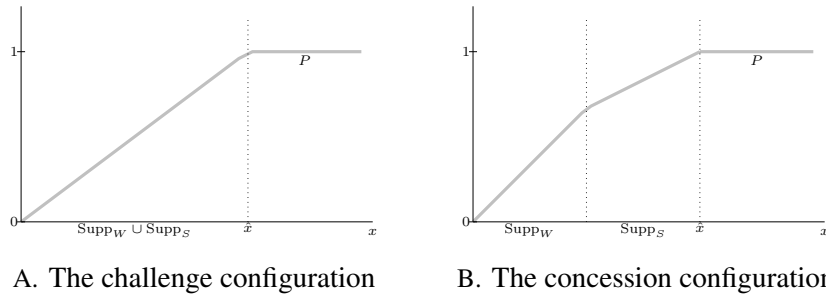


Figure 1: The probability of winning function, P , in the two equilibrium configurations.

2.3 Equilibrium configuration

The question that remains is determining the conditions under which each of these configurations can be sustained in equilibrium. Before we embark on systematic investigation of this question, we illustrate the determinants of sustainability with a couple of examples.

Example 1 (Challenge equilibrium). Suppose there are two contestants and one selection place, i.e., $n = 2$ and $m = 1$. Ex ante, each contestant is equally likely to be strong or weak ($\theta = 0.5$);

the expected performance of the two contestant types is given by $\mu_W = 2$ and $\mu_S = 5$. Because, by Lemma 1, the probability of winning function, P , is continuous, the probability that a given contestant wins a place with realized performance, x , equals the probability that x is no less than the realized performance level of the other contestant. Because there are only two contestants, the probability of winning function, P , for each contestant, is simply the probability of topping the contestant's rival. Each contestant believes that his rival is equally likely to be weak or strong. Thus,

$$P(x) = \frac{F_S(x) + F_W(x)}{2}. \quad (1)$$

In the challenge configuration, P is the CDF of a uniform distribution (cf. Figure 1.A). Thus, given that the average contest ability of the two types is 3.5, in the challenge configuration, P is the CDF of the uniform distribution over $[0, 7]$. This observation and equation (1) imply that

$$P(x) = \frac{x}{7} = \frac{F_S(x) + F_W(x)}{2}, \quad x \in [0, 7]. \quad (2)$$

Note that, while P is uniquely determined, F_W and F_S are not uniquely determined by equation (2). All that is required for F_W and F_S to be equilibrium performance distributions is the satisfaction of (2) and the contest ability constraints. Figure 2.A graphs the probability of winning function and a particular choice of the equilibrium performance distributions.

By the piecewise linearity of P , equation (1), and the adjacent supports condition for concession equilibria, if a concession equilibrium were played, F_W would have to be a uniform distribution and, hence, given that $\mu_W = 2$, weak contestants would play a uniform distribution over $[0, 4]$. For the same reasons, because $\mu_S = 5$, strong contestants would play a uniform distribution over $[4, 6]$. In this case, equation (1) implies that the slope of P would equal $1/8$ between 0 and 4 and equal $1/4$ between 4 and 6. Hence, P would be convex and nonlinear over its support, which implies that a weak contestant's best reply would not be to concede but rather to adopt the high-risk strategy of gambling between 0 and the upper endpoint of P 's support.¹⁰

In Example 1, the challenge and only the challenge configuration can be sustained. However, as Example 2 illustrates, it is not always possible to sustain the challenge configuration.

Example 2 (Concession equilibria). Consider changing Example 1 by reducing the expected performance of the weak type, from $\mu_W = 2$ to $\mu_W = 1$ while retaining all other parametric assumptions of Example 1. In this case, the average contest ability of the two types equals 3. Thus, in a challenge equilibrium, P would be uniformly distributed over $[0, 6]$. Consequently, for the same reasons as noted in Example 1, for a challenge equilibrium to exist, the condition

$$\frac{x}{6} = \frac{F_S(x) + F_W(x)}{2}, \quad x \in [0, 6], \quad (3)$$

¹⁰In some cases, for instance, the case of Example 1 with μ_S reduced to $\mu_S \in (2, 4]$, the strong type's contest ability advantage is so small that no continuous performance distribution satisfying the contest ability constraint can satisfy the adjacent supports condition for concession equilibria, which implies, by definition, that the concession configuration cannot be sustained in equilibrium.

must be satisfied for performance distributions that respect the contest-ability constraints. Given that $F_S \geq 0$, equation (3) implies that $x/3 \geq F_W(x)$ for all $x \in [0, 6]$. Thus, given that F_W is a CDF and is thus bounded above by 1, it would have to be that $\min[x/3, 1] \geq F_W(x)$ for all $x \geq 0$. Hence, given that $\min[x/3, 1]$ is the CDF of a uniform distribution over $[0, 3]$, F_W would stochastically dominate this uniform distribution, which is impossible because the mean of F_W , μ_W , equals 1, whereas the mean of the uniform distribution over $[0, 3]$ equals 1.5.

Consequently, the challenge configuration cannot be sustained: the weak type's contest ability is too small relative to the strong type's to generate a probability of winning function that is linear over the union of both types' supports. In this example, only the concession configuration can be sustained: weak types play a uniform distribution with a lower endpoint of 0 and an upper endpoint determined by the weak type's contest ability constraint while strong types play a uniform distribution whose lower endpoint is the upper endpoint of the support of the weak type's performance distribution and whose upper endpoint is determined by the strong type's contest ability constraint. This case is graphed in Figure 2.B.

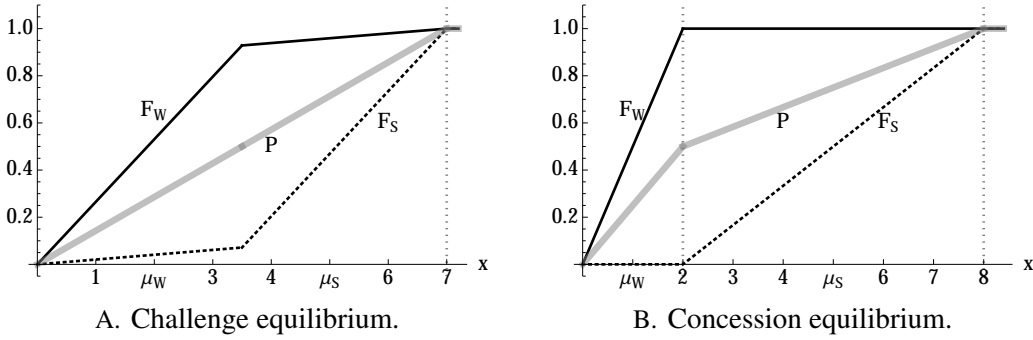


Figure 2: Examples of a challenge equilibrium (Panel A) and a concession equilibrium (Panel B). In the two examples, the common parameters are $n = 2$, $m = 1$, $\mu_S = 5$, and $\theta = 0.5$. In Panel A, $\mu_W = 2$. Weak contestants play the uniform distribution over $[0, 3.5]$ with probability $13/14$ and the uniform distribution over $[3.5, 7]$ with probability $1/14$; strong contestants play the uniform distribution over $[0, 3.5]$ with probability $1/14$ and the uniform distribution over $[3.5, 7]$ with probability $13/14$. In Panel B, $\mu_W = 1$. Weak contestants play the uniform distribution over $[0, 2]$; strong contestants play the uniform distribution over $[2, 8]$.

In Examples 1 and 2, one and only one, configuration emerges in equilibrium for each parameterization of the model. This result holds generally. The key to establishing this assertion as well as to identifying the conditions under which the two configurations sustain equilibria is provided by considering probability of winning to the weak contestant. First, consider the concession configuration. Let p_t^C be the probability of being selected for a contestant of type $t \in \{S, W\}$ in the concession configuration. For a given contestant i , let \tilde{S}_n^{-i} be the number of strong rivals to i . \tilde{S}_n^{-i} is Binomially distributed with parameters $n - 1$ and θ , i.e., $\tilde{S}_n^{-i} \sim \text{Binom}(n - 1, \theta)$. Because, in the concession configuration, weak contestants never outperform strong contestants, and contestants of the same type have the same probability of winning, if contestant i is weak, contestant i has no chance of winning if $\tilde{S}_n^{-i} \geq m$ and has a

probability of winning equal to $(m - \tilde{S}_n^{-i}) / (n - \tilde{S}_n^{-i})$ if $\tilde{S}_n^{-i} < m$.¹¹ Thus,

$$p_W^C = \mathbb{E} \left[\max \left[0, \frac{m - \tilde{S}_n^{-i}}{n - \tilde{S}_n^{-i}} \right] \right]. \quad (4)$$

For the concession configuration to sustain an equilibrium, it must be that a given weak contestant has no incentive to challenge strong contestants. A simple and feasible way for a weak contestant to challenge strong contestants is to mimic strong contestants' strategy with probability μ_W / μ_S and choose 0 with the complementary probability. Under this "mimicking strategy," a weak contestant's expected performance equals $\frac{\mu_W}{\mu_S} \times \mu_S + \left(1 - \frac{\mu_W}{\mu_S}\right) \times 0 = \mu_W$ and, hence, the weak type's contest ability constraint is satisfied. For the concession configuration to sustain an equilibrium, it must be that a weak contestant has no incentive to deviate from a concession strategy to the prescribed mimicking strategy. This requires that

$$p_W^C \geq \frac{\mu_W}{\mu_S} p_S^C, \quad (5)$$

where the right-hand side is the probability of winning for a weak contestant if he deviates to the prescribed mimicking strategy.¹² Note that, in any equilibrium, the expected number of places filled must equal the selection quota, i.e.,

$$m = n(\theta p_S + (1 - \theta) p_W), \quad (6)$$

where p_t represents type- $t \in \{S, W\}$ contestant's equilibrium probability of winning. Thus, by equations (5) and (6), concession equilibria exist only if

$$p_W^C \geq \frac{m}{n} \left(\frac{1}{\theta r + 1 - \theta} \right), \quad (7)$$

where $r = \mu_S / \mu_W$ represents the *strength asymmetry* between strong and weak contestants.

In fact, the right hand side of (7) is just the weak type's probability of winning in the challenge configuration. To see this, let p_t^G be the probability of winning for a contestant of type $t \in \{S, W\}$ in the challenge configuration. Because the probability of winning function, P , is linear over its support in the challenge configuration, choosing a deterministic performance level equal to contest ability is a weakly optimal strategy for each type. Hence, we can evaluate each type's probability of winning in the challenge configuration simply by evaluating P at the type's contest ability. Thus, given that P is linear over its support and meets the origin, it must be that, in the challenge configuration, the ratio between strong and weak types' probabilities of winning, p_S^G / p_W^G , equals their strength asymmetry, $r = \mu_S / \mu_W$, i.e.,

$$\frac{p_S^G}{p_W^G} = \frac{\mu_S}{\mu_W} = r. \quad (8)$$

¹¹Suppose contestant i is weak and the concession equilibrium is played. If $\tilde{S}_n^{-i} < m$, then after the \tilde{S}_n^{-i} strong rivals all win a selection place, there are still $m - \tilde{S}_n^{-i} > 0$ selection places left to be assigned to the $n - \tilde{S}_n^{-i}$ weak contestants, including contestant i . By symmetry, these $n - \tilde{S}_n^{-i}$ weak contestants have the same probability of winning. Thus, each of these weak contestants has a probability of winning equal to $(m - \tilde{S}_n^{-i}) / (n - \tilde{S}_n^{-i})$.

¹²By deviating, the weak contestant's probability of winning equals the strong type's probability of winning in the concession configuration, p_S^C , with probability μ_W / μ_S , and equals 0 with the complementary probability.

Equation (8), combined with identity (6), implies that

$$p_W^G = \frac{m}{n} \left(\frac{1}{\theta r + 1 - \theta} \right). \quad (9)$$

Equations (7) and (9) thus imply that concession equilibria exist only if $p_W^C \geq p_W^G$.

If $p_W^C < p_W^G$, only challenge equilibria can exist. In fact, the condition that $p_W^C < p_W^G$ is necessary for challenge equilibria to exist. This is because, in challenge equilibria, weak contestants not only have a chance of winning by beating weak rivals but also by beating strong rivals. Because the necessary condition for the existence of challenge equilibria, $p_W^C < p_W^G$, and the one for the existence of concession equilibria, $p_W^C \geq p_W^G$, are complementary, and because, as we show in the Appendix, an equilibrium always exists, these necessary conditions are also sufficient conditions. We thus obtain the following proposition.

Proposition 1. *Let p_t be type- $t \in \{S, W\}$ contestant's equilibrium probability of winning.*

- (a) *The concession (challenge) configuration sustains an equilibrium if and only if $p_W^C \geq p_W^G$ ($p_W^C < p_W^G$), where p_W^C and p_W^G are given by equations (4) and (9).*
 (b) *A weak contestant's equilibrium probability of winning, p_W , is given by*

$$p_W = \max \left[p_W^C, p_W^G \right]. \quad (10)$$

- (c) *A strong contestant's equilibrium probability of winning, p_S , is determined by p_W through equation (6).*

Proposition 1 implies that configuration of the equilibrium, for a given parameterization of the model, is determined by weak contestants' preferences. Through adopting high-risk strategies, weak contestants are able to sometimes challenge strong contestants for selection places. However, because of the contest-ability constraint, such challenges require increasing the probability of low performance, performance that is likely to be topped even by weak rivals. High-risk strategies can be sustained in equilibrium only when the benefits of such high-risk strategies outweigh their costs.

The next lemma shows that increasing competition tends to induce weak contestants to challenge strong contestants, and that weak contestants will challenge strong contestants if the selection contest is sufficiently competitive.

Lemma 2. *Everything else being equal, the challenge configuration will be played in equilibrium if (a) contest size, n , is sufficiently large, or (b) strength asymmetry, r , is sufficiently small (i.e., sufficiently close to 1).*

If the challenge configuration is played in equilibrium, the challenge configuration will also be played in equilibrium if (a) contest size, n , increases, (b) selection quota, m , decreases, or (c) strength asymmetry, r , decreases.

Increasing contest size or decreasing the selection quota increases the proportion of rivals that must be topped to win a place. Both these parameter changes make it less likely that

besting only weak rivals is sufficient for a weak contestant to be selected. This increases weak contestants' incentive to challenge strong contestants through high-risk strategies. Reducing strength asymmetry increases weak contestants' contest ability relative to strong contestants', making it easier for weak contestants to best strong contestants. This also increases weak contestants' incentive to challenge strong contestants.

3 Risk-taking and meritocracy

For a fixed contest design—contestant pool, n and the selection quota, m —the contest designer's welfare is determined by the ability of contestants selected. As discussed in the introduction, if the designer's welfare function is meritocratic, she prefers to select strong contestants and deselect weak contestants. However, the designer does not know, *ex ante*, which contestants are strong and which are weak. Moreover, as we have seen from the results in Section 2, the outcome of the contest may not perfectly reveal ability. Thus, whether the selection of a contestant of unknown type will increase designer welfare depends on the tradeoff between the increase in welfare that results if the contestant turns out to be strong versus the decrease in welfare that results if the contestant turns out to be weak. We incorporate this tradeoff into the analysis with a simple linear specification—we assume that the designer maximizes the expectation of

$$\# \text{Strong Selected Contestants} - \sigma \times \# \text{Weak Selected Contestants}, \quad \sigma > 0. \quad (11)$$

In equation (11), σ measures the relative costs of the two types of misselection the designer might make, overselection, selecting weak contestants, and underselection, deselecting strong contestants. If $\sigma < (>)1$, then underselection is more (less) costly than overselection. If $\sigma = 1$, then both types of misselection are equally costly. Thus, we think of σ as a measure of the designer's overselection aversion.

For a given selection quota, m , and contestant pool, n , let $u(m, n)$ represent the designer's welfare and let $\Pi(m, n)$ be a selected contestant's probability of being strong. Consistent with equation (11), the designer's welfare, u , is given by

$$u(m, n) = m\Pi(m, n) - \sigma m(1 - \Pi(m, n)) = m((1 + \sigma)\Pi(m, n) - \sigma). \quad (12)$$

An application of Bayes rule shows that, in the risk-taking contest, the probability that a selected contestant is strong, Π , is given by

$$\Pi(m, n) = \frac{\theta p_S}{\theta p_S + (1 - \theta)p_W} = 1 - \frac{n}{m}(1 - \theta)p_W = 1 - \frac{n}{m}(1 - \theta) \max[p_W^C, p_W^G], \quad (13)$$

where the second equality follows from equation (6) and the last equality from (10).

The fundamental question we aim to address is how contestant risk-taking affects the attainment of the meritocratic ideal embodied in the designer's objective function—select the best and only the best. Because holding a contest requires the designer to commit to fixed number of selection places and the realized number of strong contestants is random, there is always

a possibility that some places will be filled by weak contestants and some strong contestants will not receive a place. Thus, the very fact that selection is contest-based implies that contest selection cannot attain the meritocratic ideal.

If we abstract from the problem of an imperfect relationship between contest performance and ability, but respect the constraint imposed by selection being contest-based, the best possible selection strategy for the designer is to prioritize strong contestants, i.e., select weak contestants to fill the quota only after all strong contestants have been selected. We term this policy *merit-based selection* and represent designer welfare under merit-based selection with u_M . Merit-based selection of contestants thus represents the closest approximation of the meritocratic ideal possible in a contest setting. Thus, merit-based selection will be the benchmark that we use to measure the effects of risk-taking on meritocracy.

Under the concession configuration of the risk-taking contest, weak contestants never outperform strong contestants. Although weak contestants do take risks in the sense of choosing random performance, their risk-taking is not “excessive” in the sense that it is undertaken only to top other weak contestants and thus does not affect the performance ranking of strong contestants relative to weak contestants. Thus, because places are allocated based on performance ranking, concession equilibria implement merit-based selection.

In contrast, if the challenge configuration is played, it is possible for the performance of a weak contestant to top the performance of a strong contestant. Because of such “excessive” risk-taking of weak contestants, contest performance ranking is not always consistent with ability ranking. Because places are allocated based on performance ranking, strong contestants are not always prioritized. Given the designer’s meritocratic preferences, this implies that designer’s welfare in challenge equilibria is lower than her welfare under merit-based selection. These observations are formalized below.

Lemma 3. *For a given selection quota, m , and contestant pool size, n , the designer’s welfare, $u(m, n)$, and welfare under merit-based selection, $u_M(m, n)$, satisfy the following conditions:*

- i. if $p_W^C \geq p_W^G$, only concession equilibria exist, and $u(m, n) = u_M(m, n)$;*
 - ii. if $p_W^C < p_W^G$, only challenge equilibria exist, and $u(m, n) < u_M(m, n)$,*
- where p_W^C and p_W^G are defined by equations (4) and (9) respectively.*

Lemma 3 shows that, for a fixed contest design, the noise generated by strategic risk-taking can lower designer welfare by making contest selection less meritocratic. In the following two sections, we examine how meritocratic contest designers respond to contestants’ strategic risk-taking when designing contests.

4 Risk-taking and the size of the contestant pool

In this section, we consider the effect on designer welfare of varying the size of the contestant pool, n , for a fixed selection quota, m . Inspection of equation (12) shows that designer welfare, for a fixed selection quota, only depends on winner quality, Π .

To identify the effect of risk taking, we first examine the effect of varying n under merit-based selection. Let \tilde{S}_n be the number of strong contestants in the pool of n contestants. Note that $\tilde{S}_n \sim \text{Binom}(n, \theta)$. Under merit-based selection, the designer first attempts to fill the m places with strong contestants and then fills any residual places with weak contestants. Thus, under merit-based selection, the number of strong selected contestants equals $\min[\tilde{S}_n, m]$. Now suppose we add a contestant to the pool. Let \tilde{s} be an independent Bernoulli random variable, which equals 1 with probability θ and 0 otherwise. Under merit-based selection, after adding a contestant to the pool, the number of strong selected contestants equals $\min[\tilde{S}_n + \tilde{s}, m]$. Clearly, $\min[\tilde{S}_n + \tilde{s}, m] \geq \min[\tilde{S}_n, m]$ and, with positive probability, $\min[\tilde{S}_n + \tilde{s}, m] > \min[\tilde{S}_n, m]$.¹³

Thus, under merit-based selection, adding contestants increases the expected number of strong selected contestants and thus designer welfare.

Now consider the effect of pool expansion in the risk-taking contest. As shown by Lemma 3, in the concession configuration, the designer's welfare, u , equals her welfare under merit-based selection, u_M . Given that $n \mapsto u_M(m, n)$ is strictly increasing, pool expansion increases the designer's welfare as long as the concession configuration is played. In the challenge configuration, equations (9) and (13) imply that, winner quality satisfies

$$\Pi = \frac{\theta r}{\theta r + 1 - \theta}. \quad (14)$$

Equation (14) reveals that, in contrast to the concession configuration, in the challenge configuration, winner quality is independent of the size of the contestant pool. By Lemma 2, once contest size is sufficiently large, the challenge configuration will be played. Thus, adding further contestants to a sufficiently large contestant pool will not increase winner quality and, hence, will not increase designer welfare. This argument yields the following characterization.

Theorem 1. *For any fixed selection quota, m , there exists n^c , such that the marginal effect of increasing the contestant pool on designer welfare is positive, i.e., $u(m, n+1) > u(m, n)$, only if $n < n^c$. Moreover,*

- i. *if $n \geq n^c$, increasing the contestant pool does not increase designer welfare, i.e., $u(m, n+1) = u(m, n)$.*
- ii. *If $n > n^c$, designer welfare in the risk-taking contest is less than designer welfare under merit-base selection, i.e., $u(m, n) < u_M(m, n)$.*

The basic implication of Theorem 1 is that risk-taking caps the gains from inclusivity. When making the contestant pool more inclusive is costly because of outreach, advertisement, or search costs, the optimal contest size under risk-taking selection will tend to be smaller than under merit-based selection. As we will show in Section 6.3, even when increasing pool size is costless, when the pool of potential new contestants is, on average, of lower quality than the incumbent candidate pool, the designer may strictly gain from excluding the potential contestants

¹³Because $m \geq 1$, $\min[\tilde{S}_n + \tilde{s}, m] > \min[\tilde{S}_n, m]$ if $\tilde{S}_n = 0$ and $\tilde{s} = 1$. Given that $\tilde{S}_n \sim \text{Binom}(n, \theta)$ and $\mathbb{P}\{\tilde{s} = 1\} = \theta$, $0 < \theta < 1$, the state of $\tilde{S}_n = 0$ and $\tilde{s} = 1$ occurs with positive probability.

from the contest. In such cases, the gain from expanding the pool produced by increasing the expected number of strong contestants is overwhelmed by the cost of increased risk taking. In contrast, under merit-based selection, the designer always strictly gains from inclusion because adding contestants increases the expected number of strong contestants.

Thus, in risk-taking contests, even meritocratic designers who are not biased toward specific contestants have little incentive to expand candidate fields and sometime will deliberately restrict consideration to candidates who, *ex ante*, look promising, even if considering a wider field is costless. Consequently, if social welfare values inclusivity, ensuring inclusive competition for rewards requires policy intervention, even if private decision-makers are not inherently biased toward exclusivity or toward specific groups of potential candidates.

5 Risk-taking and the selection quota

In this part, we fix the size of the contestant pool, n , and consider the welfare effects of varying the selection quota, m . Given that n is fixed, we suppress the dependence of the designer's welfare on n . Note that, because n is fixed, the expected number of weak contestants is fixed. This implies that the sum of the expected number of weak selected contestants and the expected number of weak deselected contestants is fixed. Inspection of (11) then shows that the designer's problem is equivalent to maximizing the expectation of

$$\# \text{Strong Selected Contestants} + \sigma \times \# \text{Weak Deselected Contestants}.$$

Thus, for fixed n , the selection problem is equivalent to the classic task assignment problem. In this problem, the designer assigns a fixed pool of contestants either to a "selection task" or a "deselection task." The marginal product of strong contestants is higher when performing the selection task and the marginal product of weak contestants is higher when performing the deselection task.

We first determine the optimal quota under merit-based selection by considering the effect of reducing the number of selection places by one from $m + 1$ to m . On the one hand, if there are at least $m + 1$ strong contestants in the pool, then reducing the number of selection places will lower designer welfare by 1, the cost of underselection, because, under merit-based selection, a strong contestant would have been selected had the quota equaled $m + 1$. The probability of this event equals $\mathbb{P}[\tilde{S}_n > m] = 1 - \mathbb{P}[\tilde{S}_n \leq m]$.

On the other hand, if there are less than $m + 1$ strong contestants, lowering the quota will increase designer welfare by σ , the cost of overselection, because, had the quota equaled $m + 1$, the designer would have been forced to fill the $(m + 1)$ th place with a weak contestant. The probability of this event equals $\mathbb{P}[\tilde{S}_n \leq m]$. Therefore, the marginal gain from reducing the quota from $m + 1$ to m is given by

$$\Delta(m) = \sigma \times \mathbb{P}[\tilde{S}_n \leq m] - 1 \times (1 - \mathbb{P}[\tilde{S}_n \leq m]) = (1 + \sigma)B(m; n, \theta) - 1, \quad (15)$$

where $B(\cdot; n, \theta)$ denotes the CDF of the Binom(n, θ) distribution. Thus, when $\Delta(m) > 0$, it

is strictly optimal to reduce the selection quota from $m + 1$ to m . Note that Δ is increasing in m ; thus, there exists a smallest $m \geq 0$ at which $\Delta(m) > 0$, which we call m_M^* . At $m = m_M^*$ it is strictly optimal to reduce the quota from $m_M^* + 1$ to m_M^* . Any further reduction in the quota will not increase designer welfare because, by definition, m_M^* is the smallest m that satisfies $\Delta > 0$. Thus, setting $m = m_M^*$ is an optimal policy. It might also be the case that at $m = m_M^* - 1$, $\Delta = 0$. If this occurs, then the designer's welfare is the same at $m = m_M^*$ and at $m = m_M^* - 1$, in which case, both m_M^* and $m_M^* - 1$ are optimal quotas. If, in contrast, at $m = m_M^* - 1$, $\Delta(m) < 0$, then m_M^* is the unique optimal quota under merit-based selection. We formalize this discussion with the following definitions:

$$m_M^*(n, \theta, \sigma) = \min\{m \in \{0, 1, \dots, n\} : B(m; n, \theta) > 1/(1+\sigma)\} \quad (16)$$

$$\hat{m}_M^*(n, \theta, \sigma) = \min\{m \in \{0, 1, \dots, n\} : B(m; n, \theta) \geq 1/(1+\sigma)\}. \quad (17)$$

If $m_M^* = \hat{m}_M^*$, the optimal quota under merit-based selection is unique. For any fixed n and $\sigma > 0$, $m_M^* \neq \hat{m}_M^*$ only for a finite set of probabilities, θ . So, generically, the optimal quota under merit-based selection is unique. In the non-generic case where $m_M^* \neq \hat{m}_M^*$, $\hat{m}_M^* = m_M^* - 1$. This case represents the case of two optimal quotas discussed above. For expositional convenience, in the subsequent analysis, we assume that, in the non-generic case, the designer sets $m = m_M^*$, and call m_M^* the *merit-based selection quota*. The next result shows that the merit-based selection quota, m_M^* , is nonincreasing in σ , the designer's overselection aversion, and is equal to median number of strong contestants if $\sigma = 1$.

Lemma 4. *The merit-selection quota, m_M^* , given by equation (16), is nonincreasing in σ . If $\sigma = 1$, m_M^* equals median number of strong contestants (i.e., median of the Binom(n, θ) distribution).¹⁴*

Intuitively, increasing σ makes the designer more averse to selecting weak contestants. This tends to induce her to reduce the merit-based selection quota. When $\sigma = 1$, i.e., when the marginal gain from selecting a strong contestant equals the marginal cost of selecting a weak contestant, the merit-based selection quota is set so that the marginal contestant deselected has a probability of being strong less than $1/2$.

If $\hat{m}_M^* = 0$, then deselecting all contestants is an optimal policy even under merit-based selection. If $m_M^* = n$, then selecting all contestants is an optimal policy under merit-based selection. The examination of the effects of risk-taking on contests when, even absent risk-taking, running a contest is not optimal, is not a very interesting exercise. Thus, henceforth, we impose the following restriction.

Assumption 1. n , θ , and σ satisfy the condition that $0 < \hat{m}_M^*$ and $m_M^* < n$, where m_M^* and \hat{m}_M^* are given by equations (16) and (17), respectively.

Now consider the designer's optimal quota when selection is determined by the risk-taking contest. Lemma 3 implies that, for any selection quota, the designer's welfare, $u(m)$, is bounded

¹⁴In the non-generic case, m_M^* equals the larger of the two medians.

above by her welfare under merit-based selection, $u_M(m)$, and thus, *a fortiori*, by her welfare under optimal merit-based selection, $u_M(m_M^*)$. When does designer welfare in the risk-taking contest attain this bound, $u_M(m_M^*)$, and how does contestant risk-taking affect the designer's optimal selection quota? The following proposition answers these questions.

Theorem 2. *Let m^* be the optimal selection quota in the risk-taking contest and let m_M^* be the merit-based selection quota. The following results hold under Assumption 1:*

- i. *Designer welfare in the risk-taking contest equals designer welfare under merit-based selection, i.e., $u(m^*) = u_M(m_M^*)$, if and only if the concession configuration is played in the risk-taking contest at $m = m_M^*$; otherwise, designer welfare is lower in the risk-taking contest, i.e., $u(m^*) < u_M(m_M^*)$.*
- ii. *If $r \leq \sigma(1-\theta)/\theta$ (in which case, the challenge configuration will be played at $m = m_M^*$), then $m^* = 0$, implying that deselecting all candidates is optimal.*
- iii. *If $r > \sigma(1-\theta)/\theta$, then (a) if the concession configuration is played at $m = m_M^*$, $m^* = m_M^*$, whereas (b) if the challenge configuration is played at $m = m_M^*$, $m^* \geq m_M^*$ and $m^* = \bar{m}$ or $\bar{m} + 1$, where \bar{m} is the largest quota at which the challenge configuration is played, i.e.,*

$$\bar{m} = \max \left\{ m \in \{m_M^*, \dots, n-1\} : p_W^C(m) < p_W^G(m) \right\}, \quad (18)$$

where p_W^C and p_W^G are defined by equations (4) and (9) respectively.

The logic behind Theorem 2 is fairly straightforward. Part (i) asserts that the necessary and sufficient condition for designer welfare to attain its merit-based upper bound is that the concession configuration is played at the merit-based selection quota, m_M^* . This is because, if the challenge configuration is played at m_M^* , designer welfare will fall below its merit-based upper bound either due to the potential misalignment between ability ranking and performance ranking caused by weak contestants' high-risk strategies or due to the designer's use of a "distorted" quota to accommodate risk-taking (or even both).

Part (ii) of Theorem 2 shows that, if the strength asymmetry between strong and weak contestants, r , is smaller than the threshold, $\sigma(1-\theta)/\theta$, it is optimal to deselect all contestants, even though doing so is not optimal under merit-based selection. This result follows because, when weak contestants are only marginally weaker than strong contestants, weak contestants will not concede. In this case, only the challenge configuration sustains an equilibrium. Inspection of equation (14) shows that, when $r \leq \sigma(1-\theta)/\theta$, a selected contestant's probability of being strong under the challenge configuration is less than $\sigma/(1+\sigma)$. Hence, each contestant selected lowers the designer's welfare and it is optimal to cancel the contest by setting a zero quota. Thus, selection through risk-taking contests will not be implemented when the relation between contest ability and ability, measured by r , is sufficiently weak.

Note that the condition in part (ii) highlights a fundamental difference between selection in risk-taking contests and merit-based selection: the merit-based selection quota is strictly positive when the contestant pool, n , is sufficiently large. In contrast, the condition for canceling

the contest by setting a zero quota in part (ii) is independent of n and, for any fixed level of strength asymmetry, is always satisfied for sufficiently small ex ante probability that contestants are strong, θ . Thus, part (ii) of the theorem implies that, when the designer is faced with a pool of agents with low average prior quality, the designer has no incentive to run a competition for selection/assignment regardless of the size of pool of contestants the designer can tap. Risk-taking blocks using selective contests to identify a few high ability agents hidden in a large pool of weak candidates.

Part (iii) of Theorem 2 shows that, if the strength asymmetry, r , is larger than the threshold, $\sigma(1-\theta)/\theta$, and if the challenge configuration is played at the merit-based selection quota, m_M^* , contestant risk-taking can induce *quota inflation*, setting selection quotas greater than the merit-based selection quota. Quota inflation results from two effects: the “strong loser” and “risk mollification” effects.

The strong-loser effect is a consequence of the imperfect relationship between selection and contest performance in challenge equilibria. This imperfect relationship ensures that the expected quality of selected contestants is lower and, by the same token, the average quality of deselected contestants (“losers”) is higher than under merit-based selection. Higher loser quality encourages the designer to dip deeper into the contestant pool.¹⁵

The second effect is “risk-mollification.” A larger quota can mollify weak contestants’ risk-taking incentives.¹⁶ The reduction in risk-taking makes contest performance a better reflection of contestant ability, which benefits the meritocratic designer and thus encourages increasing the selection quota. These effects are illustrated by the following example.

Example 3. Consider a contest with $n = 8$ contestants. The designer’s overselection aversion coefficient is $\sigma = 1$. Thus, over and underselection have the same cost to the designer and the designer’s objective is to maximize the difference between the number of strong and weak selected contestants. In this example, $\theta = 1/2$, so, ex ante, contestants are equally likely to be weak or strong. The contest ability of strong contestants is $\mu_S = 2$ while the contest ability of weak contestant is $\mu_W = 1$. Designer welfare under merit-based selection and in the risk-taking contest are presented in Table 1.

Consistent with Lemma 4, the merit-based selection quota, m_M^* , equals 4, the median of $\text{Binom}(n = 8, \theta = 1/2)$ distribution. Merit-based selection rather efficiently plucks strong contestants out of the pool, leading to few strong losers at the merit-based selection quota. To see this, note that the expected number of strong losers equals the expected number of strong contestants less the expected number of selected strong contestants. The expected number of

¹⁵The condition in part (iii) of Theorem 2 implies that $\theta > \sigma/(r+\sigma)$. Thus, under this condition, the average prior quality of contestants is not too low. If this condition is violated, the high-risk strategies played by weak contestants will result in a “weak winner” rather than “strong loser” effect, in which case full deselection will be optimal, as shown by part (ii) of Theorem 2.

¹⁶Of course, changing the quota can stimulate or mollify risk-taking incentives even when changing the quota does not change the configuration played. However, the contest designer is only concerned with mollifying risk-taking when it affects selection and thus we restrict the term “risk-mollification” to reducing risk-taking that affects selection.

strong contestants in this example equals $\theta \times n = 0.5 \times 8 = 4$. Thus, using the numbers in Table 1, we see that, when the selection quota equals the merit-based selection quota, $m_M^* = 4$, under merit-based selection, the expected number of strong losers equals $4 - 3.453 = 0.547$ out of four deselected candidates.

In the risk-taking contest, the challenge configuration is played for $m \leq \bar{m} = 5$. In the challenge configuration, the odds of a strong versus a weak contestant being selected equal the strength asymmetry, $r = \mu_S/\mu_W = 2$. Thus, given that, in the example, ex ante contestants are equally likely to be strong or weak, the odds of selecting a strong contestant equal 2:1. Hence, an increase in quota will, on average, increase the number of strong contestants selected by $2/3$ and the number of weak contestants selected by $1/3$. Because the expected number of strong contestants added exceeds the expected number of weak contestants added, increasing the quota at least to $\bar{m} = 5$ is optimal. The $\bar{m} = 5$ selection quota is inflated, i.e., higher than the merit-based selection quota.

Quota inflation over the range of quotas where the challenge configuration is played results from the strong-loser effect. To see this, note that the expected number of strong selected contestants equals $2/3 \times m$ and thus the expected number of strong losers equals $\theta \times n - 2/3 \times m = 4 - 2/3 \times m$. Consequently, at the merit-based selection quota, $m_M^* = 4$, the expected number of strong losers equals, $1.33\bar{3}$, and is approximately 140% larger than the number of strong losers under merit-based selection.

Whether it is optimal to inflate the quota even more, from $m = \bar{m} = 5$ to $m = \bar{m} + 1 = 6$, involves a tradeoff whose resolution depends on the specific choice of model parameters. At $m = 5$, the challenge configuration is played; at $m = 6$, the concession configuration is played. The mollification of risk-taking in the concession configuration makes selection at $m = 6$ merit-based. This effect encourages inflating the quota further to $m = 6$. However, the $\bar{m} = 5$ selection quota already exceeds the merit-based quota. This implies that the additional place created by further increasing the quota is very likely to be filled by a weak contestant. In fact, in this example, the probability that a sixth quota place will be filled by a weak contestant is approximately 85%.¹⁷ This effect discourages further inflating the quota.

As Table 1 reveals, in this example, increasing the quota from $m = \bar{m} = 5$ to $m = \bar{m} + 1 = 6$ is, in fact, optimal. Thus, the meritocratic contest designer is willing to offer a place to a contestant who is very likely to be unworthy in order to mollify contestant risk-taking incentives and thereby, on net, further the goal of meritocratic selection. Further increases of the quota beyond six are clearly suboptimal because increasing the selection quota to $m = \bar{m} + 1 = 6$ eliminates the “assortativity” distortions caused by risk-taking and further increases in the quota will lead to marginal quota places being filled by contestants who are even more likely to be weak, and thus unworthy of selection.

¹⁷The probability that the sixth contestant selected is weak equals the probability that the number of strong contestants is less than or equal to five, i.e., $B(m = 5; n = 8, \theta = 1/2) \approx 0.85$.

m	<i>merit-based selection</i>			Configuration	<i>contest selection</i>		
	$\mathbb{E}[\# \text{ Selected}]$		welfare, u_M		$\mathbb{E}[\# \text{ Selected}]$		welfare, u
	S	W			S	W	
1	0.996	0.004	0.992		0.667	0.333	0.333
2	1.961	0.039	1.922		1.333	0.667	0.667
3	2.816	0.184	2.632	Challenge	2.000	1.000	1.000
4	3.453	0.547	2.906		2.667	1.333	1.333
5	3.816	1.184	2.632		3.333	1.667	1.667
6	3.961	2.039	1.922	Concession	3.961	2.039	1.922
7	3.996	3.004	0.992		3.996	3.004	0.992

Table 1: *The selection quota, m , and designer welfare in Example 3*

Figure 3 illustrates the relation between the strength asymmetry, r , and the optimal quota in the risk-taking contest, m^* , when $\sigma = 1$, $n = 10$, and $\theta = 0.5$. As the figure shows, the optimal quota is highly inflated for small r . As the strength asymmetry increases, m^* decreases and eventually converges to the merit-based selection quota, m_M^* . As the next proposition reveals, the convergence from above exhibited in Figure 3 is a general property of optimal contest selection quotas.

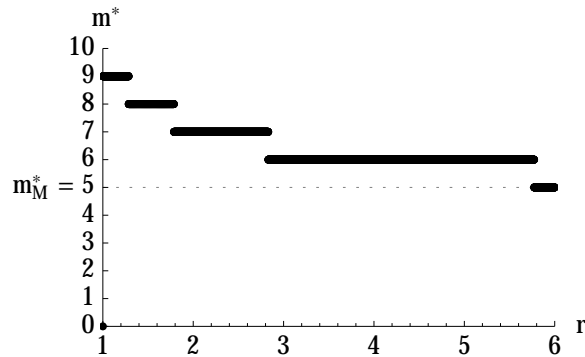


Figure 3: Optimal quota in the risk-taking contest, m^* , given strength asymmetry, r , when $\sigma = 1$, $n = 10$, and $\theta = 0.5$. The merit-based selection quota, m_M^* , equals 5.

Proposition 2. *Suppose that $r > \sigma(1-\theta)/\theta$ (otherwise, by Theorem 2, it is optimal to deselect all agents). The optimal quota in the risk-taking contest, m^* , is nonincreasing in strength asymmetry, r , and equal to m_M^* , the merit-based selection quota, for r sufficiently large.*

Proposition 2 implies that over-selection will be most pronounced in contests where contest ability is only marginally affected by (intrinsic) ability. This case is very likely to occur, as we will show in Section 6.2, where we endogenize contest ability through costly effort, when the cost of acquiring contest ability is highly convex.

6 Extensions

In this section, we consider various modifications of our baseline model. These extensions not only show that our results are quite robust but also lead to new implications.

6.1 Ex post discretionary selection and performance riskiness

In our baseline model, we assumed that the contest designer commits to fill the selection quota by best performers. Such commitment, however, is sometimes hard to enforce in practice, because contest performance sometimes depends on complex evaluations that are difficult for outsiders to verify.¹⁸ Thus, an important question to address is whether our contest game has an equilibrium in which the designer, even if ex post filling the quota at her discretion, has no incentive not to fill the quota by best performers.¹⁹ To answer this question, we need to first give a closer look at contestant risk-taking strategies, because whether better performance is a stronger signal of ability depends on the strategies played by the two types of contestants. The next result further characterizes equilibrium risk-taking strategies and shows that increasing competition can increase performance riskiness without benefiting meritocratic selection.

We measure performance riskiness by *mean-preserving spread* (MPS). It is well known that, if a distribution undergoes an MPS, the distribution has higher risks in the Rothschild and Stiglitz (1970) sense (second-order stochastic dominance).

Proposition 3. *Recall that F_t denotes the performance distribution chosen by a type- $t \in \{S, W\}$ contestant and Π denotes a selected contestant's probability of being strong. Let f_t be the density function for F_t , $t \in \{S, W\}$, and $F = \theta F_S + (1 - \theta) F_W$ be the unconditional performance distribution. The following results hold ceteris paribus:*

- i. *there exists n^c such that, for $n \geq n^c$, increasing contest size, n , has no effect on winner quality, Π , but makes the unconditional performance distribution, F , undergo an MPS.*
- ii. *If the challenge configuration is played at quota m^c , then for $m \leq m^c$, reducing the quota, m , has no effect on winner quality, Π , but makes the unconditional performance distribution, F , undergo an MPS.*
- iii. *There always exists an equilibrium in which F_S and F_W satisfy the monotone likelihood ratio property (MLRP), i.e.,*

$$\text{for every } x'' > x' \geq 0, \quad \frac{f_S(x'')}{f_W(x'')} \geq \frac{f_S(x')}{f_W(x')},$$

where we treat $f_S(x)/f_W(x) = +\infty$ if $f_W(x) = 0$.

¹⁸In contrast, the number of selected contestants is easy to verify in practice, which justifies the assumption that the designer can commit to her choice of the selection quota.

¹⁹If the designer had an incentive ex post to fill the quota by worse performers, rational contestants, who aim to maximize the probability of being selected, would fully anticipate the designer's ex post incentive and would accommodate to such an incentive by playing alternative strategies. If this happened, our previous welfare analysis, which was based on contestants' beliefs that only best performers are selected, would be problematic.

Parts (i) and (ii) of Proposition 3 imply that, if the contest is already very selective (such that the challenge configuration is played), making it even more selective by either expanding the contestant pool or reducing the selection places has no effect on winner quality but only increases performance riskiness. Thus, if performance riskiness *per se* causes direct social costs or if fierce competition imposes psychological costs on contestants, our result implies that, for competitions that are naturally fierce, e.g., elite-university admissions, reducing competition by, e.g., using a “relaxed” selection policy followed by a lottery process as proposed by Schwartz (2007), can reduce the side effect of selection contests without sacrificing meritocracy.

Part (iii) of Proposition 3 shows that there always exists an equilibrium in which the strong and weak types’ performance distributions satisfy the MLRP. The satisfaction of the MLRP implies that better performance is always a weakly better signal of ability. This is simply because, by Bayes rule, a contestant’s probability of being strong conditional on his performance level x is given by

$$\mathbb{P}[S|x] = \frac{\theta f_S(x)}{\theta f_S(x) + (1 - \theta) f_W(x)} = \frac{\theta (f_S(x)/f_W(x))}{\theta (f_S(x)/f_W(x)) + 1 - \theta},$$

which is non-decreasing in performance level x if the MLRP holds.

The satisfaction of the MLRP is obvious if the concession configuration is played in equilibrium, because under concession equilibria, strong contestants’ performance levels are uniformly higher than weak contestants’. If the challenge configuration is played in equilibrium, as discussed in Section 2.2, each type’s risk-taking strategy cannot be uniquely identified. However, as we show in Lemma A-4 in the Appendix, if the condition for challenge equilibria holds, one can always construct a challenge equilibrium in which the MLRP is satisfied. The next result is immediate from part (iii) of Proposition 3.

Corollary 1. *Filling the selection quota, m , by the m best performers is a credible commitment.*

6.2 Endogenous contest ability

In our baseline model, we assumed that the contest ability of each contestant is fixed and positively related to his (intrinsic) ability. In this subsection, we endogenize contest ability by allowing each contestant to acquire contest ability through costly effort. To do so, we assume that, after the selection quota and the size of the contestant pool are announced to the contestants, the contestants first simultaneously exert effort which determines their contest ability. We assume that the effort cost function is a strictly convex power function. Specifically, the cost of choosing contest ability μ for a type- $t \in \{S, W\}$ contestant is $c(\mu) = \frac{\mu^\alpha}{a_t}$, where $\alpha > 1$ and a_t is an ability parameter that satisfies $0 < a_W < a_S$. After the contestants acquire their contest ability, the contestants, without knowing each other’s contest ability, simultaneously choose nonnegative random performance subject to their contest-ability constraint. Selection is still based on the ranking of realized performance. Without loss of generality, we assume that the reward from being selected equals 1 and the reward from being deselected equals 0. A contestant’s payoff equals the reward he receives less his effort cost.

The next proposition shows that this modified game has symmetric equilibria in which contestants of the same type choose the same level of contest ability and play the same performance distribution. These equilibria still feature one of two configurations, concession or challenge, and the condition for each configuration to sustain an equilibrium is similar to our baseline model except that now the strength asymmetry, $r = \mu_S/\mu_W$, is endogenized.

Proposition 4. *Define p_W^C as in (4) and define $p_W^G(r)$ as in (9) and as a function of r , where $r = \mu_S/\mu_W$. The modified game in which contest ability is acquired through costly effort, with the cost function for type- $t \in \{S, W\}$ being $c(\mu) = \frac{\mu^\alpha}{a_t}$, where $\alpha > 1$ and $0 < a_W < a_S$, has either concession or challenge equilibria.*

i. *Concession equilibria exist if and only if*

$$p_W^C \geq p_W^G(r^*), \quad \text{where } r^* = \left(\frac{a_S}{a_W} \right)^{\frac{1}{\alpha-1}}. \quad (19)$$

ii. *If $p_W^C < p_W^G(r^*)$, challenge equilibria exist with endogenous strength asymmetry equal to r^* given in equation (19).*

iii. *Everything else being equal, if the challenge configuration is played in a contest with $\alpha = \alpha'$, the challenge configuration will also be played in contests with $\alpha > \alpha'$.*

In the concession configuration, the weak type's probability of winning equals p_W^C given by equation (4), which is independent of contest ability. In the challenge configuration, the weak type's probability of winning is given by $p_W^G(r^*)$, where r^* , given in equation (19), is endogenous strength asymmetry in the challenge configuration.²⁰ Thus, Proposition 4 implies that, in our modified game with endogenous contest ability, the equilibrium configuration is still determined by the one that gives the weak type a better chance of winning. Because r^* does not depend on contest size or selection quota, our previous analysis of how contest size and selection quota affect contestant risk-taking and how risk-taking in turn affects the design of selection contests is robust to the extension here.

Moreover, part (iii) of Proposition 4 leads to a new implication. Note that, α , the power coefficient of the effort cost function, measures effort cost convexity. Increasing α reduces r^* given in (19), and the reduction in r^* increases $p_W^G(r^*)$, making it less likely that condition (19) holds. Thus, consistent with part (iii), increasing effort cost convexity makes it less likely that weak contestants will concede to strong contestants and, conditional on that weak contestants challenge strong contestants, a reduction in strength asymmetry due to an increase in cost convexity will further increase weak contestants' chance of besting strong contestants. Thus, the performance/ability relation is most noisy if effort costs are highly convex, in which case, to reduce strategic noise, selection contests tend to be highly "clubby."²¹ As is well known that,

²⁰If condition (19) holds, then concession equilibria exist. In concession equilibria, endogenous strength asymmetry is different from r^* but does not enter p_W^C .

²¹If the power coefficient α tends to 1, in which case the cost function lacks convexity, r^* will tend to infinity and thus, by equation (9), $p_W^G(r^*)$ will tend to 0. In this case, by Proposition 4, weak contestants will always concede and, hence, risk-taking does not create any inefficiency.

in the mutual fund industry, it is hard for fund managers to generate “alpha,” i.e., risk-adjusted abnormal returns (Fama and French, 2010), this fact suggests that the cost of improving mean performance is highly convex for mutual fund managers. Thus, our result might offer a rational explanation for why retention contests in the mutual fund industry are highly clubby—only 14% of managers in the lowest performance decile are replaced (Tab. 4 Khorana, 1996).

6.3 Pool expansion by including less promising candidates

In Section 4, we studied the effect of risk-taking on the optimal size of the contestant pool. We showed that, if the contestant pool is sufficiently large, adding more contestants who are equally likely to exhibit ability as the contestants in the original pool does not affect the expected ability of contest winners. Thus, a meritocratic contest designer has no incentive to expand the contestant pool if the pool is already sufficiently large. The next result shows that, in fact, expanding a large pool hurts a meritocratic designer if the external candidates are inferior to the contestants in the original pool.

Proposition 5. *Suppose that the designer can only expand the contestant pool by including external candidates whose ex ante probability of being strong (measured by θ) is strictly lower than the internal candidates’. If the contest with only the internal candidates has challenge equilibria, pool expansion strictly reduces designer welfare.²²*

Proposition 5 implies that, even without any direct cost of pool expansion, as long as the external candidates are ex ante less promising than internal ones, a meritocratic designer strictly prefers “limiting the field” only to internal candidates if the internal competition already triggers the play of the challenge configuration. Proposition 5 might shed some light on why many real-world selection contests limit participation by requiring, sometimes in a *de facto* way, candidates to have certain qualifications to be eligible for contest participation.

6.4 Scoring caps

Many real-life contests have a scoring cap, such as a full score in examinations and many sports games, that caps the highest performance a contestant can possibly obtain. Even in cases in which performance is unbounded, a contest designer can impose a scoring cap if she can credibly specify that all performance levels no less than a threshold will be treated the same for the purpose of determining contest winners. Under this specification, that threshold will effectively be the scoring cap. The next proposition shows that imposing a scoring cap has no effect on any contestant’s probability of winning if the scoring cap is no less than the strong type’s contest ability, whereas if otherwise, imposing a scoring cap (weakly) benefits the weak type at the cost of the strong type.

²²Because, by assumption, the external candidates are *ex ante* different from the internal candidates, the concept of symmetric equilibria in such a case refers to equilibria in which all type- t internal candidates play the same strategy and all type- t external candidates play the same strategy, $t \in \{S, W\}$.

Proposition 6. *Imposing a scoring cap no less than μ_S does not affect any type's probability of winning but weakly reduces contestant risk-taking in the sense that the unconditional performance distribution without the cap equals or is an MPS of that with the cap. Imposing a scoring cap strictly less than μ_S weakly increases the weak type's probability of winning and weakly reduces the strong type's probability of winning.*

Proposition 6 implies that imposing a highly restrictive scoring cap, i.e., a scoring cap that is strictly less than the strong type's contest ability, weakly reduces selection efficiency. Thus, a designer who cares about selection has no incentive to impose a highly restrictive scoring cap. Proposition 6 also implies that imposing a not-so-restrictive scoring cap, i.e., a scoring cap that is no less than the strong type's contest ability, has no effect on selection efficiency while reduces contestant risk-taking. Thus, if a designer cares about selection but, at the same time, is averse to performance riskiness, she can reduce risk-taking without any side effect on selection by imposing a not-so-restrictive scoring cap. Because performance at the cap cannot be topped, placing point mass on the cap can be a rational choice for a contestant. In fact, as we show in the proof of Proposition 6, if the not-so-restrictive scoring cap is strictly lower than the upper endpoint of the union of the supports of weak and strong types' performance distributions chosen in the contest without the cap, imposing such a scoring cap will lead contestants to move any probability weight originally placed on performance levels above the cap and some weight below the cap to the cap, thereby reducing performance riskiness.²³

7 Conclusion

In this paper, we studied selection contests in which contestants of private types are strategic risk takers. We showed that increasing competition, either by expanding the contestant pool or reducing the selection quota, increases weak contestants' tendency to play high-risk strategies to challenge potential strong contestants, which limits the gains in selected applicant quality produced by intensifying competition. Consequently, even meritocratic designers have an incentive to limit competition by adopting "clubby" contests, contests that feature less inclusive contestant pools and over-promotion of marginal candidates. Our model implies that many seemingly anti-meritocratic practices and proposals are in fact in line with meritocracy, such as the use of "Peter Principle" promotion policies in companies and organizations (Peter and Hull, 1969), the running of "in-house" competition instead of "open competition" for leader selection, and the advocate of using a "relaxed" selection policy followed by a lottery process for elite-university admissions (Schwartz, 2007).

²³If the scoring cap is higher than the upper endpoint of the union of Supp_W and Supp_S in the contest without the cap, imposing the cap does not affect any contestant's strategy.

References

- Barmby, Tim, Barbara Eberth, and Ada Ma**, “Incentives, learning, task difficulty, and the Peter Principle: Interpreting individual output changes in an organisational hierarchy,” *Labour Economics*, 2012, 19 (1), 76–81.
- Barut, Yasar and Dan Kovenock**, “The symmetric multiple prize all-pay auction with complete information,” *European Journal of Political Economy*, 1998, 14 (4), 627–644.
- Baye, Michael R., Dan Kovenock, and Casper G. de Vries**, “Rigging the lobbying process: An application of the all-pay auction,” *American Economic Review*, 1993, 83 (1), 289–294.
- Becker, Gary S., Kevin M. Murphy, and Iván Werning**, “The equilibrium distribution of income and the market for status,” *Journal of Political Economy*, 2005, 113 (2), 282–310.
- Che, Yeon-Koo and Ian Gale**, “Optimal design of research contests,” *The American Economic Review*, 2003, 93 (3), 646–671.
- Chevalier, Judith and Glenn Ellison**, “Risk taking by mutual funds as a response to incentives,” *Journal of Political Economy*, 1997, 105 (6), 1167–1200.
- Costin, Caudia**, “Education lessons from South Korea,” November 2014. <https://www.weforum.org/agenda/2014/11/education-lessons-from-south-korea/>.
- Cover, Thomas M.**, *Universal gambling schemes and the complexity measures of Kolmogorov and Chaitin* Technical Report No. 12, Stanford Statistics Department, 1974.
- Deeprise, Donna**, *How to Recognize & Reward Employees*, American Management Association, 2006.
- Diamond, Peter A. and Joseph E. Stiglitz**, “Increases in Risk and in Risk Aversion,” *Journal of Economic Theory*, 1974, 8 (3), 337–360.
- Dillon, Eleanor Wiske and Jeffrey Andrew Smith**, “The determinants of mismatch between students and colleges,” 2013. National Bureau of Economic Research.
- Drugov, Mikhail and Dmitry Ryvkin**, “Tournament rewards and heavy tails,” *Mimeo*, 2018.
- Ekins, Emily**, “The Reason-Rupe Spring 2014 Millennial Survey,” June 2014. <http://reason.com/assets/db/2014-millennials-report.pdf>.
- Fairburn, James A. and James M. Malcomson**, “Performance, promotion, and the Peter Principle,” *The Review of Economic Studies*, 2001, 68 (1), 45–66.
- Fama, Eugene F and Kenneth R French**, “Luck versus skill in the cross-section of mutual fund returns,” *The Journal of Finance*, 2010, 65 (5), 1915–1947.

- Fang, Dawei and Thomas H. Noe**, “Skewing the odds: Taking risks for rank-based rewards,” *Mimeo*, 2016.
- , **Thomas Noe, and Philipp Strack**, “Turning Up the Heat: The Demoralizing Effect of Competition in Contests,” *Mimeo*, 2018.
- Frost, Stephen**, “How to build a meritocracy,” January 2017. <https://minutehack.com/guides/how-to-build-a-meritocracy>.
- Fu, Qiang and Jingfeng Lu**, “Micro foundations of multi-prize lottery contests: A perspective of noisy performance ranking,” *Social Choice and Welfare*, 2012, 38, 497–517.
- Gaba, Anil and Ajay Kalra**, “Risk behavior in response to quotas and contests,” *Marketing Science*, 1999, 18 (3), 417–434.
- , **Ilia Tsetlin, and Robert L. Winkler**, “Modifying Variability and Correlations in Winner-Take-All Contests,” *Operations Research*, 2004, 52 (3), 384–395.
- Ghanem, Dalia and Junjie Zhang**, “Effortless Perfection: Do Chinese cities manipulate air pollution data?,” *Journal of Environmental Economics and Management*, 2014, 68 (2), 203–225.
- Gilpatric, Scott M.**, “Risk Taking in Contests and the Role of Carrots and Sticks,” *Economic Inquiry*, 2009, 47 (2), 266–277.
- Gürtler, Oliver and Matthias Kräkel**, “Optimal tournament contracts for heterogeneous workers,” *Journal of Economic Behavior & Organization*, 2010, 75 (2), 180–191.
- Hart, Sergiu**, “Allocation games with caps: From Captain Lotto to all-pay auctions,” *International Journal of Game Theory*, 2015, pp. 1–25.
- Hess, Gregory D and Athanasios Orphanides**, “War politics: An economic, rational-voter framework,” *The American Economic Review*, 1995, pp. 828–846.
- Hvide, Hans K.**, “Tournament Rewards and Risk Taking,” *Journal of Labor Economics*, 2002, 20 (4), 877–898.
- **and Eirik G. Kristiansen**, “Risk taking in selection contests,” *Games and Economic Behavior*, 2003, 42 (1), 172–179.
- Jimerson, Shane R, Sarah MW Pletcher, Kelly Graydon, Britton L Schnurr, Amanda B Nickerson, and Deborah K Kundert**, “Beyond grade retention and social promotion: Promoting the social and academic competence of students,” *Psychology in the Schools*, 2006, 43 (1), 85–97.

- Jones, M. C.**, “The Complementary Beta Distribution,” *Journal of Statistical Planning and Inference*, 2002, 104 (2), 329–337.
- Kalra, Ajay and Mengze Shi**, “Designing optimal sales contests: A theoretical perspective,” *Marketing Science*, 2001, 20 (2), 170–193.
- Kamphorst, Jurjen J.A. and Otto H. Swank**, “Don’t demotivate, discriminate,” *American Economic Journal: Microeconomics*, 2016, 8 (1), 140–165.
- Kaniel, Ron and Robert Parham**, “WSJ Category Kings—The impact of media attention on consumer and mutual fund investment decisions,” *Journal of Financial Economics*, 2017, 123 (2), 337–356.
- Kawamura, Kohei and Inés Moreno de Barreda**, “Biasing Selection Contests with Ex-Ante Identical Agents,” *Economics Letters*, 2014, 123 (2), 240–243.
- Khorana, Ajay**, “Top management turnover: An empirical investigation of mutual fund managers,” *Journal of Financial Economics*, 1996, 40 (3), 403–429.
- , “Performance changes following top management turnover: Evidence from open-end mutual funds,” *Journal of Financial and Quantitative Analysis*, 2001, 36 (3), 371–393.
- Knoeber, Charles R. and Walter N. Thurman**, “Testing the theory of tournaments: An empirical analysis of broiler production,” *Journal of Labor Economics*, 1994, 12 (2), 155–179.
- Lazear, Edward P and Sherwin Rosen**, “Rank-order tournaments as optimum labor contracts,” 1979. National Bureau of Economic Research Cambridge, Mass., USA.
- Lazear, Edward P. and Sherwin Rosen**, “Rank-order tournaments as optimum labor contracts,” *Journal of Political Economy*, 1981, 89, 841–864.
- Lizzeri, Alessandro**, “Budget deficits and redistributive politics,” *Review of Economic Studies*, 1999, 66 (4), 909–928.
- Luenberger, David**, *Optimization by Vector Space Methods*, Wiley, 1969.
- Morgan, John and Felix Várdy**, “Diversity in the Workplace,” *American Economic Review*, 2009, 99 (1), 472–85.
- , **Dana Sisak, and Felix Várdy**, “The ponds dilemma,” *The Economic Journal*, 2015.
- Myerson, Roger B.**, “Incentives to cultivate favored minorities under alternative electoral systems,” *American Political Science Review*, 1993, 87 (04), 856–869.
- Nafziger, Julia**, “Motivational job assignments,” *Economica*, 2011, 78 (312), 676–696.

- Olszewski, Wojciech and Ron Siegel**, “Pareto Improvements in the Contest for College Admissions,” *Mimeo*, 2018.
- Peter, Laurence J and Raymond Hull**, *The Peter Principle*, Souvenir Press London, 1969.
- Piketty, Thomas**, “Top income shares in the long run: An overview,” *Journal of the European Economic Association*, 2005, 3 (2-3), 382–392.
- Prendergast, Canice**, “The insurance effect of groups,” *International Economic Review*, 1992, 33 (3), 567–581.
- Ray, Debraj and Arthur Robson**, “Status, Intertemporal Choice, and Risk-Taking,” *Econometrica*, 2012, 80 (4), 1505–1531.
- Robson, Arthur J.**, “Status, the Distribution of Wealth, Private and Social Attitudes to Risk,” *Econometrica*, 1992, 60 (4), 837–857.
- Rothschild, Michael and Joseph Stiglitz**, “Increasing Risk: I. A Definition,” *Journal of Economic Theory*, 1970, 2, 225–243.
- Rykin, Dmitry**, “The Selection Efficiency of Tournaments,” *European Journal of Operational Research*, 2010, 206 (3), 667–675.
- **and Andreas Ortmann**, “The Predictive Power of Three Prominent Tournament Formats,” *Management Science*, 2008, 54 (3), 492–504.
- Schwartz, Barry**, “Why the best schools can’t pick the best kids — and vice versa,” March 2007. <http://articles.latimes.com/2007/mar/18/opinion/op-schwartz18/2>.
- Seel, Christian and Philipp Strack**, “Gambling in contests,” *Journal of Economic Theory*, 2013, 148 (5), 2033–2048.
- Serrato, Juan Carlos Suárez, Xiao Yu Wang, and Shuang Zhang**, “The limits of meritocracy: screening bureaucrats under imperfect verifiability,” Technical Report, National Bureau of Economic Research 2016.
- Shleifer, Andrei**, “Does competition destroy ethical behavior?,” *American Economic Review*, 2004, 94 (2), 414–418.
- Strack, Philipp**, “Risk-taking in contests: The impact of fund-manager compensation on investor welfare,” *Mimeo*, 2016.
- Taylor, Jonathan**, “Risk-Taking Behavior in Mutual Fund Tournaments,” *Journal of Economic Behavior & Organization*, 2003, 50 (3), 373–383.

Wadsworth, George P and Joseph G Bryan, *Introduction to Probability and Random Variables*, McGraw-Hill, 1960.

Weinstein, Deena, “Fraud in science,” *Social Science Quarterly*, 1979, 59 (4), 639–652.

Wooldridge, Adrian, “Ever higher society, ever harder to ascend,” December 2005.
<https://www.economist.com/node/3518560>.

Appendix A: Proofs of results

Proof of Lemma 1. In the main text, we provided an informal but intuitive discussion of the proof of Lemma 1. In what follows, we give a formal proof by using duality theory. The proof consists of several steps.

Step 1: The first step, which is the key, is to develop an algorithm that can be used to characterize each type's best reply in a symmetric equilibrium. The algorithm is given in Lemma A-1. Note that, for every type- $t \in \{S, W\}$ contestant, his problem is to choose a performance distribution, F_t , for his nonnegative random performance, X_t , so as to maximize $\mathbb{E}[P(X_t)]$, subject to the contest ability constraint, i.e., $\mathbb{E}(X_t) = \mu_t$. More conveniently, we can formulate the problem as one of choosing a performance measure, dF_t , to use against P . The performance measure, dF_t , has to satisfy two constraints: (a) it has to be a probability measure and (b) its expectation equals μ_t . The solution to this problem coincides with the solution to the following relaxed problem of choosing a measure dF_t supported by $[0, \infty)$:

$$\max_{dF_t \geq 0} \int_{0-}^{\infty} P(x) dF_t(x) \quad \text{s.t. (i) } \int_{0-}^{\infty} dF_t(x) \leq 1 \text{ \& (ii) } \int_{0-}^{\infty} x dF_t(x) \leq \mu_t. \quad (\text{P}_F)$$

The Lagrangian associated with problem (P_F) is given by

$$\mathcal{L}(dF_t, \alpha_t, \beta_t) = \int_{0-}^{\infty} P(x) dF_t(x) - \alpha_t \left(\int_{0-}^{\infty} dF_t(x) - 1 \right) - \beta_t \left(\int_{0-}^{\infty} x dF_t(x) - \mu_t \right), \quad (\text{A-1})$$

where α_t and β_t are nonnegative dual variables. Rewrite equation (A-1) as

$$\mathcal{L}(dF_t, \alpha_t, \beta_t) = \int_{0-}^{\infty} [P(x) - (\alpha_t + \beta_t x)] dF_t(x) + \alpha_t + \beta_t \mu_t. \quad (\text{A-2})$$

Our next result shows the existence of a solution to problem (P_F).

Result A-1. Suppose that P is nonnegative, nondecreasing, bounded, and upper semicontinuous, with $P(\mu_t) < P(\infty)$ (which we will show to be the case in any symmetric equilibrium). Problem (P_F) has a solution and the support of this solution is bounded.

Proof. Define

$$v_t^* = \sup \left\{ \int_{0-}^{\infty} P(x) dF_t(x) : \int_{0-}^{\infty} dF_t(x) \leq 1 \text{ \& } \int_{0-}^{\infty} x dF_t(x) \leq \mu_t \right\}. \quad (\text{A-3})$$

Because P is bounded and the contest ability constraint and the unit mass constraint are not mutually exclusive, v_t^* clearly exists. The map $dF_t \mapsto \int_{0-}^{\infty} P(x) dF_t(x)$ is linear. The set of nonnegative measures satisfying the contest ability and the unit mass constraints is convex. Thus, by basic duality theory (Luenberger, 1969, §8.3, Theorem 1), there exist $\alpha_t^o \geq 0$ and $\beta_t^o \geq 0$ such that

$$\sup_{dF_t \geq 0} \mathcal{L}(dF_t, \alpha_t^o, \beta_t^o) = v_t^*.^{24} \quad (\text{A-4})$$

²⁴The other conditions for the existence of the nonnegative dual variables are clearly satisfied as the constraint space is simply \mathfrak{R}^2 and thus its nonnegative cone has a nonempty interior and the feasible set contains a point where the contest ability and the unit mass constraints are strictly satisfied.

Because P is nondecreasing and $P(\mu_t) < P(\infty)$, increased contest ability has value. Hence, the contest ability constraint cannot be slack, which implies that $\beta_t^o > 0$. Because P is bounded, $P(\infty)$ exists. Thus, given that $\alpha_t^o \geq 0$ and $P(x) \leq P(\infty)$ for all $x \geq 0$, we must have that, for all $x > x^* = P(\infty)/\beta_t^o$, $P(x) - (\alpha_t^o + \beta_t^o x) < 0$. Inspecting equation (A-2) then shows that placing positive weight on any performance level $x \in (x^*, \infty)$ lowers the Lagrangian, \mathcal{L} . Hence, restricting the probability measure to $[0, x^*]$ will not lower its supremum. Thus,

$$\sup\{\mathcal{L}(dF_t, \alpha_t^o, \beta_t^o) : dF_t \geq 0\} = \sup\{\mathcal{L}(dF_t, \alpha_t^o, \beta_t^o) : dF_t \geq 0 \ \& \ dF_t\{(x^*, \infty)\} = 0\}. \quad (\text{A-5})$$

The set of measures with support in $[0, x^*]$ is compact in the weak topology and P is upper semicontinuous. Thus, the supremum of the Lagrangian is attained over the restricted set of measures, i.e., there exists dF_t^o such that

$$\mathcal{L}(dF_t^o, \alpha_t^o, \beta_t^o) = \sup\{\mathcal{L}(dF_t, \alpha_t^o, \beta_t^o) : dF_t \geq 0 \ \& \ dF_t\{(x^*, \infty)\} = 0\}. \quad (\text{A-6})$$

Thus, equations (A-3), (A-4), (A-5), (A-6), and basic duality theory (Luenberger, 1969, §8.4, Theorem 1) imply that dF_t^o solves problem (P_F). \square

The next lemma presents the algorithm that can be used to characterize the solution to problem (P_F).

Lemma A-1. *Suppose that P satisfies all the conditions in Result A-1. A probability distribution function, F_t , solving problem (P_F) exists. For any such solution, its support is bounded and there exist dual variables $\alpha_t \geq 0$ and $\beta_t > 0$ such that α_t and β_t satisfy*

$$P(x) \leq \alpha_t + \beta_t x \quad \forall x \geq 0 \quad \& \quad dF_t\{x \geq 0 : P(x) < \alpha_t + \beta_t x\} = 0, \quad (\text{A-7})$$

and, if $\mathbf{v}(P, \mu_t)$ represents the optimal value of problem (P_F),

$$\mathbf{v}(P, \mu_t) = \alpha_t + \beta_t \mu_t. \quad (\text{A-8})$$

Conversely, if a probability distribution, F_t , satisfies (A-7) and makes the contest ability constraint, (P_F-ii), bind, it is a solution to (P_F).

Proof. The dual variables associated with an optimal solution are the solutions to the following dual problem:

$$\min_{\alpha_t, \beta_t \geq 0} \sup_{dF_t \geq 0} \mathcal{L}(dF_t, \alpha_t, \beta_t). \quad (\text{D}_F)$$

Equation (A-2) implies that the optimal dual variables that solve (D_F) must satisfy

$$P(x) - (\alpha_t + \beta_t x) \leq 0 \quad \forall x \geq 0, \quad (\text{A-9})$$

since otherwise $\sup_{dF_t \geq 0} \mathcal{L}(dF_t, \alpha_t, \beta_t)$ tends to positive infinity. Thus, $\alpha_t + \beta_t x$ is an upper bound for $P(x)$.

When condition (A-9) is satisfied, the value of the dual problem (D_F) equals

$$\alpha_t + \beta_t \mu_t,$$

which is strictly increasing in both α_t and β_t . Thus, the nonnegative optimal dual variables must

minimize $\alpha_t + \beta_t \mu_t$ subject to condition (A-9). Hence, given that P is upper semicontinuous, condition (A-9) must be binding for some $x \geq 0$, i.e., there exists some point(s) $x' \geq 0$ such that $P(x') - (\alpha_t + \beta_t x') = 0$. Thus, $\alpha_t + \beta_t x$ is an *upper support line* for P . To distinguish this line from other upper support lines that P might have, we call this line the *type- t 's upper support line*. Placing any probability weight on points at which $P(x) - (\alpha_t + \beta_t x) < 0$ lowers the Lagrangian. Thus, the optimal performance distribution for type- t must place no weight on such points. Therefore, the optimal performance measure for type- t is always concentrated on points at which the type- t 's support line, $\alpha_t + \beta_t x$, meets the contest payoff function, P . Thus, the optimal performance measure for type- t , dF_t , and the associated optimal dual variables, α_t and β_t , must satisfy (A-7).

Note that the maps $dF_t \mapsto \int_{0-}^{\infty} dF_t(x)$ and $dF_t \mapsto \int_{0-}^{\infty} x dF_t(x)$ are linear. Thus, the set of non-negative measures satisfying the unit mass constraint, (P_F-i), and the contest ability constraint, (P_F-ii), is convex. Given that the map $dF_t \mapsto \int_{0-}^{\infty} P(x) dF_t(x)$ is linear, by basic duality theory (Luenberger, 1969, §8.6, Theorem 1), strong duality holds. Thus, the primal problem (P_F) must have its optimal value equal to that of the dual problem (D_F). Hence, the optimal value of problem (P_F) satisfies equation (A-8). Relaxing the unit mass constraint, (P_F-i), by $\varepsilon > 0$ increases the type- t contestant's payoff by at least $\varepsilon \times P(0) \geq 0$. Thus, a solution to problem (P_F) in which the unit mass constraint is satisfied as an equality (i.e., the optimal measure is a probability measure) always exists and α_t , the dual variable associated with the unit mass constraint, is at least equal to 0. Similarly, since $P(\mu_t) < P(\infty)$, the type- t contestant does not have sufficient contest ability to guarantee the largest possible payoff. Thus, the contest ability constraint, (P_F-ii), must be binding at the optimum and hence, $\beta_t > 0$. \square

Step 2: The second step is to argue that

Result A-2. In any symmetric equilibrium, the contest payoff function, P , is nonnegative, nondecreasing, bounded, and continuous, with $P(\mu_t) < P(\infty) = 1$ and $P(0) = 0$.

Proof. Continuity and the result that $P(0) = 0$ follow from the argument in the main text that no one places point mass in any symmetric equilibrium. The result that $P(\mu_t) < 1$ follows from the fact that no contestant ensures winning in any symmetric equilibrium. The rest is obvious. \square

Because continuity implies upper semicontinuity, Result A-2 implies that, in any symmetric equilibrium, P satisfies all the conditions in Lemma A-1. This enables us to apply Lemma A-1. Lemma A-1 implies that the optimal performance measure for type- t is always concentrated on points at which the type- t 's upper support line, $\alpha_t + \beta_t x$, meets the contest payoff function, P . Because P is continuous, the set of points at which P meets the upper support line must be closed. Thus, given that supports of distributions are by definition closed, we must have

$$\forall t \in \{S, W\}, \quad \text{Supp}_t \subset \{x \geq 0 : P(x) = \alpha_t + \beta_t x\}, \quad (\text{A-10})$$

where Supp_t represents the support of F_t , the performance distribution selected by type t .

Define ψ as the *concave lower envelope* of the two upper support lines, $\{\alpha_t + \beta_t x\}_{t=S,W}$, associated with the two types of contestants, i.e.,

$$\psi(x) = \min[\alpha_S + \beta_S x, \alpha_W + \beta_W x]. \quad (\text{A-11})$$

By Lemma A-1 and the definition of the concave lower envelope, ψ , we must have

$$\forall t \in \{S, W\}, \quad \alpha_t + \beta_t x \geq \psi(x) \geq P(x). \quad (\text{A-12})$$

Step 3: The third step in the proof is to show that

Lemma A-2. *Define $\psi(x)$ as in (A-11). There exist constants $\alpha_t \geq 0$ and $\beta_t > 0$, $t \in \{S, W\}$, such that, in any symmetric equilibrium, $P(x) = \psi(x)$ over its support $[0, \hat{x}]$ ($P(x) = 1$ for $x > \hat{x}$), where \hat{x} is implicitly determined by*

$$\psi(\hat{x}) = 1. \quad (\text{A-13})$$

Proof. First, by Lemma A-1, the type- t contestant's expected payoff by playing a best reply to P is $\alpha_t + \beta_t \mu_t$, which, in any symmetric equilibrium, must be strictly less than one. Also, by Lemma A-1, $\beta_t > 0$. Thus, \hat{x} , defined implicitly by (A-13), must exist and $\psi(x) > 1$ for $x > \hat{x}$.

Next, note that $P(x) \leq 1$ for all $x \geq 0$. Thus, for $x > \hat{x}$, $P(x) < \psi(x)$. This implies, by (A-12), that $P(x) < \alpha_t + \beta_t x$ for $x > \hat{x}$. Thus, by (A-10), no contestant places any weight over (\hat{x}, ∞) . This further implies, given that no one places point mass in any symmetric equilibrium, that a contestant ensures winning if his performance equals \hat{x} , i.e., $P(\hat{x}) = 1$. Thus, $P(\hat{x}) = 1 = \psi(\hat{x})$ and $P(x) = 1 < \psi(x)$ for $x > \hat{x}$.

Finally, to show that $P(x) = \psi(x)$ for $x \leq \hat{x}$, we establish the following technical result, which will also be used in a later proof:

Result A-3. In a symmetric equilibrium, for any $a > 0$, if P is continuous over $[0, a]$ and if there exists $x' \in (0, a)$ such that $P(x') = \psi(x')$, then it must be that $P(x) = \psi(x)$ for all $x \in [0, x']$.

Proof. We prove the result by way of contradiction. Let $S = \{x \in [0, x'] : P(x) \neq \psi(x)\}$. Suppose, contrary to the result, that $S \neq \emptyset$. Then S must contain a point, say x_o . Let $x_1 = \min\{x \in (x_o, x'] : P(x) = \psi(x)\}$. Because, by hypothesis, $P(x') = \psi(x')$ and P is continuous for $x \leq x'$, x_1 is well defined. By the definition of x_1 and equation (A-12), $P(x) < \psi(x)$ for all $x \in (x_o, x_1)$. Because all contestants face the same P in a symmetric equilibrium, by (A-10) and (A-12), no contestant places any weight on (x_o, x_1) . Thus, $P(x) = P(x_o)$ for all $x \in (x_o, x_1)$. Thus, by continuity of P for $x \leq x'$, $P(x_o) = P(x_1)$. However, because ψ is strictly increasing, $\psi(x_o) < \psi(x_1)$. Thus, given that $P(x_o) \leq \psi(x_o)$ and $P(x_o) = P(x_1)$, we must have $P(x_1) < \psi(x_1)$, which contradicts the definition of x_1 . Thus, $S = \emptyset$ and the result follows. \square

Because, by Result A-2, P is continuous in any symmetric equilibrium and because, as has been proved, $P(\hat{x}) = \psi(\hat{x})$, Result A-3 implies that $P(x) = \psi(x)$ for $x \leq \hat{x}$. This completes the proof of Lemma A-2. \square

Step 4: The last step is to use Lemma A-2 to complete the proof. By Lemma A-2, $P(x) = \psi(x)$ over P 's support $[0, \hat{x}]$. Because ψ , given by (A-11), is continuous, increasing, concave, and is either linear or piecewise linear with one kink, it must be that, over its support $[0, \hat{x}]$, P is continuous, increasing, concave, and is either linear or piecewise linear with one kink. The fact that $P(0) = 0$ follows from Result A-2. This completes the proof of Lemma 1. \square

Proof of Proposition 1. The “only if” part and each type’s equilibrium payoff are both established in the main text.

Below we establish the “if” part by constructing the symmetric equilibrium. Note that, in equilibrium, no one places any point mass. Thus, given a contestant’s performance level x , the contestant’s probability of besting a given rival of unknown type equals $F(x)$, where

$$F(x) = \theta F_S(x) + (1 - \theta) F_W(x) \quad (\text{A-14})$$

is the unconditional performance distribution. To win a selection place, a contestant has to best at least $(n - m)$ out of his $(n - 1)$ rivals. Thus, in any symmetric equilibrium, a contestant’s probability of winning function, P , and the unconditional performance distribution, F , satisfy that

$$P(x) = \sum_{i=n-m}^{n-1} \binom{n-1}{i} F(x)^i (1 - F(x))^{n-1-i}. \quad (\text{A-15})$$

The next lemma shows the existence of a concession equilibrium when $p_W^C \geq p_W^G$, and gives a construction of P and F_t , $t \in \{S, W\}$, in the concession equilibrium.

Lemma A-3. *When $p_W^C \geq p_W^G$, there exists a unique concession equilibrium. In this equilibrium, the probability of winning function, P , is given by*

$$P(x) = \begin{cases} \beta_W x, & x \in [0, \tilde{x}] \\ \alpha_S + \beta_S x, & x \in [\tilde{x}, \hat{x}] \\ 1, & x \geq \hat{x} \end{cases}, \quad (\text{A-16})$$

where β_W , \tilde{x} , β_S , α_S , and \hat{x} are determined by contest parameters as follows:

$$\beta_W = \frac{p_W^C}{\mu_W} \quad (\text{A-17})$$

$$\tilde{x} = \frac{\tilde{p} \mu_W}{p_W^C} \quad (\text{A-18})$$

$$\beta_S = \frac{p_S^C - \tilde{p}}{\mu_S - (\tilde{p} \mu_W / p_W^C)} \quad (\text{A-19})$$

$$\alpha_S = \frac{\tilde{p} (\mu_S - (p_S^C \mu_W / p_W^C))}{\mu_S - (\tilde{p} \mu_W / p_W^C)} \quad (\text{A-20})$$

$$\hat{x} = \frac{(1 - \tilde{p}) \mu_S - (1 - p_S^C) (\tilde{p} \mu_W / p_W^C)}{p_S^C - \tilde{p}} \quad (\text{A-21})$$

with p_W^C given by (4), p_S^C determined by p_W^C through equation (6), and \tilde{p} given by

$$\tilde{p} = \sum_{i=n-m}^{n-1} \binom{n-1}{i} (1-\theta)^i \theta^{n-1-i}. \quad (\text{A-22})$$

The constants, α_S , β_W , and β_S , given by (A-20), (A-17), and (A-19), respectively, satisfy the following: if $p_W^C = p_W^G$, then $\alpha_S = 0$ and $\beta_W = \beta_S > 0$. If $p_W^C > p_W^G$, then $\alpha_S > 0$ and $\beta_W > \beta_S > 0$.

In the concession equilibrium, the weak type's performance distribution, F_W , is supported by $[0, \tilde{x}]$, the strong type's performance distribution, F_S , is supported by $[\tilde{x}, \hat{x}]$, and F_W and F_S are given implicitly by

$$\sum_{i=n-m}^{n-1} \binom{n-1}{i} [(1-\theta)F_W(x)]^i [1 - (1-\theta)F_W(x)]^{n-1-i} = \beta_W x, \quad x \in [0, \tilde{x}] \quad (\text{A-23})$$

$$\sum_{i=n-m}^{n-1} \binom{n-1}{i} [1 - \theta + \theta F_S(x)]^i [\theta(1 - F_S(x))]^{n-1-i} = \alpha_S + \beta_S x, \quad x \in [\tilde{x}, \hat{x}], \quad (\text{A-24})$$

where β_W , \tilde{x} , α_S , β_S , and \hat{x} are given by equations (A-17)–(A-21), respectively.

Proof. The proof consists of several steps.

Step 1: We first show that, in the concession configuration, P must have the structure given by (A-16). Note that, in the concession configuration, $\text{Supp}_W = [0, \tilde{x}]$ and $\text{Supp}_S = [\tilde{x}, \hat{x}]$, where \tilde{x} represents the upper endpoint of Supp_W , which is also the lower endpoint of Supp_S in the concession configuration. Thus, by (A-10), in the concession configuration, it must be that $P(x) = \alpha_W + \beta_W x$ for $x \in [0, \tilde{x}]$ and $P(x) = \alpha_S + \beta_S x$ for $x \in [\tilde{x}, \hat{x}]$. Because, by Lemma 1, $P(0) = 0$. Thus, given that $P(x) = \alpha_W + \beta_W x$ for $x \in [0, \tilde{x}]$, it must be that $\alpha_W = 0$ and, hence, $P(x) = \beta_W x$ for $x \in [0, \tilde{x}]$. Thus, in the concession configuration, P must have the structure given by (A-16).

Step 2: The second step is to show that the five constants, β_W , \tilde{x} , β_S , α_S , and \hat{x} , must satisfy equations (A-17)–(A-21) in the concession configuration. First, the continuity of P , combined with (A-16), implies that

$$\beta_W \tilde{x} = \alpha_S + \beta_S \tilde{x} \quad (\text{A-25})$$

$$\alpha_S + \beta_S \hat{x} = 1. \quad (\text{A-26})$$

Next, by Lemma A-1, in the concession configuration, it must be that $p_t^C = \alpha_t + \beta_t \mu_t$, $t \in \{S, W\}$. Thus, given that $\alpha_W = 0$, we must have

$$\beta_W \mu_W = p_W^C \quad (\text{A-27})$$

$$\alpha_S + \beta_S \mu_S = p_S^C. \quad (\text{A-28})$$

Third, note that $\text{Supp}_W = [0, \tilde{x}]$ and $\text{Supp}_S = [\tilde{x}, \hat{x}]$. Thus, given that no one places any point mass in equilibrium, for a given contestant, if his performance equals \tilde{x} , he will outperform all weak rivals but be outperformed by all strong rivals. Given that each rival is strong with probability θ , the given contestant's probability of winning by having performance equal to \tilde{x}

in the concession configuration is given by \tilde{p} in (A-22). Thus, it must be that $P(\tilde{x}) = \tilde{p}$. Given that $P(\tilde{x}) = \beta_W \tilde{x}$, we must have

$$\beta_W \tilde{x} = \tilde{p}. \quad (\text{A-29})$$

Equations (A-25)–(A-29) imply equations (A-17)–(A-21).

Step 3: The above construction of P ensures that P is continuous, bounded, and piecewise linear with one kink over its support and with $P(0) = 0$. We know that, in equilibrium, P must also be strictly increasing over its support and weakly concave. P , constructed in (A-16), will be strictly increasing over its support and weakly concave if and only if $\alpha_S \geq 0$ and $\beta_W \geq \beta_S > 0$. We now show that

$$p_W^C \geq p_W^G \implies \alpha_S \geq 0 \quad \& \quad \beta_W \geq \beta_S > 0, \quad (\text{A-30})$$

with $\alpha_S = 0$ and $\beta_W = \beta_S$ if and only if $p_W^C = p_W^G$.

Note that, p_S^C is the probability of winning if a contestant always bests weak rivals and ties with strong rivals. \tilde{p} , given by (A-22), is the probability of winning if a contestant always bests weak rivals but is always beaten by strong rivals. p_W^C is the probability of winning if a contestant is always beaten by strong rivals and ties with weak rivals. It is thus clear that

$$p_S^C > \tilde{p} > p_W^C > 0. \quad (\text{A-31})$$

Also note that, by identity (6), $p_W^C \geq p_W^G$ implies that $p_S^C \leq p_S^G$. Thus, given that $p_W^G/p_S^G = \mu_W/\mu_S$, we must have

$$p_W^C \geq p_W^G \implies p_W^C \geq \frac{p_S^C \mu_W}{\mu_S}, \quad (\text{A-32})$$

with equality if and only if $p_W^C = p_W^G$. By (A-31), $p_S^C > \tilde{p}$. Thus, $p_W^C \geq \frac{p_S^C \mu_W}{\mu_S}$ implies that $p_W^C > \frac{\tilde{p} \mu_W}{\mu_S}$. Thus, by (A-32),

$$p_W^C \geq p_W^G \implies p_W^C > \frac{\tilde{p} \mu_W}{\mu_S}. \quad (\text{A-33})$$

By (A-17) and (A-19),

$$\beta_W - \beta_S = \frac{p_W^C \mu_S - p_S^C \mu_W}{\mu_W (\mu_S - (\tilde{p} \mu_W / p_W^C))}. \quad (\text{A-34})$$

Suppose $p_W^C \geq p_W^G$. By (A-32), the numerators of the right hand sides of (A-20) and (A-34) are both nonnegative and are zero if and only if $p_W^C = p_W^G$. By (A-33), the denominators of the right hand sides of (A-19), (A-20), and (A-34) are positive. By (A-31), the numerator of the right hand side of (A-19) is positive. These facts imply (A-30). Thus, by (A-30), if $p_W^C \geq p_W^G$, then P , constructed in (A-16), is strictly increasing over its support and weakly concave, while if $p_W^C = p_W^G$, P is strictly increasing and linear over its support. The latter represents the boundary case in which P satisfies the linearity condition for the challenge configuration but weak types still concede to strong types.

Step 4: The above analysis shows that, when $p_W^C \geq p_W^G$, we can always construct, according to (A-16), a continuous, piecewise linear, weakly concave P that intersects the origin with β_W , \tilde{x} , α_S , β_S , and \hat{x} given by equations (A-17)–(A-21), respectively. The next step is to construct

the CDFs, F_W and F_S , that produce such a P . Note that, because in the concession configuration, $\text{Supp}_W = [0, \tilde{x}]$ and $\text{Supp}_S = [\tilde{x}, \hat{x}]$, and because F_W and F_S are continuous distributions, it must be that $F_W(x) = 1$ for $x \geq \tilde{x}$, $F_S(x) = 0$ for $x \leq \tilde{x}$, and $F_S(x) = 1$ for $x \geq \hat{x}$. Thus, by equations (A-14), (A-15) and (A-16), F_W and F_S must satisfy (A-23) and (A-24).

Step 5: The final step is to show that F_W and F_S , given by (A-23) and (A-24), respectively, are continuous CDFs that satisfy their associated contest ability constraints. We first show that F_W , implicitly given by (A-23), is a continuous CDF without any point mass and has support given by $[0, \tilde{x}]$ and mean equal to μ_W . Continuity follows simply from the facts that the right hand side of (A-23) is continuous in x and the left hand side of (A-23) is continuous in F_W , implying that F_W must be continuous in x . Equation (A-23) also implies that $F_W(0) = 0$. Thus, F_W has no point mass on 0. Equations (A-22), (A-23), and (A-29) imply that $F_W(\tilde{x}) = 1$. Thus, by construction, the support of F_W is within $[0, \tilde{x}]$. In fact, the support of F_W is exactly $[0, \tilde{x}]$. This is because the right hand side of (A-23) is increasing in x for $x \in [0, \tilde{x}]$, implying that the left hand side of (A-23) must be increasing in x for $x \in [0, \tilde{x}]$. Note that the left hand side of (A-23) is increasing in F_W .²⁵ Thus, for $x \in [0, \tilde{x}]$, F_W must be increasing in x . Hence, F_W is a continuous CDF without any point mass and with support given by $[0, \tilde{x}]$.

To see that the mean of F_W equals μ_W , integrate both sides of (A-23) with respect to F_W over the interval $[0, \tilde{x}]$. This yields that

$$\int_0^{\tilde{x}} \sum_{i=n-m}^{n-1} \binom{n-1}{i} [(1-\theta)F_W(x)]^i [1 - (1-\theta)F_W(x)]^{n-1-i} dF_W(x) = \int_0^{\tilde{x}} \beta_W x dF_W(x). \quad (\text{A-35})$$

By integration by substitution, using $p = (1-\theta)F_W(x)$, and taking the constants outside the integral, we can rewrite the left hand side of (A-35) as

$$\begin{aligned} \int_0^{\tilde{x}} \sum_{i=n-m}^{n-1} \binom{n-1}{i} [(1-\theta)F_W(x)]^i [1 - (1-\theta)F_W(x)]^{n-1-i} dF_W(x) \\ = \frac{1}{1-\theta} \sum_{i=n-m}^{n-1} \binom{n-1}{i} \int_0^{1-\theta} p^i (1-p)^{n-1-i} dp \\ = \frac{1}{1-\theta} \sum_{j=0}^{m-1} \binom{n-1}{j} \int_0^{1-\theta} p^{n-1-j} (1-p)^j dp, \end{aligned} \quad (\text{A-36})$$

where the last line follows from re-indexing using $j = n-1-i$ and the fact that $\binom{n-1}{n-1-j} = \binom{n-1}{j}$. By Wadsworth and Bryan (1960, p.52), the CDF of a Binom(n, θ) random variable can be expressed as

$$B(j; n, \theta) = (n-j) \binom{n}{j} \int_0^{1-\theta} p^{n-1-j} (1-p)^j dp. \quad (\text{A-37})$$

²⁵To see this, let $B(\cdot; n, p)$ denote the CDF of a Binom(n, p) random variable. Because, for any $p' > p$, a Binom(n, p') random variable dominates a Binom(n, p) random variable in the likelihood ratio ordering and, *a fortiori*, dominates a Binom(n, p) random variable in the stochastic order. Thus, for any fixed k and n , $B(k; n, p)$ is decreasing in p . Note that the left hand side of (A-23) equals $1 - B(n-m-1; n-1, (1-\theta)F_W)$. Thus, given that $B(k; n, p)$ is decreasing in p , it must be that the left hand side of (A-23) is increasing in F_W .

By (A-37) and the fact that $\binom{n-1}{j} = \frac{1}{n}(n-j)\binom{n}{j}$,

$$\begin{aligned} \frac{1}{1-\theta} \sum_{i=n-m}^{n-1} \binom{n-1}{i} \int_0^{1-\theta} p^i (1-p)^{n-1-i} dp \\ = \frac{1}{1-\theta} \sum_{j=0}^{m-1} \frac{1}{n} B(j; n, \theta) = \frac{1}{1-\theta} \sum_{j=0}^{m-1} \left(\frac{j+1}{n} - \frac{j}{n} \right) B(j; n, \theta). \end{aligned} \quad (\text{A-38})$$

Let $b(\cdot; n, \theta)$ be the probability mass function of a Binom(n, θ) random variable. Because $B(j; n, \theta) - B(j-1; n, \theta) = b(j; n, \theta)$ and because $B(0; n, \theta) = 0$, applying Abel summation by parts to (A-38) yields

$$\begin{aligned} \frac{1}{1-\theta} \sum_{i=n-m}^{n-1} \binom{n-1}{i} \int_0^{1-\theta} p^i (1-p)^{n-1-i} dp \\ = \frac{1}{1-\theta} \left(\frac{m}{n} B(m-1; n, \theta) - \sum_{j=1}^{m-1} \frac{j}{n} b(j; n, \theta) \right) \\ = \frac{m}{n(1-\theta)} \sum_{j=0}^{m-1} \binom{n}{j} \theta^j (1-\theta)^{n-j} - \frac{1}{1-\theta} \sum_{j=1}^{m-1} \frac{j}{n} \binom{n}{j} \theta^j (1-\theta)^{n-j} \\ = \sum_{j=0}^{m-1} \frac{m}{n-j} \binom{n-1}{j} \theta^j (1-\theta)^{n-1-j} - \sum_{j=1}^{m-1} \frac{j}{n-j} \binom{n-1}{j} \theta^j (1-\theta)^{n-1-j} \\ = \sum_{j=0}^{m-1} \frac{m-j}{n-j} \binom{n-1}{j} \theta^j (1-\theta)^{n-1-j} = \sum_{j=0}^{m-1} \frac{m-j}{n-j} b(j; n-1, \theta) = p_W^C = \beta_W \mu_W, \end{aligned} \quad (\text{A-39})$$

where the fourth line follows from the facts that $\frac{m}{n} \binom{n}{j} = \frac{m}{n-j} \binom{n-1}{j}$ and $\frac{j}{n} \binom{n}{j} = \frac{j}{n-j} \binom{n-1}{j}$, and in the last line, the first equality follows from combining common factors and the fact that $\frac{j}{n-j} \binom{n-1}{j} \theta^j (1-\theta)^{n-1-j} = 0$ if $j = 0$, the second equality follows from equation (4), and the last from equation (A-27). Equations (A-36) and (A-39), combined with the fact that, in the concession configuration, $\text{Supp}_W = [0, \tilde{x}]$, imply that $\int_0^{\tilde{x}} x dF_W(x) = \mu_W$. Thus, our construction of F_W satisfies the weak type's contest ability constraint.

Next, we show that F_S , implicitly defined by (A-24), satisfies all the conditions required for F_S to be the strong type's performance distribution played in the concession equilibrium. Because the proof is analogous to the one used for the case of F_W except for the proof of the satisfaction of the contest ability constraint, in what follows, we only give the proof that F_S , given by (A-24), satisfies the strong type's contest ability constraint. Integrate both sides of (A-24) with respect to F_S over the interval $[\tilde{x}, \hat{x}]$. This yields

$$\int_{\tilde{x}}^{\hat{x}} \sum_{i=n-m}^{n-1} \binom{n-1}{i} [1-\theta + \theta F_S(x)]^i [\theta(1-F_S(x))]^{n-1-i} dF_S(x) = \int_{\tilde{x}}^{\hat{x}} \alpha_S + \beta_S x dF_S(x). \quad (\text{A-40})$$

By integration by substitution, using $p = 1 - \theta + \theta F_S(x)$, and taking the constants outside the

integral, we can rewrite the left hand side of (A-40) as

$$\begin{aligned}
& \int_{\tilde{x}}^{\hat{x}} \sum_{i=n-m}^{n-1} \binom{n-1}{i} [1 - \theta + \theta F_S(x)]^i [\theta(1 - F_S(x))]^{n-1-i} dF_S(x) \\
&= \frac{1}{\theta} \sum_{i=n-m}^{n-1} \binom{n-1}{i} \int_{1-\theta}^1 p^i (1-p)^{n-1-i} dp \\
&= \frac{1}{\theta} \sum_{i=n-m}^{n-1} \binom{n-1}{i} \left(\mathcal{B}(i+1, n-i) - \int_0^{1-\theta} p^i (1-p)^{n-1-i} dp \right) \\
&= \frac{1}{\theta} \frac{m}{n} - \frac{1}{\theta} \sum_{i=n-m}^{n-1} \binom{n-1}{i} \int_0^{1-\theta} p^i (1-p)^{n-1-i} dp = \frac{1}{\theta} \frac{m}{n} - \frac{1-\theta}{\theta} p_W^C = p_S^C = \alpha_S + \beta_S \mu_S,
\end{aligned} \tag{A-41}$$

where $\mathcal{B}(a, b)$ denotes the Beta function with parameters a and b . The third line follows from the fact that $\int_0^1 p^i (1-p)^{n-1-i} dp = \mathcal{B}(i+1, n-i)$. In the last line, the first equality follows from the fact that $\binom{n-1}{i} \mathcal{B}(i+1, n-i) = \frac{1}{n}$, the second from (A-39), the third from (6), and the last from (A-28). Equation (A-40) and (A-41), combined with the fact that, in the concession configuration, $\text{Supp}_S = [\tilde{x}, \hat{x}]$, imply that $\int_{\tilde{x}}^{\hat{x}} x dF_S(x) = \mu_S$. Thus, our construction of F_S satisfies the strong type's contest ability constraint. \square

Lemma A-3 shows the construction of the concession equilibrium when $p_W^C \geq p_W^G$. Now we turn to the case in which $p_W^C < p_W^G$. The next lemma shows the existence of challenge equilibria when $p_W^C < p_W^G$. In its proof, we construct one challenge equilibrium in which the two types' performance distributions satisfy the monotone likelihood ratio property, defined in the lemma. The existence of such an equilibrium will be used for a later proof.

Lemma A-4. *When $p_W^C < p_W^G$, there exist challenge equilibria. All of these challenge equilibria produce the same probability of winning function, P , given by*

$$P(x) = \begin{cases} \frac{m}{n\bar{\mu}} x & x \in [0, n\bar{\mu}/m] \\ 1, & x \geq n\bar{\mu}/m \end{cases}, \tag{A-42}$$

where $\bar{\mu} = \theta \mu_S + (1 - \theta) \mu_W$ is the expected contest ability of a contestant of unknown type. At least in one of these challenge equilibria, F_S and F_W satisfy the monotone likelihood ratio property (MLRP), i.e.,

$$\text{for every } x'' > x' \geq 0, \quad \frac{f_S(x'')}{f_W(x'')} \geq \frac{f_S(x')}{f_W(x')},$$

where we treat $f_S(x)/f_W(x) = +\infty$ if $f_W(x) = 0$. (The construction of such a challenge equilibrium is provided in the proof.)

Proof. First, we establish equation (A-42). Note that, in the challenge configuration, P is linear over its support with $P(0) = 0$. Thus, by Lemmas A-1 and A-2, in any challenge equilibrium, the optimal dual variables in the Lagrangian (A-1) must satisfy that $\alpha_W = \alpha_S = 0$ and $\beta_W = \beta_S$.

Thus, by Lemma A-2, in any challenge equilibrium, over its support,

$$P(x) = \beta_W x. \quad (\text{A-43})$$

Thus, by Lemma A-1 and the fact that $\alpha_W = 0$, we have $p_W^G = P(\mu_W) = \beta_W \mu_W$. Thus, by (9) and the fact that, in any challenge equilibrium, $\beta_W = \beta_S$, it must be that, in any challenge equilibrium,

$$\beta_W = \beta_S = \frac{m}{n\bar{\mu}}, \quad (\text{A-44})$$

where $\bar{\mu} = \theta\mu_S + (1 - \theta)\mu_W$. Equation (A-42) then follows from equations (A-43) and (A-44).

The next step is to find a pair of CDFs, F_S and F_W , that produce P constructed in (A-42) and satisfy their contest ability constraints under the condition $p_W^C < p_W^G$. The following argument gives such a construction.

Step 1: Identify the equilibrium performance distributions in the boundary case. Note that, by (4), $p_W^C \in (0, m/n)$ is independent of r , where $r = \mu_S/\mu_W$. By (9), p_W^G is strictly decreasing in r , tends to m/n as $r \rightarrow 1$ and tends to 0 as $r \rightarrow \infty$. Thus, fixing n , m , and θ , there exists a unique $r^o > 1$ such that $p_W^C = p_W^G(r^o)$ and

$$p_W^C < p_W^G(r) \iff r < r^o. \quad (\text{A-45})$$

Note that, given μ_S , μ_W , and r^o , there uniquely exist μ_S^o and μ_W^o such that

$$\theta\mu_S^o + (1 - \theta)\mu_W^o = \theta\mu_S + (1 - \theta)\mu_W \quad (\text{A-46})$$

$$\mu_S^o/\mu_W^o = r^o. \quad (\text{A-47})$$

By construction and the fact that $\mu_W < \mu_S$,

$$r = \frac{\mu_S}{\mu_W} < r^o \implies \mu_W^o < \mu_W < \mu_S < \mu_S^o. \quad (\text{A-48})$$

Now let F_S^o and F_W^o be the equilibrium performance distributions played in the case where the strong and the weak type's contest abilities equal μ_S^o and μ_W^o , respectively. Because, by construction, $\mu_S^o/\mu_W^o = r^o$, and because $p_W^C = p_W^G(r^o)$, this case is the boundary case, in which F_S^o and F_W^o jointly produce a uniform probability of winning function but F_S^o and F_W^o have adjacent supports. Lemma A-3 shows the construction of equilibrium in this boundary case. By Lemma A-3, in this boundary case, there uniquely exists a pair of CDFs, F_S^o and F_W^o , such that the mean of F_S^o equals μ_S^o , the mean of F_W^o equals μ_W^o , the lower endpoint of the support of F_S^o equals the upper endpoint of the support of F_W^o , and F_S^o and F_W^o jointly produce a uniform probability of winning function through (A-15).

Step 2: Mix the boundary equilibrium performance distributions. By (A-45) and (A-48), if $p_W^C < p_W^G(r)$, there must exist $\rho_W \in (0, 1)$ and $\rho_S \in (0, 1)$ such that

$$\rho_W \mu_W^o + (1 - \rho_W) \mu_S^o = \mu_W \quad \& \quad \rho_S \mu_W^o + (1 - \rho_S) \mu_S^o = \mu_S, \quad (\text{A-49})$$

in which case

$$\rho_W = \frac{\mu_S^o - \mu_W}{\mu_S^o - \mu_W^o} \quad \& \quad \rho_S = \frac{\mu_S^o - \mu_S}{\mu_S^o - \mu_W^o}. \quad (\text{A-50})$$

Equations (A-46) and (A-49) imply that ρ_W and ρ_S satisfy that

$$\theta\rho_S + (1 - \theta)\rho_W = 1 - \theta. \quad (\text{A-51})$$

Now we argue that, when $p_W^C < p_W^G(r)$, there exists a challenge equilibrium in which strong and weak types play the following strategy:

$$\text{S-strategy} = \begin{cases} F_W^o & \text{w. p. } \rho_S \\ F_S^o & \text{w. p. } 1 - \rho_S \end{cases} \quad \text{W-strategy} = \begin{cases} F_W^o & \text{w. p. } \rho_W \\ F_S^o & \text{w. p. } 1 - \rho_W \end{cases}, \quad (\text{A-52})$$

where ρ_W and ρ_S are given in (A-50). Note that, by (A-49), the prescribed strategies satisfy each type's contest ability constraint. Also, by construction, a contestant of unknown type will play F_W^o with probability $\theta\rho_S + (1 - \theta)\rho_W = 1 - \theta$, where the equality follows from (A-51), and play F_S^o with the complementary probability, θ . Thus, by construction, the unconditional performance distribution equals the one in the boundary case, i.e., $\theta F_S^o + (1 - \theta)F_W^o = \theta F_S^o + (1 - \theta)F_W^o$. Because, by (A-15), the shape of P only depends on the unconditional performance distribution, and because $\theta F_S^o + (1 - \theta)F_W^o$ produces a uniform P , it must be that $\theta F_S^o + (1 - \theta)F_W^o$ also produces a uniform P . Thus, the prescribed strategies constitute a challenge equilibrium.

Finally, we show that the prescribed strategies satisfy MLRP. Note that, by construction, the support of F_W^o is below the support of F_S^o . Also note that, by (A-48) and (A-50), $\rho_S < \rho_W$. Thus, by (A-52), in the prescribed equilibrium, the strong type plays F_W^o with a lower probability and plays F_S^o with a higher probability compared to the weak type. Thus, the prescribed strategies satisfy MLRP. \square

Lemmas A-3 and A-4 establish the “if” part. This completes the proof. \square

Proof of Lemma 2. To establish the first part, note that equation (4) implies that

$$p_W^C = \mathbb{E} \left[\frac{m - \tilde{S}_n^{-i}}{n - \tilde{S}_n^{-i}} \mid \tilde{S}_n^{-i} \leq m - 1 \right] \mathbb{P}[\tilde{S}_n^{-i} \leq m - 1] \leq \frac{m}{n} \mathbb{P}[\tilde{S}_n^{-i} \leq m - 1]. \quad (\text{A-53})$$

Because $r = \mu_S/\mu_W > 1$, equation (9) implies that

$$p_W^G = \frac{m}{n} \frac{1}{r\theta + (1 - \theta)} > \frac{m}{n} \frac{1}{r}. \quad (\text{A-54})$$

By Proposition 1, the challenge configuration will be played in equilibrium if and only if $p_W^G > p_W^C$. Thus, by (A-53) and (A-54), the challenge configuration will be played whenever

$$\mathbb{P}[\tilde{S}_n^{-i} \leq m - 1] \leq \frac{1}{r}. \quad (\text{A-55})$$

Everything else being equal, $\mathbb{P}[\tilde{S}_n^{-i} \leq m - 1] \rightarrow 0$ as $n \rightarrow \infty$, and $1/r \rightarrow 1$ as $r \rightarrow 1$. Thus, by (A-55), the challenge configuration will be played if either n is sufficiently large or r is sufficiently close to 1. This establishes the first part of the lemma.

To establish the second part, let $\tilde{S}_n \sim \text{Binom}(n, \theta)$. Note that, by (4), we can rewrite p_W^C as

$$\begin{aligned} p_W^C &= \sum_{s=0}^{n-1} \max \left[0, \frac{m-s}{n-s} \right] \binom{n-1}{s} \theta^s (1-\theta)^{n-1-s} \\ &= \frac{1}{n(1-\theta)} \sum_{s=0}^n \max[0, m-s] \binom{n}{s} \theta^s (1-\theta)^{n-s} = \frac{1}{n(1-\theta)} \mathbb{E} [\max[0, m - \tilde{S}_n]], \end{aligned} \quad (\text{A-56})$$

where the second equality follows from the binomial coefficient identity, $\binom{n-1}{s} = \frac{n-s}{n} \binom{n}{s}$, and the fact that $\max[0, m-n] = 0$. Thus, by equations (9) and (A-56),

$$\frac{p_W^C}{p_W^G} = \mathbb{E} \left[\max \left[1 - \frac{\tilde{S}_n}{m}, 0 \right] \right] \left(\frac{r\theta + 1 - \theta}{1 - \theta} \right). \quad (\text{A-57})$$

Equation (A-57) allows us to evaluate the effect of a parameter change on the equilibrium configuration. First consider a change of n . Note that $s \mapsto \max[1 - (s/m), 0]$ is nonincreasing. Also note that, when n increases, the distribution of \tilde{S}_n after the increase stochastically dominates the one before the increase. Thus, $n \mapsto \mathbb{E} \left[\max \left[1 - \frac{\tilde{S}_n}{m}, 0 \right] \right]$ is nonincreasing. Hence, by equation (A-57), $n \mapsto p_W^C(n)/p_W^G(n)$ must be nonincreasing. Thus, by Proposition 1, an increase in n favors the play of the challenge configuration.

Next, consider a change of m . Note that, for any fixed s , $m \mapsto \max[1 - (s/m), 0]$ is nondecreasing. Because a change in m does not change the distribution of \tilde{S}_n , $m \mapsto \mathbb{E} \left[\max \left[1 - \frac{\tilde{S}_n}{m}, 0 \right] \right]$ must be nondecreasing. Hence, by equation (A-57), $m \mapsto p_W^C(m)/p_W^G(m)$ must be nondecreasing. Thus, by Proposition 1, a decrease in m favors the play of the challenge configuration.

Finally, consider a change of r . Because p_W^C is independent of r while $r \mapsto p_W^G(r)$ is decreasing, it is clear that a decrease in r increases $p_W^G(r)$ relative to p_W^C . Thus, by Proposition 1, a decrease in r favors the play of the challenge configuration. \square

Proof of Lemma 3. Follows immediately from Proposition 1 and the fact that only concession equilibria implement merit-based selection. \square

Proof of Theorem 1. By Lemma 2, the challenge configuration will be played if n is sufficient large, in which case any further increase in n will not change the equilibrium configuration. The theorem then follows immediately from Lemma 3 and the fact that winner quality in the challenge configuration, given by equation (14), is independent of n . \square

Proof of Lemma 4. Obvious. \square

Proof of Theorem 2. We first prove part i of the theorem. Because selection under the concession configuration implements merit-based selection, if the concession configuration is played at $m = m_M^*$, it must be that $u(m_M^*) = u_M(m_M^*)$. Because $u(m) \leq u_M(m_M^*)$, it is then optimal to choose $m = m_M^*$. Thus, $u(m^*) = u(m_M^*) = u_M(m_M^*)$. This establishes the “if” part. To establish the “only if” part, suppose that the challenge configuration is played at $m = m_M^*$. Then by Lemma 2, the challenge configuration will also be played at $m < m_M^*$. Note that only m_M^* and (in the non-generic case) $m_M^* - 1$ (see the discussion around equations (16) and (17) in the main

text) can be optimal merit-based selection quotas. Thus, if the optimal contest selection quota, m^* , equals either m_M^* or $m_M^* - 1$, given that the challenge configuration is played at $m = m_M^*$ and, *a fortiori*, played at $m = m_M^* - 1$, by Lemma 3, it must be that $u(m^*) < u_M(m^*) \leq u_M(m_M^*)$. If m^* differs from m_M^* and $m_M^* - 1$ and, thus, differs from an optimal quota under merit-based selection, by Lemma 3, it must be that $u(m^*) \leq u_M(m^*) < u_M(m_M^*)$. Thus, in both cases, $u(m^*) < u_M(m_M^*)$. This establishes the “only if” part.

Next, we prove part ii. Note that winner quality, $\Pi(m, n)$, has the following upper bound:

$$\Pi(m, n) \leq 1 - \frac{n}{m}(1 - \theta)p_W^G = \frac{\theta r}{\theta r + 1 - \theta},$$

where the inequality follows from (13) and the equality from (9). Thus, by equation (12), designer welfare in the risk-taking contest, u , has the following upper bound:

$$u(m, n) \leq m \left(\frac{(1 + \sigma)\theta r}{\theta r + 1 - \theta} - \sigma \right) = m \left(\frac{\theta r - \sigma(1 - \theta)}{\theta r + 1 - \theta} \right). \quad (\text{A-58})$$

The last expression will be nonpositive if $r \leq \sigma(1 - \theta)/\theta$. Thus, if $r \leq \sigma(1 - \theta)/\theta$, $u(m, n) \leq 0$ and, clearly, this zero upper bound will be attained by choosing $m = 0$. This establishes part ii.

Finally, consider part iii. If the concession configuration is played at m_M^* , then by part i, which has been proved, it is optimal to choose $m = m_M^*$. Now suppose that the challenge configuration is played at $m = m_M^*$ and that $r > \sigma(1 - \theta)/\theta$. Note that, in the challenge configuration, designer welfare, u , is given by the last expression in (A-58), which, given the hypothesis that $r > \sigma(1 - \theta)/\theta$, is strictly increasing in m . This, combined with the fact that, by Lemma 2, the challenge configuration is played at any $0 < m \leq \bar{m}$, where \bar{m} is the largest quota at which the challenge configuration is played, implies that choosing $m = \bar{m}$ strictly dominates choosing $m < \bar{m}$. By the definition of \bar{m} , the concession configuration will be played at $m > \bar{m}$. Selection under the concession configuration implements merit-based selection. Thus, by (15), the marginal gain from increasing the quota from $m > \bar{m}$ to $m + 1$ is $1 - (1 + \sigma)B(m; n, \theta)$. By the definition of m_M^* in (16), $1 - (1 + \sigma)B(m; n, \theta) < 0$ for $m > m_M^*$. Thus, given that $\bar{m} \geq m_M^*$, the marginal gain from increasing the quota from $m > \bar{m}$ to $m + 1$ is strictly negative. Thus, choosing $m = m_M^* + 1$ strictly dominates choosing $m > m_M^* + 1$. Hence, it is optimal to choose either $m = m_M^*$ or $m = m_M^* + 1$. \square

Proof of Proposition 2. Throughout this proof, we fix n and θ . Let $m^*(r)$ be the optimal selection quota conditioned on r . By (A-45), for any fixed m , $p_W^C \geq p_W^G(r)$ for r sufficiently large. Thus, by Proposition 1, for r sufficiently large, the concession configuration will be played at $m = m_M^*$. Thus, by Theorem 2, $m^*(r) = m_M^*$ for r sufficiently large.

To complete the proof, all we need to show is that

$$m^*(r') \geq m^*(r''), \quad 1 < r' < r'', \quad \text{where } r' > \sigma(1 - \theta)/\theta. \quad (\text{A-59})$$

Note that, by Lemma 2, if the concession configuration is played at $m = m_M^*$ under $r = r'$, the concession configuration will also be played at $m = m_M^*$ under $r = r'' > r'$, in which case, by Theorem 2, $m^*(r') = m^*(r'') = m_M^*$ and, hence, (A-59) holds. Also note that, by Theorem 2,

for $r' > \sigma(1-\theta)/\theta$, $m^*(r') \geq m_M^*$. Moreover, by Theorem 2, if the concession configuration is played at $m = m_M^*$ under $r = r''$, we must have $m^*(r'') = m_M^*$ and, hence, $m^*(r') \geq m_M^* = m^*(r'')$. Again, (A-59) holds.

The only case to consider is the one in which the challenge configuration is played at $m = m_M^*$ both under $r = r' > \sigma(1-\theta)/\theta$ and under $r = r'' > r'$. Let $\bar{m}(r)$ be the largest quota that sustains a challenge configuration under r , defined in (18). Because, for any fixed m , the play of the concession configuration under $r = r'$ implies the play of the concession configuration under $r = r'' > r'$, it must be that

$$\bar{m}(r') \geq \bar{m}(r''), \quad r' < r''.$$

By Theorem 2 and the hypothesis that $\sigma(1-\theta)/\theta < r' < r''$, if the challenge configuration is played at $m = m_M^*$ both under $r = r' > \sigma(1-\theta)/\theta$ and under $r = r'' > r'$, $m^*(r')$ equals either $\bar{m}(r')$ or $\bar{m}(r') + 1$, and $m^*(r'')$ equals either $\bar{m}(r'')$ or $\bar{m}(r'') + 1$. Thus, if $\bar{m}(r') > \bar{m}(r'')$, (A-59) obviously holds. Now suppose that $\bar{m}(r') = \bar{m}(r'')$. To establish (A-59), the only situation we need to rule out is the one in which $m^*(r') = \bar{m}(r')$, $m^*(r'') = \bar{m}(r'') + 1$, and $\bar{m}(r'') = \bar{m}(r')$. Let $u(m, r)$ be designer welfare under m and r . If $m^*(r') = \bar{m}(r')$, it must be that

$$u(\bar{m}(r'), r') > u(\bar{m}(r') + 1, r'). \quad (\text{A-60})$$

If $m^*(r'') = \bar{m}(r'') + 1$ and $\bar{m}(r'') = \bar{m}(r')$, it must be that

$$u(\bar{m}(r') + 1, r'') = u(\bar{m}(r'') + 1, r'') \geq u(\bar{m}(r''), r'') = u(\bar{m}(r'), r''). \quad (\text{A-61})$$

Because, by definition, $\bar{m}(r')$ and $\bar{m}(r'')$ are the largest quotas under which the challenge configuration sustains an equilibrium under $r = r'$ and $r = r''$, respectively, it must be that $\bar{m}(r') + 1$ and $\bar{m}(r'') + 1$ are the smallest quotas under which the concession configuration sustains an equilibrium under $r = r'$ and $r = r''$, respectively. Thus, if $\bar{m}(r') = \bar{m}(r'')$, it must be that the concession configuration is played at $m = \bar{m}(r') + 1$ both under $r = r'$ and under $r = r''$. In this case, given that, for any fixed m , if the concession configuration is played, designer welfare is independent of r , we must have $u(\bar{m}(r') + 1, r') = u(\bar{m}(r') + 1, r'')$. This implies, by (A-60) and (A-61), that $u(\bar{m}(r'), r') > u(\bar{m}(r'), r'')$ for $r' < r''$, which, however, is impossible, because for any fixed m , an increase in r weakly reduces the weak type's probability of winning and, hence, cannot make the designer strictly worse off. Thus, (A-59) is established. This completes the proof. \square

Proof of Proposition 3. We first prove part i. By Lemma 2 and Theorem 1, everything else being equal, there exists n^c such that for any $n'' > n' \geq n^c$, the challenge configuration is played both under $n = n''$ and under $n = n'$ and winner quality is the same under $n = n''$ as under $n = n'$.

Let $F_{m:n}$ be the equilibrium unconditional performance distribution when n contestants compete for m selection places. By equation (A-15) and Lemma A-4, if the challenge configuration is played, $F_{m:n}$ is given, over its support $[0, n\bar{\mu}/m]$, by

$$\sum_{i=n-m}^{n-1} \binom{n-1}{i} F_{m:n}(x)^i (1 - F_{m:n}(x))^{n-1-i} = \frac{m}{n\bar{\mu}} x, \quad (\text{A-62})$$

where $\bar{\mu} \equiv \theta\mu_S + (1 - \theta)\mu_W$. Thus, by Jones (2002), $F_{m:n}$ in the challenge configuration is a Complementary Beta distribution. Complementary Beta distributions are smooth and have positive derivatives on the interior of their supports. Thus, the inverse function, $F_{m:n}^{-1}$, in the challenge equilibrium is smooth and has positive derivatives over the open interval $(0, 1)$. Given that $F_{m:n'}$ and $F_{m:n''}$, $n'' > n'$, are two non-identical distributions with the same mean, to show that $F_{m:n''}$ is an MPS of $F_{m:n'}$, it suffices to show that $F_{m:n''}$ and $F_{m:n'}$ satisfy a single-crossing condition (cf. Diamond and Stiglitz, 1974): there exists x' such that $F_{m:n''}(x) - F_{m:n'}(x) \leq (\geq) 0$ when $x \geq (\leq) x'$. This single-crossing condition can be equivalently expressed in terms of the quantile functions: there exists $\hat{q} \in (0, 1)$ such that $F_{m:n''}^{-1}(q) - F_{m:n'}^{-1}(q) \geq (\leq) 0$, when $q \geq (\leq) \hat{q}$. We prove this below.

Note that equation (A-62) implies that

$$F_{m:n}^{-1}(q) = \frac{n\bar{\mu}}{m} \sum_{i=n-m}^{n-1} \binom{n-1}{i} q^i (1-q)^{n-1-i}, \quad q \in (0, 1). \quad (\text{A-63})$$

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

$$\begin{aligned} F_{m:n''}^{-1}(q) - F_{m:n'}^{-1}(q) &= \frac{n''\bar{\mu}}{m} \sum_{i=n''-m}^{n''-1} \binom{n''-1}{i} q^i (1-q)^{n''-1-i} \\ &\quad - \frac{n'\bar{\mu}}{m} \sum_{i=n'-m}^{n'-1} \binom{n'-1}{i} q^i (1-q)^{n'-1-i}. \end{aligned} \quad (\text{A-64})$$

Differentiate (A-64) with respect to q , apply the result that $(i+1)\binom{n-1}{i+1} = (n-1-i)\binom{n-1}{i}$ to cancel the common terms, and combine the common factors. This yields

$$\frac{d(F_{m:n''}^{-1}(q) - F_{m:n'}^{-1}(q))}{dq} = \frac{\bar{\mu}q^{n''-1-m}(1-q)^{m-1}}{m} K(q), \quad (\text{A-65})$$

where $K(q) = n''(n''-m)\binom{n''-1}{m-1}q^{n''-n'} - n'(n'-m)\binom{n'-1}{m-1}$. When $q \in (0, 1)$, the sign of (A-65) is determined by the sign of K . Note that $K < 0$ when $q = 0$, $K > 0$ when $q = 1$, and K is continuous and strictly increasing for $q \geq 0$. Thus, there exists $q^* \in (0, 1)$ such that K crosses the horizontal axis from below at $q = q^*$. This implies, by (A-65), that $F_{m:n''}^{-1} - F_{m:n'}^{-1}$ is strictly decreasing for $q \in (0, q^*)$ and strictly increasing for $q \in (q^*, 1)$. Since $F_{m:n''}^{-1}(0) = F_{m:n'}^{-1}(0) = 0$, it follows that $F_{m:n''}^{-1}(q) - F_{m:n'}^{-1}(q) < 0$ for $q \in (0, q^*]$. This result, together with the facts that $F_{m:n''}^{-1}(1) - F_{m:n'}^{-1}(1) = (n'' - n')\bar{\mu}/m > 0$ and $F_{m:n''}^{-1} - F_{m:n'}^{-1}$ is continuous and strictly increasing for $q \in (q^*, 1)$, implies the satisfaction of the single-crossing condition. Thus, $F_{m:n''}$ is an MPS of $F_{m:n'}$.

Part ii can be proved by a similar argument. Note that, by Lemma 2, everything else being equal, if the challenge configuration is played at $m = m^c$, the challenge configuration will also be played at any $m \leq m^c$. Thus, if the challenge configuration is played at $m = m^c$, for any $m'' < m' \leq m^c$, the challenge configuration will be played both under $m = m''$ and under $m = m'$.

Hence, $F_{m'',n}$ and $F_{m',n}$ will both satisfy equation (A-62). Thus, by (A-63),

$$F_{m'',n}^{-1}(q) - F_{m',n}^{-1}(q) = \frac{n\bar{\mu}}{m''} \sum_{i=n-m''}^{n-1} \binom{n-1}{i} q^i (1-q)^{n-1-i} - \frac{n\bar{\mu}}{m'} \sum_{i=n-m'}^{n-1} \binom{n-1}{i} q^i (1-q)^{n-1-i}. \quad (\text{A-66})$$

Differentiate (A-66) with respect to q , apply the result that $(i+1)\binom{n-1}{i+1} = (n-1-i)\binom{n-1}{i}$ to cancel the common terms, and combine the common factors. This yields

$$\frac{d(F_{m'',n}^{-1}(q) - F_{m',n}^{-1}(q))}{dq} = n\bar{\mu}q^{n-1-m'}(1-q)^{m''-1}J(q), \quad (\text{A-67})$$

where $J(q) = \frac{n-m''}{m''} \binom{n-1}{n-m''} q^{m'-m''} - \frac{n-m'}{m'} \binom{n-1}{n-m'} (1-q)^{m'-m''}$. When $q \in (0, 1)$, the sign of (A-67) is determined by the sign of J . Note that $J < 0$ when $q = 0$, $J > 0$ when $q = 1$, and J is continuous and strictly increasing for $q \in [0, 1]$. Thus, there exists $q^o \in (0, 1)$ such that J single crosses the horizontal axis from below at $q = q^o$. This implies, by (A-67), that $F_{m'',n}^{-1} - F_{m',n}^{-1}$ is strictly decreasing for $q \in (0, q^o)$ and strictly increasing for $q \in (q^o, 1)$. Since $F_{m'',n}^{-1}(0) = F_{m',n}^{-1}(0) = 0$, it follows that $F_{m'',n}^{-1}(q) - F_{m',n}^{-1}(q) < 0$ for $q \in (0, q^o]$. This result, together with the facts that $F_{m'',n}^{-1}(1) - F_{m',n}^{-1}(1) = n\bar{\mu}(1/m'' - 1/m') > 0$ and $F_{m'',n}^{-1} - F_{m',n}^{-1}$ is continuous and strictly increasing for $q \in (q^o, 1)$, implies the satisfaction of the single-crossing condition. Thus, $F_{m'',n}$ is an MPS of $F_{m',n}$.

Part iii has been established in Lemma A-4. □

Proof of Corollary 1. Established by the argument in the main text before Corollary 1. □

Proof of Proposition 4. In a symmetric equilibrium, at the contest ability acquisition stage, every weak type chooses the same contest ability μ_W and every strong type chooses the same contest ability μ_S . Note that, in any symmetric equilibrium, it must be that $\mu_W > 0$ and $\mu_S > 0$. This is because choosing zero contest ability would imply placing point mass on 0. By the same argument as the one used in Section 2.2 for showing the continuity of P and by the fact that the cost of choosing $\varepsilon > 0$ contest ability can be made arbitrarily small by shrinking ε to zero, it is clear that placing point mass on 0 cannot be sustained in a symmetric equilibrium.

Given $\mu_W > 0$ and $\mu_S > 0$, the weak and the strong types' performance distributions, F_W and F_S , and the probability of winning function, P , at the risk-taking stage are characterized by Lemmas A-3 and A-4. Let $P(\cdot; \mu_W, \mu_S)$ be the probability of winning function at the risk-taking stage when, at the contest ability acquisition stage, the weak type chooses μ_W and the strong type chooses μ_S . By Lemma 1, in any symmetric equilibrium, P is weakly concave.²⁶ Because taking no risk is a weakly optimal best reply to a weakly concave P , by choosing contest ability

²⁶Weak concavity of P holds even if $\mu_S \leq \mu_W$, because if $\mu_S < \mu_W$, we can simply treat the high-ability type as the weak type and the low-ability type as the strong type at the risk-taking stage and all the arguments in the proof of Lemma 1 apply. If $\mu_S = \mu_W$, one can treat every contestant as of the same type at the risk-taking stage by treating either $\theta = 0$ or $\theta = 1$. The proof of Lemma 1 does not rely on the value of θ . Thus, the concavity of P still holds under $\mu_S = \mu_W$.

μ , a contestant's probability of winning is given by $P(\mu; \mu_W, \mu_S)$. In a symmetric equilibrium, it must be that a type- $t \in \{S, W\}$ contestant's expected payoff, $P(\mu; \mu_W, \mu_S) - (\mu^\alpha/a_t)$ is maximized at $\mu = \mu_t$. Thus, by the first-order conditions, we must have

$$P'(\mu_W; \mu_W, \mu_S) = \frac{\alpha \mu_W^{\alpha-1}}{a_W} \quad \& \quad P'(\mu_S; \mu_W, \mu_S) = \frac{\alpha \mu_S^{\alpha-1}}{a_S}. \quad (\text{A-68})$$

We first argue that, in any symmetric equilibrium, it must be that $\mu_S > \mu_W$. This is because, if, to the contrary, $\mu_S \leq \mu_W$, the fact that $a_W < a_S$ would imply that $\alpha \mu_W^{\alpha-1}/a_W > \alpha \mu_S^{\alpha-1}/a_S$. Thus, by (A-68), it would have to be that $P'(\mu_W; \mu_W, \mu_S) > P'(\mu_S; \mu_W, \mu_S)$, which, given the weak concavity of P , could only happen if $\mu_W < \mu_S$, contradicting the hypothesis that $\mu_S \leq \mu_W$.

Next, by Proposition 1, P can only be of one of two configurations: concession or challenge. In the challenge configuration, by Lemma A-4, $P' = m/(n\bar{\mu})$ over the support of P , where $\bar{\mu} = \theta \mu_S + (1 - \theta) \mu_W$. Thus, by (A-68), in the challenge configuration, it must be that

$$\frac{m}{n\bar{\mu}} = \frac{\alpha \mu_t^{\alpha-1}}{a_t}, \quad t \in \{S, W\}, \quad (\text{A-69})$$

which implies that

$$r = \frac{\mu_S}{\mu_W} = \left(\frac{a_S}{a_W}\right)^{\frac{1}{\alpha-1}} \equiv r^*. \quad (\text{A-70})$$

By Proposition 1, equation (A-70), and the fact that p_W^C is independent of r , a challenge equilibrium exists only if $p_W^C < p_W^G(r^*)$.

To show that $p_W^C < p_W^G(r^*)$ is sufficient for the existence of a challenge equilibrium, we only need to verify that no contestant has an incentive to deviate from acquiring contest ability according to (A-69) when each type- t rival chooses μ_t according to (A-69) and plays risk-taking strategies to produce a uniform P with $P' = m/(n\bar{\mu})$ over its support. This verification is easy. Note that, given $P' = m/(n\bar{\mu})$ over its support, the choice of μ_t given by (A-69) satisfies both the first-order and the second-order conditions. Also note that the individual rationality constraint is also satisfied, because, when $P(x) = mx/(n\bar{\mu})$ over its support, by choosing μ_t according to (A-69), a type- t contestant has expected payoff equal to

$$\frac{m\mu_t}{n\bar{\mu}} - \frac{\mu_t^\alpha}{a_t} = \mu_t \left(\frac{m}{n\bar{\mu}} - \frac{\mu_t^{\alpha-1}}{a_t} \right) = \frac{\mu_t m}{n\bar{\mu}} \left(1 - \frac{1}{\alpha} \right) > 0,$$

where the second equality follows from (A-69) and the inequality from the assumption that $\alpha > 1$. Thus, a challenge equilibrium exists if and only if $p_W^C < p_W^G(r^*)$ and the endogenous strength asymmetry in challenge equilibria equals r^* .

Now we show that there is no concession equilibria if $p_W^C < p_W^G(r^*)$. Suppose, to the contrary, that a concession equilibrium exists given $p_W^C < p_W^G(r^*)$. By Lemma A-3, in the concession equilibrium, the marginal benefit of contest ability is higher for the weak type than for the strong type, i.e., $\beta_W \geq \beta_S$. By the first-order conditions,

$$\frac{\alpha \mu_W^{\alpha-1}}{a_W} = \beta_W \quad \& \quad \frac{\alpha \mu_S^{\alpha-1}}{a_S} = \beta_S, \quad (\text{A-71})$$

implying, given $\beta_S \leq \beta_W$, that

$$r = \frac{\mu_S}{\mu_W} = \left(\frac{a_S \beta_S}{a_W \beta_W} \right)^{\frac{1}{\alpha-1}} \leq \left(\frac{a_S}{a_W} \right)^{\frac{1}{\alpha-1}} = r^*. \quad (\text{A-72})$$

Because $r \mapsto p_W^G(r)$ is decreasing, the hypothesis $p_W^C < p_W^G(r^*)$ and equation (A-72) imply that $p_W^C < p_W^G(r)$. Thus, by Proposition 1, at the risk-taking stage, only the challenge configuration can sustain an equilibrium, a contradiction. This contradiction implies that concession equilibria exist only if $p_W^C \geq p_W^G(r^*)$.

Next, we show that there exists a concession equilibrium if $p_W^C \geq p_W^G(r^*)$. We first show the expression for the strength asymmetry in the boundary case in the next lemma.

Lemma A-5. $p_W^C = p_W^G(r)$ if and only if $r = p_S^C/p_W^C$. $p_W^C > (<) p_W^G(r)$ if and only if $r > (<)$ p_S^C/p_W^C .

Proof. By equation (6),

$$p_W^C = p_W^G(r) \iff p_S^C = p_S^G(r) \iff p_S^C/p_W^C = p_S^G(r)/p_W^G(r), \quad (\text{A-73})$$

where $p_S^G(r)/p_W^G(r) = \mu_S/\mu_W = r$. Thus, $p_W^C = p_W^G(r)$ if and only if $p_S^C/p_W^C = r$. The rest of the result follows from the fact that p_W^C is independent of r while $p_W^G(r)$ is decreasing in r . \square

Let $r^o = p_S^C/p_W^C$. By Lemma A-5,

$$p_W^C \geq p_W^G(r^*) \iff r^* \geq r^o. \quad (\text{A-74})$$

Note that the optimal dual variables, β_W and β_S , in the concession equilibrium are given by (A-17) and (A-19), respectively. Thus, (A-17), (A-19), and (A-71) imply that, in any concession equilibrium, it must be that

$$\frac{\alpha \mu_W^{\alpha-1}}{a_W} = \frac{p_W^C}{\mu_W} \quad (\text{A-75})$$

$$\frac{\alpha \mu_S^{\alpha-1}}{a_S} = \frac{p_S^C - \tilde{p}}{\mu_S - (\tilde{p} \mu_W / p_W^C)}, \quad (\text{A-76})$$

where p_W^C and \tilde{p} are given by (4) and (A-22), respectively, and p_S^C is determined by p_W^C through (6). Recall that p_W^C and p_S^C are independent of μ_W and μ_S . Now we show that, when $r^* \geq r^o$, there exist $\mu_W > 0$ and $\mu_S \geq \mu_W r^o$ that satisfy (A-75) and (A-76). The value of $\mu_W > 0$ that satisfies (A-75) is uniquely given by

$$\mu_W = \mu_W^o \equiv \left(\frac{a_W p_W^C}{\alpha} \right)^{\frac{1}{\alpha}} > 0. \quad (\text{A-77})$$

Define

$$\mathcal{K}(\mu_S) = \frac{\alpha \mu_S^{\alpha-1}}{a_S} - \frac{p_S^C - \tilde{p}}{\mu_S - (\tilde{p} \mu_W / p_W^C)}. \quad (\text{A-78})$$

For fixed μ_W , let $p_W^G(\mu_S)$ denote the weak type's probability of winning in the challenge configuration conditioned on μ_S . By Lemma A-5, for $\mu_S \geq \mu_W r^o$, $p_W^C \geq p_W^G(\mu_S)$. By (A-31), $p_S^C > \tilde{p}$. Thus, by (A-33), for any fixed μ_W , if $\mu_S \geq \mu_W r^o$, it must be that $\frac{p_S^C - \tilde{p}}{\mu_S - (\tilde{p} \mu_W / p_W^C)} > 0$. Thus, by

(A-78), it is clear that, fixing μ_W , $\mu_S \leftrightarrow \mathcal{K}(\mu_S)$ is strictly increasing. It is also obvious that $\mathcal{K}(\mu_S) \rightarrow \infty$ as $\mu_S \rightarrow \infty$. Also note that, when $\mu_S = \mu_W^o r^o$, where μ_W^o is given by (A-77), we must have

$$\mathcal{K}(\mu_S) = \mathcal{K}(\mu_W^o r^o) = \frac{p_W^C}{\mu_W^o} \left(\frac{(r^o)^{\alpha-1} a_W}{a_S} - \frac{p_S^C - \tilde{p}}{p_W^C r^o - \tilde{p}} \right), \quad (\text{A-79})$$

where the second equality follows from substituting $(\mu_W^o)^{\alpha-1}$ using equation (A-75). Because, by definition, $r^* = (a_S/a_W)^{\frac{1}{\alpha-1}}$,

$$r^* \geq r^o \implies \frac{(r^o)^{\alpha-1} a_W}{a_S} \leq 1. \quad (\text{A-80})$$

Because, by definition, $r^o = p_S^C/p_W^C$,

$$\frac{p_S^C - \tilde{p}}{p_W^C r^o - \tilde{p}} = 1. \quad (\text{A-81})$$

Thus, by (A-79), when $\mu_S = \mu_W^o r^o$, $\mathcal{K}(\mu_S) \leq 0$. Hence, given that $\mu_S \leftrightarrow \mathcal{K}(\mu_S)$ is strictly increasing for $\mu_S \geq \mu_W r^o$ and tends to infinity as $\mu_S \rightarrow \infty$, when $\mu_W = \mu_W^o$, there must uniquely exist a $\mu_S^o \geq \mu_W r^o$ such that $\mathcal{K}(\mu_S^o) = 0$. Thus, (A-75) and (A-76) are simultaneously satisfied with $\mu_W = \mu_W^o$ and $\mu_S = \mu_S^o$, where $\mu_S^o \geq \mu_W^o r^o$. By Lemma A-5 and the fact that $\mu_S^o \geq \mu_W^o r^o$, $p_W^C \geq p_W^G(\mu_S^o/\mu_W^o)$. Thus, the condition for the concession equilibrium is satisfied. The choices of $\mu_W = \mu_W^o$ and $\mu_S = \mu_S^o$ also satisfy each type's second-order condition and the individual rationality condition. Thus, there exists a concession equilibrium if $p_W^C \geq p_W^G(r^*)$. \square

Proof of Proposition 5. Let θ (θ') be the probability of being strong for every internal (external) candidate, where $\theta > \theta'$. Consider adding $n' > 0$ external candidates to the contest with $n > m$ internal candidates. Throughout, suppose that the contest with only the n internal candidates produces a challenge equilibrium.

We show that, in any ‘‘symmetric’’ equilibrium of the expanded contest (‘‘symmetric’’ in the sense that every type- $t \in \{S, W\}$ internal candidate plays the same strategy and every type- $t \in \{S, W\}$ external candidate plays the same strategy), the expanded contest has lower winner quality than the contest with only the n internal candidates.

Consider the expanded contest. Let \hat{p}_t and \hat{p}'_t be the equilibrium probability of winning for a type- t internal candidate and for a type- t external candidate, respectively, in the expanded contest, $t \in \{S, W\}$. Because a weak internal candidate always has the option of mimicking the strong internal type with probability μ_W/μ_S and choosing zero performance with the complementary probability, it must be that

$$\hat{p}_W \geq \frac{\mu_W}{\mu_S} \hat{p}_S = \frac{\hat{p}_S}{r}. \quad (\text{A-82})$$

Analogously,

$$\hat{p}'_W \geq \frac{\mu_W}{\mu_S} \hat{p}'_S = \frac{\hat{p}'_S}{r}. \quad (\text{A-83})$$

Let $\hat{\Pi}(n, n')$ be winner quality in the expanded contest. Note that

$$\begin{aligned} \hat{\Pi}(n, n') &= \frac{n\theta\hat{p}_S + n'\theta'\hat{p}'_S}{n\theta\hat{p}_S + n'\theta'\hat{p}'_S + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W} \\ &\leq \frac{n\theta\hat{p}_{Wr} + n'\theta'\hat{p}'_{Wr}}{n\theta\hat{p}_{Wr} + n'\theta'\hat{p}'_{Wr} + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W}, \quad (\text{A-84}) \end{aligned}$$

where the first line follows from the fact that winner quality equals the expected number of strong winners divided by the sum of expected number of strong winners and the expected number of weak winners, and the second line follows from (A-82), (A-83), and the fact that, for any fixed $b > 0$, $f(a) = \frac{a}{a+b}$ is increasing in a for $a > 0$. Let $\Pi(n)$ be winner quality in the contest with only n internal candidates. If the contest with only internal candidates has a challenge equilibrium, $\Pi(n)$ is given by (14). By (14) and (A-84), for any $\theta' < \theta$ and $n, n' > 0$,

$$\begin{aligned} \Pi'(n, n') - \Pi(n) &\leq \frac{n\theta\hat{p}_{Wr} + n'\theta'\hat{p}'_{Wr}}{n\theta\hat{p}_{Wr} + n'\theta'\hat{p}'_{Wr} + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W} - \frac{r\theta}{r\theta + 1 - \theta} = \\ &\frac{n'\hat{p}'_{Wr}}{(n\theta\hat{p}_{Wr} + n'\theta'\hat{p}'_{Wr} + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W)(r\theta + 1 - \theta)} (\theta' - \theta) < 0. \end{aligned}$$

Thus, winner quality in any symmetric equilibrium of the expanded contest is strictly lower than in the original contest. \square

Proof of Proposition 6. Let \bar{x} be the scoring cap. Applying the argument for the continuity of P in Section 2.2 to all $x \in [0, \bar{x})$ establishes the following result:

Lemma A-6. *In any symmetric equilibrium, the contest payoff function, P , intersects the origin and is continuous over $[0, \bar{x})$, where \bar{x} represents the scoring cap.*

Now suppose $\bar{x} \geq \mu_S$. Although a discontinuity of P can occur at \bar{x} , by Lemma A-6 and the fact that P is bounded above by $P(\bar{x})$, P is upper semicontinuous. Given that P is nondecreasing and bounded, the upper semicontinuity of P guarantees the existence of a best reply to P and makes the argument used in the proof of Lemma 1 applicable here. Hence, Lemma A-1 and equation (A-10) still hold, which implies that the support of contestants' performance distributions must still fall in the range where the probability of winning function, P , meets ψ , the concave lower envelope of the two upper support lines, $\{\alpha_t + \beta_t x\}_{t=S,W}$, associated with the two types of contestants.²⁷ In this case, there are still only two candidate equilibrium configurations: concession and challenge. Because the weak type's configuration-conditioned payoffs, expressed by equations (4) and (9), are determined by the configuration-conditioned selection rules that are unaffected by the scoring cap $\bar{x} \in [\mu_S, \infty)$, imposing such a scoring cap does not change any type's configuration-conditioned payoff. By the same argument used in the proof of Proposition 1 for showing that the equilibrium configuration is the one that favors the weak type, after imposing $\bar{x} \geq \mu_S$, the equilibrium configuration is still the one that favors the weak

²⁷With a scoring cap, tie might occur at the scoring cap. Because we have assumed a symmetric tie-breaking rule, there are still only two contestant types, distinguished by contest ability but not by the tie-breaking rule. Thus, there are still only two upper support lines, S-support line and W-support line.

type. Thus, given that each type's configuration-conditioned payoffs are unaffected by $\bar{x} \geq \mu_S$, neither the equilibrium configuration nor each type's equilibrium payoff is affected by $\bar{x} \geq \mu_S$.

To show that imposing a cap $\bar{x} \geq \mu_S$ reduces risk taking, note that the optimal dual variables, α_t and β_t , for $t \in \{S, W\}$, are constant in $\bar{x} \in [\mu_S, \infty)$.²⁸ This is because, as discussed above, the equilibrium configuration is unaffected by the cap $\bar{x} \in [\mu_S, \infty)$. The optimal dual variables in challenge equilibria are given by $\alpha_W = \beta_W = 0$ and $\beta_S = \beta_W = m/(n(\theta\mu_S + (1 - \theta)\mu_W))$ (see the argument in the proof of Lemma A-4), which are unaffected by the cap $\bar{x} \in [\mu_S, \infty)$. The optimal dual variables in concession equilibria are given by equations (A-17), (A-19), and (A-20) and $\alpha_W = 0$ (see the argument in the proof of Lemma A-3), which are again unaffected by the cap $\bar{x} \in [\mu_S, \infty)$. Thus, the optimal dual variables must be constant in $\bar{x} \in [\mu_S, \infty)$.

Next, note that constant optimal dual variables imply that the concave lower envelope, ψ , defined by (A-11), is constant in $\bar{x} \geq \mu_S$. Thus, when the cap is not binding, i.e., when $\bar{x} \geq \hat{x}$, where \hat{x} is defined by (A-13), equilibrium distributions are unaffected.

Now consider the case where the cap is binding, i.e., where $\bar{x} \in [\mu_S, \hat{x})$. Because the contest payoff function, P , is bounded above by ψ and because ψ is strictly increasing, we must have $P(\bar{x}) \leq \psi(\bar{x}) < \psi(\hat{x}) = 1$, which implies point mass on $x = \bar{x}$. Thus, by the random resolution of ties, P is discontinuous at $x = \bar{x}$. Moreover, because probability weight is placed only on points where P meets ψ , we must have

$$P(\bar{x}) = \psi(\bar{x}). \quad (\text{A-85})$$

Given that P is discontinuous at $x = \bar{x}$ while ψ is increasing and continuous, equation (A-85) implies

$$P(\bar{x}-) < \psi(\bar{x}-). \quad (\text{A-86})$$

Note that P must also meet ψ at some point $x' \in (0, \bar{x})$, because otherwise, P could only meet ψ at 0 and at $\bar{x} \geq \mu_S$, which, by Lemma A-1, would imply that weak contestants place point mass on 0, contradicting Lemma A-6. Thus, given that P meets ψ at some point $x' \in (0, \bar{x})$ and given the continuity of P over the interval $[0, \bar{x})$, Result A-3 and equations (A-85) and (A-86) imply the existence of $\tilde{x} \in (0, \bar{x})$ such that

$$P(x) = \begin{cases} \psi(x) & \text{if } x \in [0, \tilde{x}] \\ \psi(\tilde{x}) & \text{if } x \in [\tilde{x}, \bar{x}) \\ \psi(\bar{x}) & \text{if } x \geq \bar{x} \end{cases}. \quad (\text{A-87})$$

Thus, imposing the cap $\bar{x} \in [\mu_S, \hat{x})$ induces each contestant to transfer mass over $(\tilde{x}, \bar{x}]$ to the point mass on \bar{x} , leaving the mean of the distribution and the distribution over $[0, \tilde{x}]$ unchanged. In such a case, it must be that the unconditional performance distribution without the cap is an MPS of the one with the cap.

²⁸When $\bar{x} = \mu_S$, the optimal dual variables are not unique. However, S 's strategy is unique: S places all the mass on μ_S when $\bar{x} = \mu_S$. Thus, without loss of generality, we redefine the values of α_S and β_S when $\bar{x} = \mu_S$ by their limiting values when $\bar{x} \downarrow \mu_S$.

Finally, to show that imposing a cap $\bar{x} < \mu_S$ weakly increases the weak type's probability of winning and weakly reduces the strong type's probability of winning, simply note that, by a similar argument used above, there are again only two candidate equilibrium configurations, concession and challenge, and the equilibrium configuration is determined by the weak type's preference. Because each type's performance is capped by \bar{x} , imposing the cap $\bar{x} < \mu_S$ essentially reduces the strong type's contest ability (i.e., mean performance) relative to the weak type's, which reduces the strength asymmetry, r . Thus, given that the weak type's payoff in the challenge configuration is decreasing in r , imposing the cap $\bar{x} < \mu_S$ increases the weak type's payoff in the challenge configuration and reduces the strong type's payoff in the challenge configuration. Because each type's payoff in the concession configuration does not depend on the cap \bar{x} and the equilibrium configuration is the one that favors the weak type, it must be that imposing the cap $\bar{x} < \mu_S$ weakly increases the weak type's payoff and weakly reduces the strong type's payoff. \square