

Less competition, more meritocracy?*

Dawei Fang[†]
University of Gothenburg

Thomas Noe[‡]
University of Oxford

Abstract

Uncompetitive contests for grades, promotions, retention, and job assignments, which feature lax standards and limited candidate pools, are often criticized for being unmeritocratic. We show that, when contestants are strategic, lax standards and exclusivity can make selection more meritocratic. When many contestants compete for a few promotions, strategic contestants adopt high-risk strategies. Risk-taking reduces the correlation between performance and ability. Through reducing the effects of risk-taking, “Peter-Principle” promotion policies, which entail promoting some contestants that are unlikely to be worthy, can increase the overall correlation between selection and ability, and thus further the goal of meritocratic selection.

*We thank Li Chen, Qiang Fu, Ed Hopkins, Jingfeng Lu, John Quah, Christian Seel, Marco Serena, Ron Siegel, Philipp Strack, and seminar audiences at UC Berkeley, Chapman University, University of Gothenburg, National University of Singapore, Singapore Management University, Stockholm University, Econometric Society Asian Meeting 2019, and SAET 2017 and 2019 for helpful comments and suggestions. We would like to also thank the editor, Peter Kuhn, and two anonymous reviewers for the *Journal of Labor Economics* for extremely insightful comments which have greatly contributed to the presentation of our results. Dawei Fang gratefully acknowledges financial support from Browaldh Stiftelsen (BFh18-0010) and VINNOVA (grant 2010-02449). Earlier versions of the paper were circulated under the title “Selection contests are always clubby” and the title “Lowering the bar and limiting the field: The effect of strategic risk-taking on selection contests.”

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

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

1 Introduction

Meritocratic selection contests are a pervasive feature of modern life. Elite firms, like Google, fill entry level positions based on candidates' ranking in interviews which are essentially oral examinations aimed to identify cognitive ability (Popomaronis, 2019). University admissions around the world are profoundly affected by scores on scholastic aptitude tests. Investor fund selection is largely determined by league-table rankings of mutual funds compiled by rating services such as Morningstar (Guercio and Tkac, 2008; Armstrong et al., 2019).

The most important feature that distinguishes meritocratic selection contests from other contests is that the outcomes of these contests primarily affect contestants and society through their effects on *selection* rather than through the outputs the contests engender.¹ Selection is based on ranked performance and expected contestant performance is positively correlated with ability, where “ability” is defined as any trait desired by the selector, frequently the designer of the contest. These contests are “meritocratic” because selecting the most able is a defining feature of meritocracy.² Thus, increasing the correlation between selection and ability can be viewed as making contests more meritocratic.

The design of selection contests is frequently shaped by the perspective that competition and high standards make contests more meritocratic (Frost, 2017). In fact, “meritocratic society” is sometimes even treated as a synonym for “competitive society” (Ekins, 2014). However, this paper shows that, when success in a selection contest can be a product of strategic risk taking, meritocratic selection can often be furthered by anti-competitive policies such as low selection bars and restricted candidate fields.³

These implications are developed in a parsimonious model of contest design. In the model, n contestants compete for selection. The number of contestants selected is determined by the contest's *selection quota*, m . Contestants prefer selection to deselection. Each contestant is endowed with *ability*: strong contestants have high ability and weak contestants have low ability. Ability is private information. Selection is based on the ranking of contestants' performance (e.g., portfolio returns, examination scores). The expected performance of a contestant, which we term the contestant's performance *capacity*, is fixed by the contestant's ability. However, contestants can take risks, i.e., submit performance that is a mean-preserving spread of capacity.

¹In contrast, the objective of many contests, e.g., sales contests, crowdsourcing contests, is to motivate the production of economically valuable outputs, e.g., increased sales, innovative product designs.

²The definition of “meritocracy” we adopt is consistent with standard usage: “Meritocracy: a social system, society, or organization in which people get success or power because of their abilities, not because of their money or social position” (Cambridge English Dictionary). In some contexts, “meritocracy” can also be used to describe a “fair” process of selection, one that minimizes the effect of extraneous, random characteristics (e.g., caste, social networks, ethnicity). In our setting, the only characteristic that distinguishes contestants is ability and thus this second sense of “meritocracy” is not relevant.

³Risk-taking in competitions can take many forms depending on the nature of the competition: in fund manager competitions, increasing the unsystematic risk exposure (Chevalier and Ellison, 1997; Khorana, 2001; Kaniel and Parham, 2017), in weight-lifting competitions, attempting a very heavy lift (Genakos and Pagliero, 2012), in research grant competitions, fudging data, in job interviews, making up stories about professional experience, etc.

By adopting a high-risk strategy, weak contestants have a chance of besting strong contestants.⁴

The design of the contest, i.e., the number of contestants, n , and the number of quota places, m , is entrusted to a contest designer. In contrast to the existing literature, our focus is on the effect of competitiveness on selection efficiency rather than contestant performance.⁵ In the baseline model, the designer's objective function is symmetric, i.e., the gain from selecting a strong contestant equals the loss from selecting a weak contestant. The problem the designer faces is that ability is private information and must be inferred from contestant performance rank, which is affected by risk taking.

We first consider optimal selection quotas when the number of contestants is outside the designer's control, e.g., a promotion contest within a firm in which the firm can choose the promotion rate. Our analysis shows that, when the expected quality of the contestant pool is not too low, "Peter-Principle" selection—choosing a large selection quota under which some selected candidates are likely to be weak—can be optimal. Although selecting these unworthy candidates *per se* reduces meritocratic designer welfare, a lax selection standard, by increasing the probability that weak contestants will be selected, reduces weak contestants' incentives to adopt high-risk strategies to challenge strong contestants, thereby increasing the correlation between performance and ability.

Our result contrasts sharply with other rationales for over-promotion (e.g., Prendergast, 1992; Fairburn and Malcomson, 2001; Gürtler and Kräkel, 2010). These papers consider situations in which employers are willing to make selection less meritocratic, through over-promotion, in order to increase the incentive efficiency of compensation. We show that over-promotion can be motivated simply by the objective of making selection *more* meritocratic.

When the expected quality of the contestant pool is sufficiently low, lax selection standards can result in so many weak candidates being selected that selecting no candidates dominates the lax quota. At the same time, shrinking the quota increases risk taking, which also attenuates the correlation between performance and ability. In this case, designer welfare is maximized by setting a zero quota. In such cases, risk taking renders the contest mechanism an ineffective means of selecting able candidates out of an applicant pool that is weak on average. This ineffectiveness can be costly given the many relative advantages of contests over other selection mechanisms.⁶

⁴Risk-taking strategies have been modeled as mean-preserving spreads of a fixed performance level in various contexts, such as fund manager competitions (Seel and Strack, 2013; Fang and Noe, 2016; Strack, 2016), sales contests (Gaba and Kalra, 1999), R&D races (Dasgupta and Stiglitz, 1980; Klette and de Meza, 1986), political campaigns (Myerson, 1993), and status contests (Robson, 1992; Becker et al., 2005; Ray and Robson, 2012).

⁵See Moldovanu and Sela (2001), Fang et al. (2020), and Olszewski and Siegel (2020) for the analyses of the effects of competitiveness on contestant performance.

⁶A number of researchers have argued that contest selection is more advantageous than selection based on absolute performance when performance is hard to verify (Che and Gale, 2003) or affected by common time-varying shocks (Lazear and Rosen, 1981; Knoeber and Thurman, 1994), as well as when relative performance is easier to measure (Lazear and Rosen, 1981), when performance evaluation is delegated to lenient reviewers (Letina et al., 2020), and when firms have a strong preference for offering a fixed amount of total compensation to employees (Gürtler and Kräkel, 2010).

Next we consider the case where the selection quota is fixed but the designer can vary the number of contestants, e.g., a competition for a firm’s CEO position. In such contests, firms can limit the number of contestants either by excluding external candidates, or only including candidates with references from current employees, or, expand the number of contestants by encouraging external candidates. We show that, once the number of contestants increases beyond a threshold level, adding even more contestants does not increase the expected ability of selected contestants, and thus does not increase meritocratic welfare.

In contrast, in contest models that do not feature strategic risk taking, increasing the number of contestants generally increases the expected ability of selected contestants. For example, Ryvkin (2010, Corollary 3.1) shows that, in selection contests where contest noise is exogenous, expected winner ability is always increasing in contest size.⁷ In all-pay contests, where contestant performance equals contestant effort and effort costs depend on contestant ability, a contestant’s performance rank equals her ability rank if contestants are ex-ante homogeneous with ability being private information (Moldovanu and Sela, 2001), and approximates her ability rank if contest size is large (Olszewski and Siegel, 2016, 2020). Thus, in these contests, adding contestants also increases expected winner ability.⁸

After developing the baseline results, we extend our analysis in a number of directions. Our first extension considers the effect of asymmetric designer objective functions, functions that specify a gain from selecting strong contestants that need not equal the loss from selecting weak contestants. We show that the extended model’s predictions are qualitatively identical to the baseline model’s.

Next, we extend the analysis to endogenize contestant capacity. We assume that contestants must invest effort to produce performance capacity (e.g., students studying in order to increase their capacity to perform in an examination). Total and marginal costs of capacity acquisition are lower for contestants with higher ability. When capacity is endogenous, a change of contest design can potentially affect selection efficiency by affecting contestant capacity choice. However, we show that, in many cases (e.g., power effort cost functions), our results are robust.

In summary, our analysis predicts that, because of contestant risk taking, contest designers whose only goal is meritocratic selection, will frequently have no incentive to expand an already large contestant pool and, when choosing promotion and reward schemes, will never choose schemes that are expected to reject worthy candidates but frequently adopt schemes that are expected to select unworthy candidates. Thus, when contestants are strategic, meritocracy is consistent with contests with limited contestant pools and lax selection standards.

In our setting, often criticized policies such as soft grading curves can make selection more meritocratic. For the same reasons, motivational promotions by firms can be rationalized even when such “motivational promotions” have no motivational effect.⁹ In retention contests aimed

⁷While Ryvkin (2010) restricts the quota to one, his result extends to the case of an arbitrary fixed quota.

⁸However, the objective in these papers is not to increase expected winner ability. Thus, the increase in winner quality induced by larger contest size need not imply the optimality of increasing contest size in these papers.

⁹For a discussion of motivational promotion in the workplace, see Deeprose (2006).

to optimize workforce quality, our results predict that dismissal rates will frequently be far lower under risk taking than under a rate based simply on contestant ability.¹⁰

In addition to predicting the effects of actual policies, our results have implications for policy proposals such as “lottery admission,” a reform of elite university admission systems aimed to lower the psychological stress associated with the university application process (Schwartz, 2007). Under lottery admission, $m' > m$ applicants are selected for “approval” through a rank competition and m of those approved are allocated places by lottery. Our analysis provides conditions under which lottery admission will not make selection less meritocratic.

Our results are most relevant to real world settings where weak contestants can take chances that enable them to sometimes best strong contestants. Many one-shot, high-stakes contests (e.g., job interviews, college entrance exams) have these features. While, throughout the paper, we abstract from the motivational role of contests but focus purely on the selection role, we believe that the economic force identified in our paper will influence optimal designs of contests where contest designers care about both selection efficiency and contestant effort.

Related literature

This paper studies the effect of risk on selection. Its principal departure from the extant literature is that it models “endogenous” risk, risk generated by contestant strategies rather than risk generated by an exogenous noise term that mediates the relationship between ability and performance. It is known theoretically that contestants in a disadvantaged position have an incentive to adopt high-risk strategies in order to win (Aron and Lazear, 1990; Cabral, 2003). Such behavior is also widely observed empirically (Chevalier and Ellison, 1997; Khorana, 2001; Bothner et al., 2007; Beaudoin and Swartz, 2010; Genakos and Pagliero, 2012; Kirchler et al., 2018). Starting from this point, our paper shows that uncompetitive contest designs, by ameliorating risk taking, can produce more meritocratic selection outcomes.

Lazear (2004) studies promotion in a model where risk is generated by exogenous noise affecting the performance/ability relation. His analysis, like ours, identifies a Peter-Principle effect: because a component of superior worker performance is produced by luck, i.e., a high realization of a random noise term affecting worker performance, the expected future performance of promoted workers is less than their performance before promotion. However, the implications of our analysis differ quite dramatically from Lazear (2004). In Lazear (2004), the designer, realizing that part of worker performance is the product of luck, adjusts upward the performance cutoff required for promotion. Thus, in Lazear (2004), exogenous risk leads to an increase in promotion standards relative to the no-noise case. In our analysis, endogenous risk leads the designer to lower standards by expanding the selection quota relative to the quota which would have been set in the absence of risk.

¹⁰This conclusion appears to be consistent with empirical studies of dismissals in mutual funds. Khorana (1996, Table 4) finds that only 14% of mutual-fund managers in the lowest performance decile are replaced despite the fact that, as Khorana (2001) documents, replacing low-performing managers improves mutual-fund returns.

Like Lazear (2004), Ryvkin and Ortmann (2008) and Ryvkin (2010) also consider the effect of exogenous risk on selection. These papers fix the selection quota at one, and thus, unlike our paper, they do not address the effect of risk on the selection quota. However, they do consider the effect of expanding the contestant pool. A common conclusion they reach is that expanding the contestant pool always increases the expected ability of the winner. In contrast, we show that, when contestants are strategic risk takers, pool expansion beyond a threshold number of contestants never increases the expected ability of the winner(s).

Like Ryvkin and Ortmann (2008) and Ryvkin (2010), Hvide and Kristiansen (2003) also fixes the selection quota at one. However, Hvide and Kristiansen (2003) does take into account strategic risk taking. Hvide and Kristiansen find an example in which the expected ability of the winner is not uniformly increasing in the size of the contestant pool. This example is consistent with the idea that contestant risk taking can nullify the gains from pool expansion.

However, in Hvide and Kristiansen (2003), because contestants can only randomize between a safe and an exogenously specified risk-taking strategy, once all contestants switch to the risk-taking strategy, it is not possible for them to accommodate further pool expansion by further increasing risk. Thus, in Hvide and Kristiansen (2003), when pool size reaches the threshold at which all contestants play the risk-taking strategy, the effect of further pool expansion on winner quality is the same as the effect of pool expansion in exogenous-noise settings—i.e., it always increases the expected winner quality. As our paper shows, this result no longer holds when risk-taking strategies are not exogenously specified.

More generally, our paper is related to research showing that meritocracy can be furthered by seemingly unmeritocratic policies. Meyer (1991) and Kawamura and Moreno de Barreda (2014) find that biasing the contest selection mechanism toward certain contestants can increase selection efficiency.¹¹

2 An example

In this section, we provide an example that aims to capture the fundamental intuition underlying our analysis. The example is developed heuristically, and precise statements and discussion of the model's assumptions are deferred to the following sections. The example we consider is the problem faced by an administrator who has been tasked with setting the grade distribution for an examination. The exam will be taken by 4 students. Each student is of either a more able strong type or a less able weak type. A student's type is the student's private information and is independent of the types of the other students. Each student is strong with probability $1/2$ and weak with probability $1/2$. The distribution of student types and the number

¹¹In a paper titled "The Limits of Meritocracy," Morgan et al. (2018) model contests with exogenous noise; we model contests without exogenous noise; they define "meritocratic" to mean low contest noise; the designer's objective is effort maximization not meritocratic selection. The questions addressed in Morgan et al. (2018) are interesting and important. However, the only commonalities between Morgan et al. (2018) and our paper are that both papers model contests and both use the term "meritocracy."

of students taking the exam are common knowledge.

Expected exam performance is fixed by a student's type. For simplicity, we assume that strong students always receive full marks, 100. Weak students, in contrast, receive only half marks, 50, in expectation, but can possibly receive high or even full marks by taking risks. In practice, students can choose the intensity with which they revise potential exam questions. Thus, one high-risk strategy for a student is concentrating her limited capacity on studying (or memorizing) answers to a subset of potential questions. Using this strategy, a student will receive high marks if the exam questions happen to be drawn from this subset. Of course, if the student picks the wrong subset, her marks will be quite low.

We assume that a weak student can choose *any* mark distribution over $[0, 100]$ subject to the constraint that her expected marks equal 50. For example, a uniform mark distribution on $[0, 100]$ is feasible to a weak student, as is the mark distribution that produces full marks with probability $1/2$ and zero marks with probability $1/2$. We impose the zero and full marks bounds on performance because the context of this example is an examination. Although our analysis accommodates such bounds, they are not required to obtain our results, provided that a lower bound exists.¹² The assumption of strong students always receiving full marks simplifies our discussion by allowing us to focus on only weak students' risk-taking strategies. As we show later, this simplification is without loss of generality.

Based on exam performance, students are awarded an *A* or a *B*. Each student prefers receiving an *A* to receiving a *B*. The exam is "graded on a curve" which is set by the administrator before the exam is taken and is known by the students. The curve specifies a quota, m , for the number of *A* grades awarded. Grade awards are based on relative performance: if the administrator sets the curve at m , the m students with the highest marks receive *A* grades and the rest receive *B* grades. Ties are broken by fair randomization.

The administrator's only objective is to assign *A* grades to strong students and assign *B* grades (or equivalently, not assign *A* grades) to weak students. Consistent with this objective, we assume that the administrator sets the curve to maximize his welfare, u , where

$$u = \mathbb{E}[\# \text{Strong students receiving an } A - \# \text{Weak students receiving an } A].$$

Non-strategic students If weak students do not take risks and, hence, always receive half marks, strong students, who always receive full marks, will always have priority for receiving an *A*. When the administrator sets the curve, he does not know the number of strong students taking the exam but knows that each of the 4 students is strong with probability $1/2$. Thus, he knows that, when strong students are ranked ahead of weak students, the two highest ranked students are more likely to be strong than weak and the two lowest ranked students are more likely to be weak than strong. It is hence optimal for him to set $m = 2$, awarding an *A* to and only to the top half of exam performers.

¹²See Section 3 for further discussion.

Strategic students What happens if students are strategic? We base our analysis on symmetric equilibria, in which students of the same type play the same strategy. We refer to equilibria in which weak students concede to (i.e., have zero probability of besting) strong students as *concession equilibria* and equilibria in which weak students challenge (i.e., have a positive probability of besting) strong students as *challenge equilibria*. Clearly, any symmetric equilibrium must be either a concession equilibrium or a challenge equilibrium.

In this example, strong students' performance is fixed at full marks. Thus, in the example, concession equilibria refer to equilibria in which weak students place no point mass on full marks. For concession equilibria to exist, it must be the case that no weak student has an incentive to deviate to a high-risk strategy of receiving full marks with probability $1/2$ and zero marks with probability $1/2$. It is easy to show that this condition is not satisfied for $m = 1$ or $m = 2$. In particular, given m , if all weak students concede to strong students, a weak student can receive an A only if the number of strong rivals she faces, which we denote by s , falls below the number of A grades, m . When $s < m$, after the s strong students each receive an A, the number of remaining As is $m - s$. These As will be contested by the $4 - s$ weak students. In this case, given that, in a symmetric equilibrium, weak students play the same strategy and, thus, have the same probability of receiving an A, each weak student's probability of receiving an A equals $(m - s)/(4 - s)$. In contrast, if any weak student deviates to the prescribed high-risk strategy, then when all of her weak rivals place no point mass on full marks, she will outperform all of them whenever she receives full marks. When $s < m$, outperforming all weak rivals is sufficient for her to receive an A. Thus, given that the high-risk strategy gives her a probability of receiving full marks equal to $1/2$, when $s < m$, this deviation gives her a probability of receiving an A equal to at least $1/2$. Because, when $m = 1$ or $m = 2$, $1/2 \geq (m - s)/(4 - s)$ for all $s < m$, with strict inequality for some $s < m$, this deviation makes her strictly better off. Thus, there is no concession equilibrium when $m = 1$ or $m = 2$.

However, a similar argument does not apply to $m = 3$. In fact, as we will verify in Section 3.2, when $m = 3$, which corresponds to assigning A grades to 75% of the students in the example, a concession equilibrium does exist. Thus, a softer grading curve mollifies weak students' risk-taking incentives. Fixing m , when weak students concede, the number of strong students receiving an A equals $\min[m, \tilde{s}]$, where \tilde{s} denotes the total number of strong students, which, in the example, follows the Binomial(4, $1/2$) distribution, and the number of weak students receiving an A equals $m - \min[m, \tilde{s}]$. Thus, if $u(m)$ represents administrator welfare given the curve, m , $u(3) = \mathbb{E}[\min[3, \tilde{s}] - (3 - \min[3, \tilde{s}])] = 7/8$.¹³

In contrast, when $m = 1$ or $m = 2$, only challenge equilibria exist. One way for a weak student to challenge strong students is to adopt the high-risk strategy of receiving full marks with probability $1/2$ and zero marks with probability $1/2$. Because, as we will show in Section 3, in any equilibrium, students cannot receive an A by having zero marks, and also because, in this example, strong students always receive full marks, if a weak student adopts such a high-risk

¹³When $\tilde{s} \sim \text{Binomial}(4, 1/2)$, $\mathbb{E}[\min[3, \tilde{s}] - (3 - \min[3, \tilde{s}])] = -3 + 2 \sum_{s=0}^4 \binom{4}{s} (\frac{1}{2})^4 \min[3, s] = 7/8$.

strategy, her probability of receiving an A will equal half of the probability of receiving an A for a strong student. As we will verify in Section 3, in this example, it is indeed the case that, in any challenge equilibrium, a strong student has twice as much probability of receiving an A as a weak student. Hence, given that, in the example, ex ante, each student is equally likely to be strong or weak, if the curve, m , leads to a challenge equilibrium, then in expectation, out of the m students receiving an A , two thirds will be strong and one third will be weak. Thus, $u(1) = (2/3) - (1/3) = 1/3$ and $u(2) = 2 \times ((2/3) - (1/3)) = 2/3$. Hence, given that $u(3) = 7/8$, which is larger, $m = 3$ is the optimal grading curve.¹⁴

Recall that, when weak students do not act strategically, it is optimal for the administrator to set the curve at $m = 2$. Thus, our analysis reveals that the administrator responds to strategic risk-taking by “grade inflation.” Grade inflation comes at a cost. When $m = 3$, the fact that weak students concede to strong students implies that the probability that the student ranked the third highest is strong equals the probability that the total number of strong students, \tilde{s} , exceeds 2. Because $\tilde{s} \sim \text{Binomial}(4, 1/2)$, this probability is less than $1/2$. Thus, the student ranked the third highest is, in fact, more likely to be weak than strong and is, hence, not expected to merit an A .

However, this cost is dominated by the benefit of grade inflation: setting the soft $m = 3$ curve ensures that weak students have no incentive to challenge strong students, and thereby better aligns exam marks with student ability. As shown above, when students act strategically, the administrator’s welfare under $m = 2$, the optimal curve when students are non-strategic, is $u(2) = 2/3$, whereas, under $m = 3$, the optimal curve when students act strategically, the administrator’s welfare equals $u(3) = 7/8$. Thus, despite the cost of inflating A grades by 50%, grade inflation increases the administrator’s welfare by more than 30%. As the example shows, our meritocratic administrator, whose only goal is awarding A grades to strong students and B grades to weak students, can benefit from choosing grading curves that are expected to award A grades to students who do not merit an A . Reducing competition, by setting a soft curve, furthers meritocratic selection.¹⁵

3 Risk-taking

3.1 Risk-taking selection contests

Consider a contest with $n \geq 2$ contestants; m of them will be *selected* to fill a *place*, and the remaining $n - m$ contestants will be *deselected* and not receive a place, where $0 < m < n$.

¹⁴It is obvious that $u(0) = 0$ and, given that each student is equally likely to be strong or weak, $u(4) = 0$. Thus, the curve, $m = 3$, is optimal even if the administrator has the option of giving the students the same grade regardless of exam performance.

¹⁵This result sharply contrasts with Chan et al. (2007), which models grade inflation as an equilibrium outcome of a signaling model in which the sender is a university with private information about student quality, the signal is the grading curve, and the receivers are potential student employers. In Chan et al. (2007), the designer (the university administration) aims to maximize student compensation, not meritocratic grade assignment, and grade inflation makes selection less meritocratic.

The number of places, m , which we call the *selection quota*, and the number of contestants, n , which we call *contest size*, are fixed before the contest and are common knowledge.

The contestants are of two possible types, t : strong, S , and weak, W . Types are drawn independently of each other 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.

Selection is based on performance in the contest. Each type- $t \in \{S, W\}$ contestant's expected performance is fixed at μ_t , where μ_t represents a type- t contestant's *capacity*. Strong contestants have higher capacity to perform, i.e., $\mu_S > \mu_W$. While a contestant's mean performance is fixed by her capacity, the contestant can choose the "randomness" of her performance by undertaking risky actions. We assume that, subject to the *capacity constraint* on mean performance, each contestant can choose any distribution of nonnegative performance.

The capacity constraint restricts only the first moment of the contestants' performance distributions. In Online Appendix E, we show that our results extend to risk-taking contests with more general capacity constraints which restrict other moments, such as variance. Such constraints can also accommodate cases where mean performance is reduced by risk taking.

The assumption that the feasible performance range is $[0, \infty)$ is made largely to simplify the exposition of our results. In practice, the feasible performance range will be determined by the nature of the technology used to measure performance, e.g., in our examination example, it is not possible to earn less than zero marks or more than full marks, so performance has exogenous upper and lower bounds. In a contest between fund managers, where performance is measured by returns, there is no exogenous upper bound on performance and the lower bound on performance is -100% .

However, provided that the bounds on performance do not rule out risk-taking or make satisfaction of the capacity constraint impossible, these bounds have little qualitative effect on the analysis. A contest with a nonzero lower bound, \underline{x} , satisfying $-\infty < \underline{x} < \mu_W$ given capacities $\mu_t, t \in \{S, W\}$, is strategically equivalent to a contest with a zero lower bound given capacities $\mu_t - \underline{x}$. For this reason, we normalize the lower bound to 0 in all of our analysis.

The upper bound on performance can affect the qualitative properties of equilibrium strategies. However, the focus of our analysis is on the effect of contest structure on selection, and, as we show in Online Appendix C, imposing an exogenous upper bound on performance does not change contestants' equilibrium selection probabilities provided that the upper bound is no less than a strong contestant's capacity, μ_S .

Contestant performance distributions determine the outcome of the contest: each contestant's realized performance is independently drawn from her performance distribution. The m contestants with the highest realized performances are selected and the remaining contestants are deselected. Ties are broken by fair randomization. Contestants are expected utility maximizers who strictly prefer selection to deselection. Thus, given that risk-taking is costless, each contestant chooses a performance distribution to maximize her *probability of winning* a place

given her rivals' strategies and the contest's parameters.¹⁶

3.2 Equilibria

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

A contestant's probability-of-winning function maps the contestant's realized performance, x , to her probability of being selected and is determined endogenously by her rivals' strategies. Because each contestant faces the same distribution of rival types and strategies are symmetric, all of the contestants face the same *probability-of-winning function*, $P: \mathbb{R}_+ \rightarrow [0, 1]$. Let $\text{Supp } P$ be the support of P . Because P increases and only increases at points in the support of at least one type's performance distribution, we have

$$\text{Supp } P = \text{Supp}_W \cup \text{Supp}_S. \quad (1)$$

To facilitate our equilibrium derivation, we make the following claim:

Claim 1. In any symmetric equilibrium, the probability-of-winning function, P , is (a) continuous and (b) has an interval support, $[0, \hat{x}]$, where $\hat{x} < \infty$ is endogenous.

Because many authors have established properties analogous to Claim 1 in symmetric fair-gambles contests and symmetric all-pay auctions (Barut and Kovenock, 1998; Fang and Noe, 2016), our verification of Claim 1 is not very original. So, we relegate this verification to Online Appendix A. In contests where performance is exogenously bounded above by \bar{x} , where $\bar{x} \geq \mu_S$, e.g., in the case of a student examination we saw in Section 2, contestants may place point mass on the upper bound, \bar{x} , in a symmetric equilibrium. In these contests, P may not be continuous at the upper bound. Nevertheless, as we show in Online Appendix C, excluding the upper bound, P is continuous and has an interval support starting from 0.

Equation (1) and Claim 1 immediately imply that

$$\text{Supp}_W \cup \text{Supp}_S = [0, \hat{x}]. \quad (2)$$

To proceed with our equilibrium derivation, a few definitions are required: for any two points $(x_1, p_1), (x_2, p_2)$ in \mathbb{R}_+^2 , define the *interval between the points*, $[(x_1, p_1), (x_2, p_2)]$ by

$$[(x_1, p_1), (x_2, p_2)] = \{\lambda(x_1, p_1) + (1 - \lambda)(x_2, p_2) : \lambda \in [0, 1]\}.$$

A *gamble* between performance levels x' and x'' represents a performance distribution that randomizes between x' and x'' . A *fair gamble* between x' and x'' for a contestant of type t is a gamble between x' and x'' with the property that the probability of choosing x' , π , satisfies $\pi x' +$

¹⁶Because a contestant faces a binary outcome, winning a place or winning no place, as long as she is an expected utility maximizer, she always aims to maximize her probability of winning regardless of the curvature of her utility function.

$(1 - \pi)x'' = \mu_t$. Because fair gambles are feasible performance distributions, if performance levels x' and x'' are in the support of the equilibrium performance distribution of type t , then a type- t contestant's payoff from a fair gamble between x' and x'' equals her equilibrium payoff.¹⁷ Next, note that, for performance levels x_1, x_2 , and x_3 satisfying $x_1 < \mu_t < x_2$ and $x_1 < \mu_t < x_3$, if $(P(x_3) - P(x_1))/(x_3 - x_1) < (P(x_2) - P(x_1))/(x_2 - x_1)$, the interval $[(x_1, P(x_1)), (x_3, P(x_3))]$ lies below the interval $[(x_1, P(x_1)), (x_2, P(x_2))]$. Because $\mu_t \in [x_1, x_2] \cap [x_1, x_3]$, this implies that the payoff to a type- t contestant from a fair gamble between x_1 and x_2 exceeds the payoff from a fair gamble between x_1 and x_3 . This result is illustrated by Figure 1.

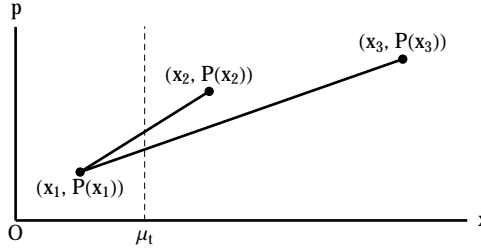


Figure 1: *Fair gambles and best replies.* In the figure, for a contestant of type $t \in \{S, W\}$, the payoff from a fair gamble between x_1 and x_3 , given by the intersection of the dashed line and the interval, $[(x_1, P(x_1)), (x_3, P(x_3))]$, yields a lower payoff than a fair gamble between x_1 and x_2 , given by the intersection of the dashed line and the interval, $[(x_1, P(x_1)), (x_2, P(x_2))]$.

Because all fair gambles in the support of a type- t contestant's performance distribution must produce the same payoff, x_3 cannot be in the support of t 's performance distribution if x_1 and x_2 are in its support. Similarly, if $(P(x_3) - P(x_1))/(x_3 - x_1) > (P(x_2) - P(x_1))/(x_2 - x_1)$, x_2 cannot be in the support of t 's performance distribution if x_1 and x_3 are in the support. Thus, the slope of the line joining any two points, $(x_1, P(x_1))$ and $(x_2, P(x_2))$, in the support of t 's performance distribution is constant. Hence, all performance/probability-of-winning pairs, $(x, P(x))$, such that $x \in \text{Supp}_t$ are collinear.¹⁸ Also note that, P must be concave over its interval support, $[0, \hat{x}]$. Otherwise, there would exist an interval $[x', x''] \subseteq [0, \hat{x}]$ over which P is convex and nonlinear. In this case, by Jensen's inequality, for any contestant who places weight over the interior of the interval, $[x', x'']$, a mean-preserving spread that moves the weight from the interior to the two endpoints of the interval, $[x', x'']$, would make her strictly better off. Thus, no performance level in the interior of $[x', x''] \subseteq [0, \hat{x}]$ could lie in the support of either type's performance distribution, contradicting equation (2). The next lemma summarizes these findings. We relegate all of the formal proofs to Online Appendix A.

Lemma 1. *In any symmetric equilibrium, for each type- $t \in \{S, W\}$ contestant, a fair gamble between x' and x'' , $x', x'' \in \text{Supp}_t$, is a best reply, and all performance/probability-of-winning pairs, $(x, P(x))$, such that $x \in \text{Supp}_t$ are collinear, and P is concave over its support.*

¹⁷The continuity of P implies that the set of optimal fair gambles is closed. Thus, the support of a contestant's equilibrium performance distribution is contained in the set of optimal fair gambles for the contestant.

¹⁸Collinearity for the single performance level where $x = \mu_t$ follows from the continuity of P .

Now consider the implications of Lemma 1 for equilibrium behavior. First, note that, by equation (2), $\text{Supp}_W \cup \text{Supp}_S$ is a connected closed set. The connectedness of $\text{Supp}_W \cup \text{Supp}_S$ and the fact that, by definition, both Supp_W and Supp_S are closed imply that the intersection, $\text{Supp}_W \cap \text{Supp}_S$, is not empty.

If the intersection consists of a single point, say \tilde{x} , then Supp_W and Supp_S must be adjacent closed intervals meeting at \tilde{x} , and because $\mu_S > \mu_W$, Supp_S must lie above Supp_W . In this case, the equilibrium is a concession equilibrium.¹⁹ The collinearity condition stated in Lemma 1 then implies that P is linear over Supp_W and over Supp_S with a possible kink at \tilde{x} . The concavity of P implies that the slope of P over Supp_W is no less than the slope of P over Supp_S . By Claim 1, P is continuous and thus $P(0) = 0$. These observations imply that, in any concession equilibrium, there exist $\hat{x} > \tilde{x} > 0$ such that $\text{Supp}_W = [0, \tilde{x}]$ and $\text{Supp}_S = [\tilde{x}, \hat{x}]$; P meets the origin and is concave over its support, $[0, \hat{x}]$, increasing over $[0, \hat{x}]$, linear over $[0, \tilde{x}]$, and linear over $[\tilde{x}, \hat{x}]$. Figure 2.A illustrates the satisfaction of these conditions.

If the intersection, $\text{Supp}_W \cap \text{Supp}_S$, contains more than one point, say at least the distinct points x' and x'' , then the probability that a weak contestant's performance exceeds a strong contestant's is positive. In this case, the equilibrium is a challenge equilibrium. The collinearity condition then implies that all pairs, $(x, P(x))$, such that $x \in \text{Supp}_t$, $t \in \{S, W\}$, are collinear with the two points, $(x', P(x'))$ and $(x'', P(x''))$. Thus, all pairs, $(x, P(x))$, such that $x \in \text{Supp}_W \cup \text{Supp}_S$ are collinear. Hence, by equations (1) and (2), P is linear over its support, $[0, \hat{x}]$. The continuity of P given by Claim 1 implies that $P(0) = 0$. These observations imply that, in any challenge equilibrium, there exists \hat{x} such that $\text{Supp}_W \cup \text{Supp}_S = [0, \hat{x}]$; P is linear over $[0, \hat{x}]$ and meets the origin. Figure 2.B illustrates the satisfaction of these conditions.

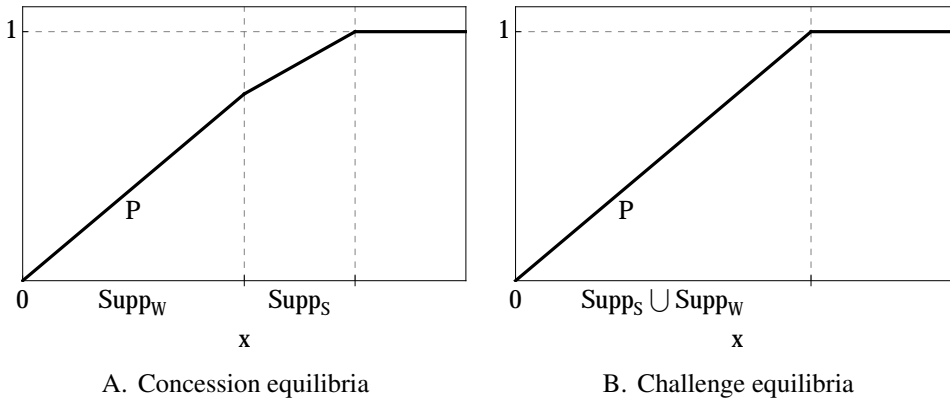


Figure 2: The possible forms of the probability-of-winning function, P , in a symmetric equilibrium.

For $t \in \{S, W\}$, let p_t be a type- t contestant's equilibrium probability of winning. Two facts are useful for deriving p_t . First, note that the expected number of selected contestants must equal the selection quota, m . Thus, in any symmetric equilibrium, p_S and p_W must satisfy

$$n(\theta p_S + (1 - \theta) p_W) = m. \quad (3)$$

¹⁹The continuity of P implies that no contestant places any mass on \tilde{x} . Thus, a weak contestant has zero probability of besting strong contestants even if her performance reaches the lower bound of Supp_S , \tilde{x} .

Second, by Lemma 1, P is concave. Thus, by Jensen's inequality, for each type- t contestant, choosing a deterministic performance level equal to her capacity level, μ_t , is a weakly best reply. Hence, we can evaluate each type- t contestant's equilibrium probability of winning simply by evaluating P at μ_t , i.e., in any symmetric equilibrium,

$$\text{for } t \in \{S, W\}, \quad p_t = P(\mu_t). \quad (4)$$

Equation (4), combined with the fact that, in challenge equilibria, P is linear over its support and meets the origin, implies that, in any challenge equilibrium (if exists), the ratio between strong and weak contestants' probabilities of winning equals their capacity ratio, i.e.,

$$\text{in any challenge equilibrium,} \quad p_S/p_W = \mu_S/\mu_W. \quad (5)$$

Equations (3) and (5) then imply that

$$\text{in any challenge equilibrium,} \quad p_W = p_o^G := \frac{m}{n} \left(\frac{\mu_W}{\theta\mu_S + (1-\theta)\mu_W} \right). \quad (6)$$

In contrast, in concession equilibria, P is concave over its support and meets the origin. Thus, by equation (4) and the fact that $\mu_S > \mu_W$, in any concession equilibrium (if exists), $p_S/p_W \leq \mu_S/\mu_W$. Hence, by equation (3) and the definition of p_o^G given in equation (6),

$$\text{in any concession equilibrium,} \quad p_W \geq p_o^G. \quad (7)$$

To determine the value of p_W in concession equilibria, simply note that, in concession equilibria, weak contestants can be selected only if all strong contestants are selected. Thus, if \tilde{S}_n denotes the number of strong contestants out of the n contestants, then when weak contestants concede, the number of weak selected contestants equals $\max[m - \tilde{S}_n, 0]$. Because, in any symmetric equilibrium, the expected number of weak selected contestants is $n(1-\theta)p_W$, it must be the case that $n(1-\theta)p_W = \mathbb{E}[\max[m - \tilde{S}_n, 0]]$ in concession equilibria. Therefore,

$$\begin{aligned} \text{in any concession equilibrium,} \quad p_W = p_o^C &:= \frac{\mathbb{E}[\max[m - \tilde{S}_n, 0]]}{n(1-\theta)} \\ &= \frac{1}{n} \sum_{s=0}^{m-1} \binom{n}{s} \theta^s (1-\theta)^{n-s-1} (m-s), \end{aligned} \quad (8)$$

where the last equality follows from the fact that $\tilde{S}_n \sim \text{Binomial}(n, \theta)$.

Given the above observations, to compute the equilibrium winning probability for each type of contestant, we only need to determine which type of equilibrium, concession or challenge, exists. Note that, in any challenge equilibrium, weak contestants can possibly be selected before any strong contestant is selected. Thus, in any challenge equilibrium, the expected number of weak selected contestants, $n(1-\theta)p_W$, must be greater than $\mathbb{E}[\max[m - \tilde{S}_n, 0]]$. Using the definition of p_o^C given in equation (8), we can express this condition as

$$\text{in any challenge equilibrium,} \quad p_W > p_o^C. \quad (9)$$

Equations (6) and (9) imply that challenge equilibria exist only if $p_o^G > p_o^C$. Equations (7) and (8) imply that concession equilibria exist only if $p_o^C \geq p_o^G$. Because these two necessary condi-

tions, $p_o^G > p_o^C$ and $p_o^C \geq p_o^G$, are complementary and because, as we show in Online Appendix B, a symmetric equilibrium always exists, these necessary conditions are also sufficient conditions. The next result, which characterizes the equilibrium winning probability for each type of contestant, is thus straightforward from the above analysis.

Lemma 2. *Concession (challenge) equilibria exist if and only if $p_o^C \geq p_o^G$ ($p_o^G > p_o^C$), where p_o^C and p_o^G are defined in equations (8) and (6), respectively. A weak contestant's equilibrium probability of winning, p_W , is given by*

$$p_W = \max \left[p_o^C, p_o^G \right].$$

A strong contestant's equilibrium probability of winning, p_S , is determined by p_W through (3).

As we show in Online Appendix C, Lemma 2 holds even if we impose an exogenous upper bound, \bar{x} , on contest performance, provided that $\bar{x} \geq \mu_S$. Thus, we can apply Lemma 2 to verify equilibria for the examination example studied in Section 2. In that example, $\mu_S = \bar{x} = 100$, $\mu_W = 50$, $\theta = 1/2$, and $n = 4$. Thus, by equations (6) and (8), if $m = 1$, $p_o^G = 1/6 > 1/32 = p_o^C$; if $m = 2$, $p_o^G = 1/3 > 3/16 = p_o^C$; if $m = 3$, $p_o^G = 1/2 < 17/32 = p_o^C$. Hence, by Lemma 2, in equilibrium, weak contestants concede to strong contestants if $m = 3$, whereas if $m = 1$ or $m = 2$, weak contestants challenge strong contestants and, by equation (5), in challenge equilibria, $p_S/p_W = \mu_S/\mu_W = 2$. The next result shows that, consistent with the example, increasing contest competitiveness, through either a decrease in the selection quota, m , or an increase in contest size, n , inclines weak contestants to challenge strong contestants.

Lemma 3. *For any given parametrization of a contest,*

- i. if challenge equilibria exist, then everything else being equal, challenge equilibria also exist if the selection quota, m , decreases or contest size, n , increases;*
- ii. if concession equilibria exist, then everything else being equal, concession equilibria also exist if the selection quota, m , increases or contest size, n , decreases.*

Lemma 3 is intuitive. Reducing the selection quota or increasing contest size makes it less likely that besting only weak rivals is sufficient for winning a place, thereby increasing weak contestants' incentives to challenge strong contestants. Because Lemmas 2 and 3 provide a characterization of equilibria that is sufficient for our following analysis of the effect of risk-taking on meritocratic selection, we defer the construction of equilibrium performance distributions to Online Appendix B.²⁰

4 Risk-taking and the selection quota

In this section, we consider how contestant risk-taking affects a meritocratic designer's choice of the selection quota, m . For this analysis, we fix contest size, n , and apply the same

²⁰The construction of equilibrium strategies is given by Result B-1 for concession equilibria and by Result B-2 for challenge equilibria in Online Appendix B.

meritocratic designer welfare function adopted in Section 2 to more general contexts—we assume that the designer sets the quota, m , to maximize his welfare, u , where

$$u = \mathbb{E}[\text{\#Strong selected contestants} - \text{\#Weak selected contestants}]. \quad (10)$$

We allow for the choices of $m = 0$ (full deselection) and $m = n$ (full selection). When $m = 0$ or $m = n$, despite no competition, designer welfare is well defined, and its calculation is straightforward and does not depend on agent behavior.

Note that, when n is fixed, the expected number of weak contestants is fixed, implying that the sum of the expected number of weak selected contestants and the expected number of weak deselected contestants is fixed. Inspection of (10) then shows that we can also interpret the designer’s problem as a task-assignment problem—the designer sets m to maximize

$$\mathbb{E}[\text{\#Strong selected contestants} + \text{\#Weak deselected contestants}].$$

In this problem, the designer assigns a fixed pool of contestants either to a more desirable “selection task” or a less desirable “deselection task.” The marginal product of strong (weak) contestants is higher when performing the selection (deselection) task. The designer’s problem is to fix the performance rank required for “promotion” to more desirable task.

Non-strategic contestants When contestants are non-strategic, weak contestants, because of their lower capacity, never best strong contestants. Hence, strong contestants are prioritized for selection. In this case, if \tilde{S}_n represents the number of strong contestants out of the n contestants, the number of strong selected contestants is $\min[m, \tilde{S}_n]$ and the number of weak selected contestants is $m - \min[m, \tilde{S}_n]$. Thus, if $\bar{u}(m)$ denotes designer welfare under the quota, m , with non-strategic contestants, $\bar{u}(m)$ can be expressed as

$$\bar{u}(m) = \mathbb{E}[\min[m, \tilde{S}_n] - (m - \min[m, \tilde{S}_n])] = 2\mathbb{E}[\min[m, \tilde{S}_n]] - m, \quad (11)$$

where $\tilde{S}_n \sim \text{Binomial}(n, \theta)$. Thus,

$$\begin{aligned} \bar{u}(m+1) - \bar{u}(m) &= 2 \left(\mathbb{E}[\min[m+1, \tilde{S}_n] - \min[m, \tilde{S}_n]] \right) - 1 \\ &= 2\mathbb{P}[\tilde{S}_n \geq m+1] - 1 = 2(1 - \mathbb{P}[\tilde{S}_n \leq m]) - 1 = 2(1/2 - B(m; n, \theta)), \end{aligned} \quad (12)$$

where $B(\cdot; n, \theta)$ denotes the CDF of the $\text{Binomial}(n, \theta)$ distribution. Inspection of (12) reveals that increasing the quota from m to $m+1$ increases (reduces) designer welfare, \bar{u} , if $B(m; n, \theta) < (>) 1/2$, whereas if $B(m; n, \theta) = 1/2$, such a quota increase does not change designer welfare. Thus, given that $m \mapsto B(m; n, \theta)$ is increasing, it is optimal to set the quota, m , at

$$m_M^*(n, \theta) = \min\{m \in \{0, 1, \dots, n\} : B(m; n, \theta) > 1/2\}, \quad (13)$$

which is a median of the $\text{Binomial}(n, \theta)$ distribution.

Binomial distributions have either one or two medians. When the $\text{Binomial}(n, \theta)$ distribution has one median, m_M^* represents the unique median. When the $\text{Binomial}(n, \theta)$ distribution has two medians, m_M^* represents the larger median. In the latter case, the smaller median equals $m_M^* - 1$, and the smaller median satisfies $B(m_M^* - 1; n, \theta) = 1/2$, which, by equation (12), implies

that both the smaller median, $m_M^* - 1$, and the larger median, m_M^* , are optimal. For expositional convenience, in the subsequent analysis, we assume that, whenever the designer is indifferent between two quotas, he chooses the larger quota.²¹ Under this assumption, the quota selected by the designer when contestants are non-strategic simply equals m_M^* .

Strategic contestants How does contestant risk-taking affect the designer's quota choice? To make our analysis non-trivial, we impose two assumptions for our subsequent analysis.

Assumption 1. The optimal quota when contestants are non-strategic is interior, i.e., $0 < m_M^* < n$.

If Assumption 1 is violated, then, when contestants are non-strategic, it is optimal for the designer to either fully deselect ($m_M^* = 0$) or fully select ($m_M^* = n$). Because introducing risk taking never increases designer welfare and also because designer welfare under full deselection or full selection is unaffected by risk taking, if Assumption 1 is violated, introducing risk taking has no effect on the designer's quota choice.

Assumption 2. Challenge equilibria exist when $m = 1$.

If Assumption 2 is violated, then even under the most competitive contest quota, $m = 1$, weak contestants concede to strong contestants. In this case, by Lemma 3, over all contest quotas, $m = 1, \dots, n - 1$, weak contestants concede. Therefore, risk taking has no effect on the designer's optimal quota choice if Assumption 2 is violated.

In contrast, as the following theorem implies, under Assumptions 1 and 2, depending on the prior quality of contestants, a meritocratic designer either tends to accommodate risk-taking by “inflating” the quota or shuts down competition by not running a contest.

Theorem 1. Suppose contest size, n , is fixed but the designer can vary the quota, m . Let p_o^C and p_o^G be defined by equations (8) and (6), respectively. Then

i. there exists \bar{m} , defined as

$$\bar{m} = \max \left\{ m \in \{1, \dots, n - 1\} : p_o^C(m) < p_o^G(m) \right\},$$

such that, over all contest quotas, $m = 1, \dots, n - 1$, challenge equilibria exist if and only if $m \leq \bar{m}$, and concession equilibria exist if and only if $m > \bar{m}$.

ii. Assuming

$$(1 - \theta)/\theta \leq \mu_S/\mu_W \tag{14}$$

holds, if $\bar{m} < m_M^*$, where m_M^* is given by (13), then the optimal quota is m_M^* , whereas if $\bar{m} \geq m_M^*$, then the optimal quota is \bar{m} or $\bar{m} + 1$.

iii. If condition (14) is violated, the optimal quota is zero.

²¹ All of our results hold if we instead assume that, whenever the designer is indifferent between two quotas, he chooses the smaller quota.

Part (i) of the theorem follows from Assumption 2 and the implication of Lemma 3 that reducing the quota inclines weak contestants to challenge strong contestants.

Part (ii) shows that, when the prior odds that a given contestant is weak, $(1 - \theta)/\theta$, are low relative to the asymmetry in contestant capacity, measured by the capacity ratio, μ_S/μ_W , a meritocratic designer tends to *inflate the quota*, setting quotas greater than the optimal quota in the absence of risk taking, m_M^* . Quota inflation is driven by two factors. First, quota inflation mollifies weak contestants' risk-taking incentives, making performance a better signal of ability. Second, risk taking reduces the correlation between ability and contest performance, thereby reducing the quality of top performers and increasing the quality of mediocre performers. When the prior odds of a contestant being weak are fairly small, this effect encourages the designer to dip deeper into the contestant pool by inflating the quota.

In the case described by part (iii), the prior odds of a contestant being weak are large relative to the capacity asymmetry between strong and weak contestants. Low capacity asymmetry makes it easier for weak contestants to challenge strong contestants through risk taking. Risk taking, given the high prior odds of a contestant being weak, makes even top performers unworthy of selection. Consequently, it is optimal not to run a contest but to set a zero quota.

Note that, when contestants are non-strategic, setting a zero quota is never optimal if contest size, n , is sufficiently large. In contrast, the condition for the optimality of a zero quota when contestants are strategic (i.e., $(1 - \theta)/\theta > \mu_S/\mu_W$) is independent of n . Thus, part (iii) implies that risk taking blocks using highly selective contests to pluck a few high-ability candidates out of a large but weak candidate pool.²²

5 Risk-taking and contest size

In this section, we consider the effect on designer welfare of varying contest size, n , given a fixed selection quota, m . The number of strong selected contestants and weak selected contestants sums to m , which is fixed by the analysis in this section. Inspection of equation (10) shows that, when m is fixed, the designer's problem is equivalent to maximizing the expected number of strong selected contestants, i.e., the quality of the m selected contestants.

Non-strategic contestants The effect of varying n when contestants are non-strategic is easy to identify. When contestants are non-strategic, strong contestants are prioritized for selection. Suppose we add a new contestant to the contestant pool. If the added contestant is strong, and if before the contestant's addition, less than m contestants were strong, the new contestant will be selected, and the number of strong selected contestants will increase. Otherwise, i.e., if the new contestant is weak or the selection quota has already been filled by strong contestants, the number of strong selected contestants will not change. Because the ability of each contestant

²²In Online Appendix I, we show that, when the total number of strong contestants in the contestant pool is fixed and is common knowledge and when, from the designer's view, the contestants are ex ante identical, it is still the case that contestant risk taking inclines the designer either to inflate the quota or not to run a contest.

is drawn independently, the probability that the pool contains less than m strong contestants is always positive. Thus, when contestants are non-strategic, adding contestants increases the expected number of strong selected contestants and, hence, designer welfare.

Strategic contestants Now consider the effect of varying n when contestants are strategic. As shown by Lemma 3, increasing n inclines weak contestants to adopt high-risk strategies that challenge strong contestants. While expanding the contestant pool increases the expected number of strong candidates in the pool, because of increased risk taking, the expected number of strong selected contestants need not increase. In fact, as our next theorem implies, when contest size is sufficiently large, further increasing size does not change the expected number of strong selected contestants.

Theorem 2. *Suppose the selection quota, m , is fixed but the designer can vary contest size, n .*

- i. There exists n^c such that challenge equilibria exist for all $n > n^c$.*
- ii. If a challenge equilibrium exists at $n = n'$, designer welfare at any $n > n'$ equals designer welfare at $n = n'$.*

The basic implication of Theorem 2 is that risk taking caps the gains from expanding the contestant pool. When pool expansion is costly because of outreach, advertisement, or search costs, the optimal contest size when contestants are strategic risk-takers will tend to be smaller than when contestants are non-strategic. As we show in Online Appendix D, even when increasing contest size is costless, when the pool of potential new contestants is, on average, of lower quality than the incumbent candidate pool, a meritocratic designer may strictly gain from excluding the potential contestants from the contest. In such cases, the gain from expanding the pool produced by increasing the expected number of strong candidates is overwhelmed by the cost of increased risk taking. In contrast, when contestants are non-strategic, the designer always strictly gains from pool expansion because adding contestants increases the expected number of strong candidates.

Thus, when contestants are strategic, even meritocratic designers, who are not biased toward specific candidates, have little incentive to expand candidate fields and sometimes will deliberately restrict consideration to candidates who, ex-ante, look promising, even if considering a wider field is costless. Our analysis also implies that the optimality of opening an in-house competition to external candidates depends on whether the in-house competition is soft and whether the external candidates are expected to be stronger than internal candidates, but does not depend on the number of potential external candidates.

6 Extensions and applications

6.1 Asymmetric designer objective

In Sections 4 and 5, we investigated the effect of risk-taking on the design of selection contests when a meritocratic designer has a symmetric welfare function given by (10). In what

follows, we show that the qualitative nature of our analysis extends to the case where designer welfare, u , has the following asymmetric specification:

$$u = \mathbb{E}[(1 - \sigma) \times \# \text{Strong selected contestants} - \sigma \times \# \text{Weak selected contestants}], \quad (15)$$

where $0 < \sigma < 1$.

Allowing for asymmetry between gains from selecting strong contestants and losses from selecting weak contestants does not change any of our result in Section 5, where the designer can vary only contest size, n . This is because the number of weak selected contestants equals the quota, m , less the number of strong selected contestants, implying that the designer's problems under the two objective functions, (10) and (15), are equivalent for any fixed m .

Now consider how introducing asymmetry into the designer's objective function might affect our analysis made in Section 4, where the designer can vary only the quota, m . As our next proposition shows, the qualitative conclusions of Theorem 1 derived under the symmetric objective specification, (10), extend to the asymmetric objective specification, (15).

Proposition 1. *When designer objective is given by (15), Theorem 1 holds under two modifications: first, modify the expression for m_M^* , the optimal quota when contestants are non-strategic, from (13) into*

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

Second, modify condition (14) into

$$(1 - \theta)/\theta \leq (\mu_S/\mu_W)(1 - \sigma)/\sigma. \quad (17)$$

Proposition 1 implies that, under the more general objective specification, (15), the designer still tends to inflate the quota if the prior odds of a contestant being weak are low relative to the capacity asymmetry, measured by μ_S/μ_W , and sets a zero quota if otherwise.²³

Application: Retention contests

Proposition 1 extends the scope of application of our analysis to settings where there are strong reasons to suspect that the designer's welfare function is asymmetric. One such setting is retention contests. In a retention contest, the designer sets a "retention quota," m , the number of incumbent workers who will be retained. The remaining workers will be dismissed. The retention quota is filled by the incumbent workers with the best performance rankings. Dismissed workers will be replaced by an equal number of new hires.

²³The qualitative nature of this result also holds if we add two more components—gains from deselecting weak contestants and losses from deselecting strong contestants—to the linear objective specification, (15). However, if we have nonlinear objective specifications, there do exist cases in which risk taking induces the designer to "under-select," i.e., to choose a positive but smaller quota. See Online Appendix F for an example. In this example, designer welfare is linear in gains from selecting strong contestants but quadratic in losses from selecting weak contestants. Quadratic losses imply that the designer cares about not only the expected number of weak selected contestants but also the variance of the number of weak selected contestants. The latter consideration is absent under our linear specifications and can render under-selection optimal.

In practice, retention contests are fairly common and are frequently engendered by “forced-ranking systems.” Forced-ranking systems, used by an estimated 20% of large U.S. corporations (Bates, 2003), assign a fixed percentage of employees (within a given comparison group) performance grades based on relative performance. The performance of employees assigned the lowest grade is deemed to be below the retention bar. In this setting, the retention quota corresponds to the number of employees in the n -employee comparison group who will not receive the lowest performance grade.

Consider a retention contest in which a firm sets a retention quota, m , for $n \geq 2$ incumbent workers, each of whom is strong with probability θ and weak with probability $1 - \theta$. A worker’s type is the worker’s private information and is independent of the other workers’ types. Under the quota, m , the $n - m$ bottom performers will be dismissed, and the firm will recruit $n - m$ workers to fill the vacancies produced by dismissal. The firm knows that each new hire will be strong with probability σ and weak with probability $1 - \sigma$. Given that, in practice, incumbent workers have retained their positions over some time period during which their performance has been observed, to some degree, incumbent workers have been better screened for quality than new hires. It thus seems natural to assume that the expected quality of a randomly selected incumbent worker is no less than the expected quality of new hires.

Assumption 3. In retention contests, $\theta \geq \sigma$.

Of course, because only the $n - m$ worst-performing incumbent workers will be replaced by new hires, Assumption 3 by no means suggests that those dismissed workers have better expected quality than the new hires. The firm’s problem is to set the retention quota, m , to optimize workforce quality, i.e., to maximize

$$\mathbb{E}[\text{\#Strong retained workers} + \text{\#Strong new hires}]. \quad (18)$$

The objective function, (18), is in fact equivalent to the asymmetric designer objective function, (15). To see this, first note that, in (18), the expected number of strong new hires equals the number of new hires, $n - m$, multiplied by the probability that a new hire is strong, σ . Thus, given that n and σ are both fixed, maximizing the objective in (18) is equivalent to maximizing

$$\mathbb{E}[\text{\#Strong retained workers}] - \sigma m.$$

Because m workers are retained, $m = \mathbb{E}[\text{\#Strong retained workers} + \text{\#Weak retained workers}]$. Thus,

$$\begin{aligned} \mathbb{E}[\text{\#Strong retained workers}] - \sigma m &= \\ \mathbb{E}[(1 - \sigma) \times \text{\#Strong retained workers} - \sigma \times \text{\#Weak retained workers}]. \end{aligned}$$

This objective function is equivalent to the asymmetric designer objective function, (15).

Given this equivalence, Proposition 1 can be used to produce a characterization of the effect of risk-taking on retention quotas. Note that the condition in Proposition 1 for selection to be biased toward quota inflation (condition (17)) is always satisfied for retention contests under

Assumption 3.

Result 1. In retention contests,

- (i) the optimal retention quota when contestants are strategic always weakly and sometimes strictly exceeds the optimal quota when contestants are non-strategic.
- (ii) When contestants are non-strategic, the optimal quota given contest size n , $m_M^*(n)$, satisfies $\lim_{n \rightarrow \infty} m_M^*(n)/n = \theta$. When contestants are strategic, the optimal quota given contest size n , $m^*(n)$, satisfies, $\lim_{n \rightarrow \infty} m^*(n)/n = \theta + (1 - \theta)(\mu_W/\mu_S) > \theta$.

Part (i) of Result 1 implies that firms tend to “over-retain” their employees, setting retention rates greater than what would be optimal if employees were non-strategic. Part (ii) shows that, if a retention contest is applied to a sufficiently large group, the firm will always inflate the retention rate regardless of the degree of capacity asymmetry. The degree of inflation equals $(1 - \theta)(\mu_W/\mu_S)$, which is proportional to the fraction of incumbent workers who are weak and inversely proportional to the capacity ratio, μ_S/μ_W . Thus, when the worker group is sufficiently large, high ability is rare, and contest performance is positively but not that strongly related to ability, strategic risk taking will substantially increase the retention quota. For example, if the capacity ratio is 3 : 2 and the prior probability that a given worker is strong equals $1/4$, asymptotically, 25% of incumbents will be retained when incumbents are non-strategic, while 75% will be retained when incumbents are strategic.

6.2 Capacity acquisition

In our baseline model, we assumed that a contestant’s capacity to perform in the contest, which fixes the contestant’s expected performance, is exogenous. Now we relax this assumption by allowing each contestant to acquire capacity through costly effort. We assume that, after the selection quota and contest size are announced to the contestants, the contestants simultaneously choose capacity by exerting costly effort. The cost of choosing the capacity level, μ , for a type- $t \in \{S, W\}$ contestant is $c_t(\mu) = c(\mu)/a_t$, where a_t is an ability parameter satisfying $a_S > a_W > 0$, and $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is twice differentiable, increasing, and strictly convex, with $c(0) = c'(0) = 0$. The assumption of $a_S > a_W$ implies that the cost and the marginal cost of acquiring any given $\mu > 0$ are both higher for weak contestants than for strong contestants. After the contestants acquire capacity, the contestants, without knowing each other’s capacity, simultaneously choose nonnegative random performance subject to their capacity constraints. Selection is still based on the ranking of realized performance. A contestant’s reward from being selected is $v > 0$ and her reward from being deselected is 0. A contestant’s payoff equals the reward she receives less her effort cost.

We first analyze how introducing capacity acquisition affects our result on optimal quotas. Following the approach used in much of the contest-design literature, we consider contest designs that feature the same reward budget, $mv = V > 0$ for all $m = 1, \dots, n - 1$. A meritocratic designer can either run a contest (i.e., choose any $m \in \{1, \dots, n - 1\}$) with a fixed budget of V

or not run a contest (i.e., choose $m = 0$ or $m = n$). The next result is the key to establishing our conclusions.

Proposition 2. *Suppose that the reward budget, V , and contest size, n , are both fixed. If a challenge equilibrium exists at the quota, $m = m'$, where $0 < m' < n$, a challenge equilibrium also exists at all quotas $m = 1, \dots, m'$, and each type of contestant's equilibrium effort (capacity) remains the same over all these quotas, $m = 1, \dots, m'$.*

Proposition 2 implies that, with a fixed reward budget, if a challenge equilibrium exists at $m = m_M^*$ and $0 < m_M^* < n$, where m_M^* represents the optimal quota when contestants are non-strategic, a challenge equilibrium also exists at all $m = 1, \dots, m_M^*$ and each type of contestant's equilibrium capacity remains the same over all these quotas, $m = 1, \dots, m_M^*$. By Theorem 1, with each type of contestant's capacity fixed, if a challenge equilibrium exists at $m = m_M^*$, choosing any $m \in (0, m_M^*)$ is always suboptimal. Thus, when capacity is endogenous, with a fixed reward budget, if a challenge equilibrium exists at $m = m_M^*$, choosing any $m \in (0, m_M^*)$ is suboptimal. If instead, a concession equilibrium exists at $m = m_M^*$, then choosing $m = m_M^*$ is clearly optimal. The following result is thus straightforward.

Corollary 1. *Suppose that contest size, n , is fixed and the optimal quota when contestants are non-strategic, m_M^* , is interior, i.e., $0 < m_M^* < n$. If the reward budget for running a contest is fixed, then when contestants are strategic and capacity is endogenous, it is optimal for the designer either to choose a contest quota no less than m_M^* or not to run a contest at all.*

Corollary 1 implies that our result on optimal quotas is robust to the introduction of capacity acquisition when a meritocratic designer faces a fixed reward budget for running a selection contest: the designer either tends to inflate the quota or finds it optimal not to run a contest.

Next, we analyze how introducing capacity acquisition affects our result on optimal contest size. In contrast to the above analysis of a quota choice, with a fixed reward budget, increasing contest size reduces per capita rewards and, hence, discourages contestant effort. However, for one class of effort cost functions, power functions, this discouragement effect does not affect meritocratic selection.

Proposition 3. *Suppose $c_t(\mu) = \mu^\alpha / a_t$, $\alpha > 1$, for $t = S, W$. Then the modified contest game in which capacity is acquired through costly effort produces the same selection outcomes, i.e., each type's probability of winning and meritocratic designer welfare, (15), as the baseline model with $\mu_S / \mu_W = (a_S / a_W)^{1/(\alpha-1)}$.*

Proposition 3 implies that, to analyze the selection outcomes of the modified game when $c_t(\mu) = \mu^\alpha / a_t$, $\alpha > 1$, for $t = S, W$, we can simply examine the selection outcomes of our baseline model, treating contestants as if their capacity ratio, μ_S / μ_W , equaled the value of $(a_S / a_W)^{1/(\alpha-1)}$.²⁴ Because this value does not depend on contest size, n , our previous result

²⁴By Lemma 2, a weak contestant's equilibrium probability of winning equals $\max[p_o^C, p_o^G]$, where p_o^C does not depend on μ_S or μ_W while p_o^G depends on μ_S / μ_W but not on μ_S or μ_W per se. Thus, as long as μ_S / μ_W is fixed, selection outcomes are independent of the absolute values of μ_S and μ_W .

on optimal contest size is robust to our extension here: there exists a threshold in contest size above which adding contestants does not increase the expected ability of contest winners.²⁵

When effort cost functions are power functions, the marginal cost function, c' , is also a power function and is thus geometrically linear (i.e., $x \mapsto xc''(x)/c'(x)$ is constant). In Online Appendix G, we show that our result on optimal contest size also extends to the case in which c' is geometrically concave (i.e., $x \mapsto xc''(x)/c'(x)$ is nonincreasing) but not to the case in which c' is strictly geometrically convex (i.e., $x \mapsto xc''(x)/c'(x)$ is increasing). The reason for the latter is because, when c' is strictly geometrically convex, once contest size is sufficiently large such that weak contestants challenge strong contestants, further increasing contest size discourages weak contestants proportionately more than it discourages strong contestants. This effect increases the capacity ratio between strong and weak contestants, μ_S/μ_W . Increased capacity ratio increases the expected ability of contest winners.

6.3 Contests for lottery tickets

Elite-university admission systems are extremely competitive. Driven by a growing concern about the associated psychological costs imposed on students, some researchers and practitioners advocate lowering the stakes of elite-university admissions competitions. One way to reduce the stakes is to introduce a lottery component into elite-university admissions decisions. For example, Schwartz (2007) argues that elite universities should adopt lottery admission systems that “approve” more applicants than can be admitted, and fill the admission quota with approved students using a random lottery. In South Korea, since 2018, colleges can no longer receive applicants’ numerical exam scores but only their exam grades on the Korean SAT. Because “Students in the same graded classification will all be considered on an equal playing field in the college admissions process” (Korea JoongAng Daily, 2015), if the number of top-graded applicants exceeds an elite college’s admission quota, a lottery process will be used for the admission.

In these situations, contestants (students) essentially compete for a lottery ticket, and the number of lottery tickets exceeds the selection (admission) quota. Schwartz (2007) argues that, because student performance can be affected by luck, in elite-university admissions competitions, the expected-quality difference between the best-performing students and the close-to-best performing students is likely to be quite small. Thus, the costs of lottery admission systems, in terms of slightly less meritocratic selection, are likely to be outweighed by their benefits of reducing psychological costs. Our next result shows that an even stronger case for lottery admission systems can be made when the random component in student performance is the product of strategic risk taking—adopting lottery admission systems can reduce the psychological costs of admissions competitions without sacrificing meritocracy.²⁶

²⁵Proposition 3 also implies that, when effort cost functions are power functions, our result on optimal quotas is robust to the introduction of capacity acquisition even if we fix the individual prize value, v , rather than fixing the reward budget, mv , for $m = 1, \dots, n - 1$.

²⁶Using a large all-pay contest setting, which focuses on contestants’ effort strategies, while abstracting from

Result 2. Suppose that a risk-taking contest with n contestants and m places has a challenge equilibrium. Then using a lottery selection policy by first approving $m' > m$ contestants based on performance ranks and then randomly selecting m out of these m' approved contestants does not affect designer welfare (defined in equation (15)), provided that the n -contestant/ m' -winner risk-taking contest also has a challenge equilibrium.

Result 2 implies that, as long as the softer competition environment created by lottery admission does not lead weak students to concede to strong students, lottery admission will not reduce the expected ability of admitted students.

7 Conclusion

In this paper, we studied selection contests designed to implement meritocracy, i.e., select, based on contest performance, strong, more able, contestants and deselect weak, less able, contestants. In contrast to much of the literature on contests and tournaments, which assumes that contest noise is the product of an exogenous “noise term” mediating the relationship between contestant actions and contest performance, in our analysis, “contest noise” is endogenously produced by strategic contestant risk taking.

Introducing strategic risk taking has fundamental effects on optimal selection-contest design: increasing competition, either by reducing the number of contestants selected or expanding the contestant pool, increases weak contestants’ tendency to play high-risk strategies that challenge stronger, more able, contestants. Because of this effect, even meritocratic designers can gain by limiting competition through adopting “clubby” contest designs, designs that feature limited contestant pools and relaxed selection standards for pool members. Our model implies that many seemingly unmeritocratic practices and proposals—e.g., “Peter-Principle” promotion and retention policies, “in-house” job competitions, and elite-university admission systems that incorporate a lottery component (Schwartz, 2007)—can, in fact, be consistent with, and sometimes even further meritocracy.²⁷

References

Armstrong, Will J, Egemen Genc, and Marno Verbeek, “Going for gold: An analysis of Morningstar analyst ratings,” *Management Science*, 2019, 65 (5), 2310–2327.

risk-taking and selection, Olszewski and Siegel (2018) show that making contests less competitive by pooling intervals of performance rankings can improve student welfare in a Pareto sense via reduced student effort, even though pooling reduces the correlation between selection and ability. Our result implies that, if students strategically take risks, reducing competition need not reduce the correlation between selection and ability.

²⁷For analytical convenience, we assumed that there are only two possible types of contestants. In Online Appendix H, we show that our contest-size result extends to cases with an arbitrary finite number of contestant types, whereas our quota-inflation result extends at least for large contests if the designer gains only by selecting the strongest type. However, if the designer gains also by selecting intermediate types, under-selection can sometimes be optimal.

- Aron, Debra J and Edward P Lazear**, “The introduction of new products,” *The American Economic Review*, 1990, 80 (2), 421–426.
- 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.
- Bates, Steve**, “Forced ranking,” *HR magazine*, 2003, 48 (6), 62–62.
- Beaudoin, David and Tim B Swartz**, “Strategies for pulling the goalie in hockey,” *The American Statistician*, 2010, 64 (3), 197–204.
- 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.
- Bothner, Matthew S, Jeong han Kang, and Toby E Stuart**, “Competitive crowding and risk taking in a tournament: Evidence from NASCAR racing,” *Administrative Science Quarterly*, 2007, 52 (2), 208–247.
- Cabral, Luis M. B.**, “R&D Competition When Firms Choose Variance,” *Journal of Economics and Management Strategy*, 2003, 12 (1), 139–150.
- Chan, William, Li Hao, and Wing Suen**, “A signaling theory of grade inflation,” *International Economic Review*, 2007, 48 (3), 1065–1090.
- Che, Yeon-Koo and Ian Gale**, “Optimal design of research contests,” *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.
- Dasgupta, Partha and Joseph Stiglitz**, “Uncertainty, Industrial Structure, and the Speed of R&D,” *Bell Journal of Economics*, 1980, 11 (1), 1–28.
- Deeprise, Donna**, *How to Recognize & Reward Employees*, American Management Association, 2006.
- Ekins, Emily**, “The Reason-Rupe Spring 2014 Millennial Survey,” *Reason*, June 2014. <http://reason.com/assets/db/2014-millennials-report.pdf>.
- Fairburn, James A. and James M. Malcomson**, “Performance, promotion, and the Peter Principle,” *Review of Economic Studies*, 2001, 68 (1), 45–66.
- 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 discouraging effect of competition in contests,” *Journal of Political Economy*, 2020, 128 (5), 1940–1975.
- Frost, Stephen**, “How to build a meritocracy,” *MinuteHack*, January 2017. <https://minutehack.com/guides/how-to-build-a-meritocracy>.
- Gaba, Anil and Ajay Kalra**, “Risk behavior in response to quotas and contests,” *Marketing Science*, 1999, 18 (3), 417–434.
- Genakos, Christos and Mario Pagliero**, “Interim rank, risk taking, and performance in dynamic tournaments,” *Journal of Political Economy*, 2012, 120 (4), 782–813.
- Guercio, Diane Del and Paula A Tkac**, “Star power: The effect of Morningstar ratings on mutual fund flow,” *Journal of Financial and Quantitative Analysis*, 2008, pp. 907–936.
- Gürtler, Oliver and Matthias Kräkel**, “Optimal tournament contracts for heterogeneous workers,” *Journal of Economic Behavior & Organization*, 2010, 75 (2), 180–191.
- Hvide, Hans K. and Eirik G. Kristiansen**, “Risk taking in selection contests,” *Games and Economic Behavior*, 2003, 42 (1), 172–179.
- 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.
- Kirchler, Michael, Florian Lindner, and Utz Weitzel**, “Rankings and risk-taking in the finance industry,” *The Journal of Finance*, 2018, 73 (5), 2271–2302.
- Klette, Tor and David de Meza**, “Is the Market Biased against Risky R&D?,” *RAND Journal of Economics*, 1986, 17 (1), 133–139.
- 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.
- Korea JoongAng Daily**, “CSAT English section to take absolute grade scale,” <https://koreajoongangdaily.joins.com/2015/10/09/etc/CSAT-English->

section-to-take-absolute-grade-scale/3010151.html 2015. Accessed: 2021-01-17.

Lazear, Edward P., “The Peter Principle: A theory of decline,” *Journal of Political Economy*, 2004, 112 (S1), S141–S163.

— **and Sherwin Rosen**, “Rank-order tournaments as optimum labor contracts,” *Journal of Political Economy*, 1981, 89 (5), 841–864.

Letina, Igor, Shuo Liu, and Nick Netzer, “Delegating performance evaluation,” *Theoretical Economics*, 2020, 15 (2), 477–509.

Meyer, Margaret A., “Learning from coarse information: Biased contests and career profiles,” *Review of Economic Studies*, 1991, 58 (1), 15–41.

Moldovanu, Benny and Aner Sela, “The optimal allocation of prizes in contests,” *American Economic Review*, 2001, 91 (3), 542–558.

Morgan, John, Justin Tumlinson, and Felix Várdy, “The limits of meritocracy,” Working paper, UC Berkeley 2018.

Myerson, Roger B., “Incentives to cultivate favored minorities under alternative electoral systems,” *American Political Science Review*, 1993, 87 (04), 856–869.

Olszewski, Wojciech and Ron Siegel, “Large contests,” *Econometrica*, 2016, 84 (2), 835–854.

— **and** — , “Pareto improvements in the contest for college admissions,” *Mimeo*, 2018.

— **and** — , “Performance-maximizing large contests,” *Theoretical Economics*, 2020, 15 (1), 57–88.

Popomaronis, Tom, “Here’s how many Google interviews it takes to hire a Googler,” <https://www.cnbc.com/2019/04/17/heres-how-many-google-job-interviews-it-takes-to-hire-a-googler.html> 2019. Accessed: 2021-02-10.

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.

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,” *Los Angeles Times*, 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.
- Strack, Philipp**, “Risk-taking in contests: The impact of fund-manager compensation on investor welfare,” *Mimeo*, 2016.

Online Appendices to “Less competition, more meritocracy?”

by Dawei Fang and Thomas Noe

Abstract

Appendix A proves the results presented in the main text. Appendix B constructs a symmetric equilibrium. Appendix C shows that our results are robust to exogenous upper bounds on performance. Appendix D considers the effect of adding contestants of lower prior quality. Appendix E shows that our results are robust to a more general capacity constraint. Appendix F shows that under-selection can be optimal if the designer’s objective function has a nonlinear specification. Appendix G extends the analysis of the effect of capacity acquisition on optimal contest size to more general effort cost functions. Appendix H considers a more general type distribution. Appendix I shows that our quota-inflation result extends to cases with no aggregate uncertainty with respect to contestant types.

A. Proofs of results

Verification of Claim 1 and proof of Lemma 1. We first verify part (a) of Claim 1, the continuity of P . Lemma 1 then follows immediately from the continuity of P and the argument developed in the main text (see the argument around Figure 1). Afterward, we verify part (b) of Claim 1.

No point mass on zero performance We first show, by way of contradiction, that in any symmetric equilibrium, no contestant places point mass on 0. Suppose, to the contrary, that in a symmetric equilibrium, there exists at least one type, say t , such that each type- t contestant places point mass on 0. Let dF_t be the performance measure associated with the function of the performance distribution, F_t . Let q_t be the probability that a given type- t contestant has zero performance, i.e.,

$$q_t = dF_t(\{0\}).$$

The hypothesis that each type- t contestant places point mass on 0 implies that

$$q_t \in (0, 1). \tag{A-1}$$

Note that we can decompose dF_t as follows:

$$dF_t = q_t \mathbb{1}_0 + (1 - q_t) dF_t^{-0}, \tag{A-2}$$

where

$$\mathbb{1}_a(x) = \begin{cases} 1 & \text{if } x = a \\ 0 & \text{otherwise} \end{cases} \quad (\text{A-3})$$

is an indicator function and

$$dF_t^{-0}(x) = \begin{cases} 0 & \text{if } x = 0 \\ \left(\frac{1}{1-q_t}\right) dF_t(x) & \text{otherwise} \end{cases}. \quad (\text{A-4})$$

This decomposition simply means that we can interpret a type- t contestant's strategy as choosing zero performance with probability q_t and choosing a random, positive performance level whose associated probability measure equals dF_t^{-0} with probability $1 - q_t$.

Now consider a type- t contestant's deviation to using an alternative performance measure $d\hat{F}_t$ given as follows:

$$d\hat{F}_t = \hat{q}_t \mathbb{1}_\varepsilon + (1 - \hat{q}_t) dF_t^{-0}, \quad (\text{A-5})$$

where $\varepsilon \in (0, \mu_t)$, $\mathbb{1}_\varepsilon$ and dF_t^{-0} are given by (A-3) and (A-4), respectively, and

$$\hat{q}_t = \frac{\mu_t/(1 - q_t) - \mu_t}{\mu_t/(1 - q_t) - \varepsilon}. \quad (\text{A-6})$$

The hypothesis that $\varepsilon < \mu_t$ and equation (A-1) imply that $\hat{q}_t \in (0, 1)$. Thus, by construction and the fact that dF_t^{-0} is a probability measure, $d\hat{F}_t$ is also a probability measure. Note that

$$\begin{aligned} \int_{0-}^{\infty} x d\hat{F}_t(x) &= \hat{q}_t \varepsilon + (1 - \hat{q}_t) \int_{0-}^{\infty} x dF_t^{-0}(x) = \hat{q}_t \varepsilon + (1 - \hat{q}_t) \int_0^{\infty} \frac{x}{1 - q_t} dF_t(x) \\ &= \hat{q}_t \varepsilon + (1 - \hat{q}_t) \left(\frac{\mu_t}{1 - q_t} \right) = \mu_t, \end{aligned}$$

where, in the first line, the first equality follows from (A-3) and (A-5) and the second equality from (A-4), and in the second line, the first equality follows from the fact that the mean of F_t equals μ_t and the second equality from (A-6). Thus, by construction, $d\hat{F}_t$ satisfies the capacity constraint for a type- t contestant.

Now we show that a type- t contestant is strictly better off deviating from dF_t to $d\hat{F}_t$ for $\varepsilon > 0$ sufficiently small. Note that

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \int_{0-}^{\infty} P(x) d\hat{F}_t(x) &= \lim_{\varepsilon \downarrow 0} \hat{q}_t P(\varepsilon) + \lim_{\varepsilon \downarrow 0} (1 - \hat{q}_t) \int_{0-}^{\infty} P(x) dF_t^{-0}(x) \\ &= q_t P(0+) + (1 - q_t) \int_{0-}^{\infty} P(x) dF_t^{-0}(x) > q_t P(0) + (1 - q_t) \int_{0-}^{\infty} P(x) dF_t^{-0}(x) = \int_{0-}^{\infty} P(x) dF_t(x), \end{aligned} \quad (\text{A-7})$$

where the first line follows from (A-5) and, in the second line, the first equality follows from the fact that, given (A-6), $\lim_{\varepsilon \downarrow 0} \hat{q}_t = q_t$, and the last equality follows from (A-2). To understand the inequality in the second line, note that, by hypothesis, each type- t contestant places point mass on 0. Thus, for a given contestant, there exists a positive probability that all of her rivals are of type t and have zero performance. Consequently, if this contestant's performance equals

0, there exists a positive probability that she will tie with all of her rivals. Hence, given the random resolution of a tie, her probability of winning by performing slightly better than 0 is higher than her probability of winning by having zero performance, i.e.,

$$P(0+) > P(0). \quad (\text{A-8})$$

The first and the last expression in (A-7), together with the strict inequality between them, imply that deviating from dF_t to $d\hat{F}_t$ for $\varepsilon > 0$ sufficiently small makes a type- t contestant strictly better off, contradicting that dF_t is an equilibrium performance measure for type- t contestants. The result that no contestant places point mass on 0 in any symmetric equilibrium thus follows.

No point mass on any positive performance level Suppose, to the contrary, that in a symmetric equilibrium, there exists at least one type, say t , such that each type- t contestant chooses a performance level equal to $a > 0$ with probability $\tau > 0$. Then consider a type- t contestant deviating to taking a mean-preserving spread of her performance by reducing her probability of choosing a from τ to 0, increasing her probability of choosing 0 by z , and increasing her probability of choosing $a + \varepsilon$ by $\tau - z$, where

$$a\tau = (\tau - z)(a + \varepsilon). \quad (\text{A-9})$$

Equation (A-9) implies that the prescribed deviation does not change the contestant's expected performance. This deviation is thus feasible to the contestant. This deviation increases the contestant's probability of winning by

$$\Delta(\varepsilon) = zP(0) + (\tau - z)P(a + \varepsilon) - \tau P(a).$$

By (A-9), z tends to 0 as $\varepsilon \downarrow 0$. Thus,

$$\lim_{\varepsilon \downarrow 0} \Delta(\varepsilon) = \tau(P(a+) - P(a)) > 0,$$

where the inequality follows from the hypotheses that $\tau > 0$ and type- t contestants place mass on a and an argument similar to the one used for establishing (A-8) (under the hypothesis we made there that type- t contestants place mass on 0). This inequality implies the existence of a profitable deviation for a type- t contestant and thus, a contradiction. The result that no contestant places point mass on any $a > 0$ in any symmetric equilibrium thus follows.

P is everywhere continuous. This follows from the fact that no contestant places point mass.

Lemma 1 holds. This result is developed in the main text. Its development relies only on the continuity of P , which has been established above.

The lower bound of the support of P is zero. Let $\underline{x} = \min \text{Supp } P$. By equation (1), there exists at least one type, say t , such that $\min \text{Supp}_t = \min \text{Supp } P = \underline{x}$. Because no contestant places point mass in any symmetric equilibrium and because the mean of F_t equals μ_t , there must exist $x^o > \mu_t$ such that $x^o \in \text{Supp}_t$. Given that $x^o, \underline{x} \in \text{Supp}_t$, Lemma 1 implies that, for a type- t contestant, a fair gamble between \underline{x} and x^o is her best reply. To obtain a contradiction, suppose, to the contrary, that $\underline{x} > 0$. Then, by equation (1) and the fact that no contestant places point mass in any symmetric equilibrium, $P(0) = P(\underline{x}) = 0$. Because $x^o > \mu_t$ and because weak

contestants have to place weight over $[0, \mu_W]$ to satisfy their capacity constraint, it must be that $P(x^o) > 0$. Thus, $P(x^o) > P(\underline{x}) = P(0)$. Hence, if $\underline{x} > 0$, for a type- t contestant, a fair gamble between 0 and x^o would produce a higher payoff than a fair gamble between \underline{x} and x^o , contradicting that the latter strategy is a best reply for a type- t contestant. The result that $\underline{x} = 0$ thus follows.

P has a connected support. Suppose, to the contrary, that P has a gap in its support. Then there must exist $0 \leq x' < x''$ such that $[x', x''] \cap \text{Supp } P = \{x', x''\}$. Then by equation (1), no contestant places weight over (x', x'') . Thus, P is flat over (x', x'') and, by continuity of P ,

$$P(x) = P(x'') \quad x \in [x', x'']. \quad (\text{A-10})$$

Because $[x', x''] \cap \text{Supp } P = \{x', x''\}$, equation (1) implies that at least one type of contestant, say t , has $x'' \in \text{Supp}_t$. Because no contestant places point mass in any symmetric equilibrium, given $x'' \in \text{Supp}_t$ and given $(x', x'') \cap \text{Supp } P = \emptyset$ and *a fortiori* $(x', x'') \cap \text{Supp}_t = \emptyset$, there must exist $\varepsilon > 0$ such that $[x'', x'' + \varepsilon] \in \text{Supp}_t$. Thus, by Lemma 1, P must be affine in x for $x \in [x'', x'' + \varepsilon]$. Because $[x'', x'' + \varepsilon] \in \text{Supp}_t$, P must be increasing in x for $x \in [x'', x'' + \varepsilon]$. Thus, by (A-10), P is constant in x for $x \in [x', x'']$ and is increasing and affine in x for $x \in [x'', x'' + \varepsilon]$, with a kink at x'' . Hence, over $x \in [x', x'' + \varepsilon]$, P is convex and nonlinear. Thus, by Jensen's inequality, placing weight in the interior of the interval $[x', x'' + \varepsilon]$ is strictly dominated by a mean-preserving spread that transfers all the weight from the interior of this interval to x' and $x'' + \varepsilon$, the two endpoints of this interval. Hence, a type- t contestant must place no weight over $(x', x'' + \varepsilon)$. This however implies, given $x' < x''$, that a type- t contestant must place no weight over $(x'', x'' + \varepsilon)$, contradicting that $[x'', x'' + \varepsilon] \in \text{Supp}_t$. The contradiction implies that P has a connected support.

P has a bounded support. For each type $t = S, W$, the fact that F_t is continuous implies the existence of x' and x'' such that $x' \neq x''$ and $x', x'' \in \text{Supp}_t$. Let β_t be the slope of the line connecting the two points, $(x', P(x'))$ and $(x'', P(x''))$. By equation (1), the fact that $x', x'' \in \text{Supp}_t$ implies that $x', x'' \in \text{Supp } P$. Thus, given that P is continuous and has a connected support and given that $x' \neq x''$, $P(x') \neq P(x'')$. Hence, given that P is nondecreasing, $\beta_t > 0$. Note that

$$\forall x \in \text{Supp}_t, \quad \beta_t (x - x') = P(x) - P(x') \leq 1 - P(x'), \quad (\text{A-11})$$

where the equality follows from Lemma 1 and the fact that $x' \in \text{Supp}_t$, and the inequality follows from the fact that P , being a probability-of-winning function, is bounded above by 1. Given that $\beta_t > 0$, (A-11) implies that

$$\forall x \in \text{Supp}_t, \quad x \leq x' + (1 - P(x')) / \beta_t,$$

which further implies that, for each type $t = S, W$, the support of F_t is bounded above. Thus, by equation (1), the support of P is bounded above.

Claim 1 follows immediately from the facts that P is continuous and has a bounded and connected support, with 0 as the lower bound of the support. \square

Proof of Lemma 2. In the main text, we proved Lemma 2 based on assuming that Claim 1 holds and a symmetric equilibrium always exists. We have just verified Claim 1. We defer the establishment of equilibrium existence to Appendix B, where we construct a symmetric equilibrium. \square

Proof of Lemma 3. We first establish the quota effect. The definitions of p_o^C and p_o^G , given in (8) and (6), respectively, imply that

$$\frac{p_o^C}{p_o^G} = \mathbb{E} \left[\max \left[1 - \frac{\tilde{S}_n}{m}, 0 \right] \right] \left(\frac{\theta \mu_S + (1 - \theta) \mu_W}{(1 - \theta) \mu_W} \right), \quad (\text{A-12})$$

where $\tilde{S}_n \sim \text{Binomial}(n, \theta)$. 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 , $m \mapsto \mathbb{E} [\max[1 - (\tilde{S}_n/m), 0]]$ must be nondecreasing. Hence, by (A-12), an increase in m weakly increases p_o^C/p_o^G . Thus, by Lemma 2, if challenge (concession) equilibria exist, then everything else being equal, challenge (concession) equilibria also exist if m decreases (increases).

Now, we establish the size effect. Consider the ratio, p_o^C/p_o^G , given by (A-12). 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 in n stochastically dominates the one before the increase. Thus, $n \mapsto \mathbb{E} [\max[1 - (\tilde{S}_n/m), 0]]$ is nonincreasing. Thus, by (A-12), an increase in n weakly decreases p_o^C/p_o^G . Hence, by Lemma 2, if challenge (concession) equilibria exist, then everything else being equal, challenge (concession) equilibria also exist if n increases (decreases). \square

Proof of Theorem 1. (i): The existence of \bar{m} follows from Assumption 2 and Lemma 2. The rest of part (i) follows from Lemma 3.

(ii): Note that the expected number of strong selected contestants equals $n\theta p_S$ and the expected number of weak selected contestants equals $n(1 - \theta)p_W$. Thus, designer welfare, u , given by (10), can be expressed as

$$u(m) = n\theta p_S - n(1 - \theta)p_W = m - 2n(1 - \theta)p_W, \quad m = 0, \dots, n, \quad (\text{A-13})$$

where the second equality follows from (3). By Lemma 2, $p_W = \max[p_o^C, p_o^G]$ whenever there is a competition (i.e., whenever $m = 1, \dots, n-1$). In fact, $p_W = \max[p_o^C, p_o^G]$ even if there is no competition (i.e., if $m = 0$ or $m = n$). To see the latter, note that, by (8) and (6), if $m = 0$, then $p_o^C = p_o^G = 0$, implying that $\max[p_o^C, p_o^G] = 0$; if $m = n$, then $p_o^C = 1 > p_o^G$, implying that $\max[p_o^C, p_o^G] = 1$. Also note that, if $m = 0$, no contestant is selected, in which case $p_W = 0$; if $m = n$, all contestants are selected, in which case $p_W = 1$. Thus, $p_W = \max[p_o^C, p_o^G]$ even if $m = 0$ or $m = n$. Hence,

$$p_W = \max[p_o^C, p_o^G], \quad m = 0, \dots, n, \quad (\text{A-14})$$

and by (A-13),

$$u(m) = m - 2n(1 - \theta) \max[p_o^C, p_o^G], \quad m = 0, \dots, n. \quad (\text{A-15})$$

Define

$$u^C(m) := m - 2n(1 - \theta)p_o^C \quad (\text{A-16})$$

$$u^G(m) := m - 2n(1 - \theta)p_o^G. \quad (\text{A-17})$$

By (A-15),

$$u(m) = \min[u^C(m), u^G(m)], \quad m = 0, \dots, n. \quad (\text{A-18})$$

Note that

$$u^C(m) = \bar{u}(m), \quad m = 0, \dots, n, \quad (\text{A-19})$$

where \bar{u} denotes designer welfare when contestants are non-strategic.

Because m_M^* , by definition, is the optimal quota when contestants are non-strategic, and also because, if there exists more than one quota that maximizes designer welfare, the optimal quota is defined as the largest quota among those that maximize designer welfare, by (A-19), it must be the case that

$$u^C(m_M^*) = \bar{u}(m_M^*) \geq \bar{u}(m) = u^C(m), \quad \text{for } m = 0, \dots, n, \text{ with strict inequality for } m > m_M^*. \quad (\text{A-20})$$

Suppose $\bar{m} < m_M^*$. Note that, by part (i), \bar{m} is the largest quota at which challenge equilibria exist. Thus, the hypothesis that $\bar{m} < m_M^*$ implies the existence of concession equilibria at m_M^* . Thus, $u(m_M^*) = u^C(m_M^*)$. In this case,

$$u(m_M^*) = u^C(m_M^*) \geq u^C(m) \geq \min[u^C(m), u^G(m)] = u(m), \quad m = 0, \dots, n, \quad (\text{A-21})$$

where the first inequality follows from (A-20) and is strict for $m > m_M^*$, and the last equality follows from (A-18). Thus, by (A-21), when contestants are strategic, if $\bar{m} < m_M^*$, then m_M^* maximizes designer welfare and there exists no $m > m_M^*$ that also maximizes designer welfare. In this case, the optimal quota when contestants are strategic equals m_M^* .

Now suppose $\bar{m} \geq m_M^*$. By part (i), challenge equilibria exist for all $m = 1, \dots, \bar{m}$. Thus, given that $u(m) = u^G(m)$ over quotas that support challenge equilibria, $u(m) = u^G(m)$ for $m = 1, \dots, \bar{m}$. Plug the expression for p_o^G given in (6) into (A-17) and simplify the result. This yields

$$u^G(m) = m \left(\frac{\theta\mu_S - (1 - \theta)\mu_W}{\theta\mu_S + (1 - \theta)\mu_W} \right). \quad (\text{A-22})$$

Given the hypothesis that condition (14) is satisfied, $m \mapsto u^G(m)$ is nondecreasing and non-negative for $m > 0$. Thus, given that $u(m) = u^G(m)$ for $m = 1, \dots, \bar{m}$ and given the fact that $u(0) = 0$, it must be the case that

$$u(\bar{m}) \geq u(m), \quad m = 0, \dots, \bar{m} - 1. \quad (\text{A-23})$$

By part (i) and the fact that $u(m) = u^C(m) = \bar{u}(m)$ over quotas that support concession equilibria,

$$u(m) = u^C(m) = \bar{u}(m), \quad m = \bar{m} + 1, \dots, n.^1 \quad (\text{A-24})$$

¹If $m = n$, all contestants will be selected. In this case, it is weakly optimal for weak contestants to play concession strategies. Thus, $u(m) = u^C(m)$ when $m = n$.

The hypothesis that $\bar{m} \geq m_M^*$ implies that $\bar{m} + 1 > m_M^*$. Thus, by (12) and (13), $\bar{u}(\bar{m} + 1) > \bar{u}(m)$ for all $m > \bar{m} + 1$. This implies, by (A-24), that

$$u(\bar{m} + 1) > u(m), \quad m = \bar{m} + 2, \dots, n. \quad (\text{A-25})$$

Equations (A-23) and (A-25) imply that, if $\bar{m} \geq m_M^*$, the optimal quota when contestants are strategic is \bar{m} or $\bar{m} + 1$. This completes the proof of part (ii).

(iii): Suppose that condition (14) is violated. Then, by (A-22), $u^G(m) < 0$ for all $m > 0$. Equation (A-18) implies that $u(m) \leq u^G(m)$ for all $m > 0$. Thus, $u(m) < 0$ for all $m > 0$. Given that $u(0) = 0$, it is optimal to choose $m = 0$. This establishes part (iii) and completes the proof of Theorem 1. \square

Proof of Theorem 2. (i): Note that

$$\mathbb{E} \left[\max \left[1 - \frac{\tilde{S}_n}{m}, 0 \right] \right] = \mathbb{P}[\tilde{S}_n < m] \mathbb{E} \left[1 - \frac{\tilde{S}_n}{m} \middle| \tilde{S}_n < m \right] \leq \mathbb{P}[\tilde{S}_n < m]. \quad (\text{A-26})$$

Because $\tilde{S}_n \sim \text{Binomial}(n, \theta)$, $\mathbb{P}[\tilde{S}_n < m] \rightarrow 0$ as $n \rightarrow \infty$. Thus, by (A-12) and (A-26), $p_o^C/p_o^G \rightarrow 0$ as $n \rightarrow \infty$. Also note that, in the proof of Lemma 3, we showed that an increase in n weakly decreases p_o^C/p_o^G . Thus, there must exist n^c such that $p_o^C/p_o^G < 1$ for all $n > n^c$. Part (i) then follows immediately from Lemma 2.

(ii): Suppose that challenge equilibria exist at n' . Then by Lemma 3, challenge equilibria also exist at any $n > n'$. By (A-22), designer welfare in challenge equilibria, u^G , does not depend on n . Thus, designer welfare at any $n > n'$ equals designer welfare at n' . Part (ii) then follows. \square

Proof of Proposition 1. Part (i) of Theorem 1 does not depend on how we specify the designer's objective. Thus, introducing asymmetry into designer objective does not change any conclusion in part (i) of the theorem.

Now consider how introducing asymmetry into designer objective changes parts (ii) and (iii) of Theorem 1. Let \hat{u} be designer welfare when designer objective is given by (15). Note that

$$\hat{u}(m) = (1 - \sigma)n\theta p_S - \sigma n(1 - \theta)p_W = m(1 - \sigma) - n(1 - \theta)p_W, \quad m = 0, \dots, n, \quad (\text{A-27})$$

where the second equality follows from equation (3). Define

$$\hat{u}^C(m) := m(1 - \sigma) - n(1 - \theta)p_o^C \quad (\text{A-28})$$

$$\hat{u}^G(m) := m(1 - \sigma) - n(1 - \theta)p_o^G. \quad (\text{A-29})$$

By (A-14), (A-27), and Lemma 2,

$$\hat{u}(m) = \min[\hat{u}^C(m), \hat{u}^G(m)], \quad m = 0, \dots, n. \quad (\text{A-30})$$

Note that $\hat{u}^C(m)$ represents designer welfare when the quota, m , supports concession equilibria. Selection when weak contestants concede is equivalent to selection when contestants are non-strategic. Because, when contestants are non-strategic, the number of strong selected

contestants is $\min[m, \tilde{S}_n]$ and the number of weak selected contestants is $m - \min[m, \tilde{S}_n]$, where $\tilde{S}_n \sim \text{Binomial}(n, \theta)$ represents the number of strong contestants out of the n contestants, it must be the case that

$$\hat{u}^C(m) = \mathbb{E}[(1 - \sigma) \min[m, \tilde{S}_n] - \sigma(m - \min[m, \tilde{S}_n])] = \mathbb{E}[\min[\tilde{S}_n, m]] - m\sigma.$$

Thus,

$$\begin{aligned} \hat{u}^C(m+1) - \hat{u}^C(m) &= \mathbb{E}[\min[\tilde{S}_n, m+1]] - \mathbb{E}[\min[\tilde{S}_n, m]] - \sigma = \mathbb{P}[\tilde{S}_n \geq m+1] - \sigma \\ &= 1 - \mathbb{P}[\tilde{S}_n \leq m] - \sigma = 1 - B(m; n, \theta) - \sigma, \end{aligned} \quad (\text{A-31})$$

where $B(\cdot; n, \theta)$ denotes the CDF of the $\text{Binomial}(n, \theta)$ distribution. (A-31) implies that

$$\hat{u}^C(m+1) > (=)(<) \hat{u}^C(m) \quad \text{if} \quad B(m; n, \theta) < (=)(>) 1 - \sigma.$$

Thus, given that $m \mapsto B(m; n, \theta)$ is increasing, \hat{u}^C is first increasing and then decreasing in m , with its maximum reached at $m = m_M^*$, where m_M^* , whose expression is given by (13) when designer objective has the symmetric specification (10), is now given by (16) when designer objective has the asymmetric specification (15). Because designer welfare when contestants are non-strategic equals designer welfare produced by concession equilibria, m_M^* , given by (16), is the optimal quota in the case of non-strategic contestants.²

In what follows, let m_M^* be defined as in equation (16). If $\bar{m} < m_M^*$, then concession equilibria exist at $m = m_M^*$. Then, by the same argument as used in the proof of part (ii) of Theorem 1, m_M^* is the optimal quota when contestants are strategic.

Now consider the case in which $\bar{m} \geq m_M^*$. Note that, by (6) and (A-29),

$$\hat{u}^G(m) = m \left(1 - \sigma - \frac{(1 - \theta) \mu_W}{\theta \mu_S + (1 - \theta) \mu_W} \right). \quad (\text{A-32})$$

Thus, if condition (17) is satisfied, $m \mapsto \hat{u}^G(m)$ is nondecreasing and nonnegative. Then, by essentially the same argument as used in the proof of part (ii) of Theorem 1, if $\bar{m} \geq m_M^*$ and condition (17) is satisfied, the optimal quota given strategic contestants is either \bar{m} or $\bar{m} + 1$. If condition (17) is violated, then by (A-32), $\hat{u}^G(m) < 0$ for all $m > 0$. In this case, by (A-30), $\hat{u}(m) < 0$ for all $m > 0$. Then, by the fact that $\hat{u}(0) = 0$, it is optimal to choose a zero quota when contestants are strategic. \square

Proof of Result 1. We first establish part (i). Because $\mu_S > \mu_W$ and because, by hypothesis, $\theta \geq \sigma$ in retention contests, condition (17) is satisfied. Part (i) then follows immediately from (a) Proposition 1, (b) the satisfaction of condition (17), and (c) the equivalence between designer objective given by (15) and designer objective in retention contests given by (18).

Next, we establish part (ii). Because $\tilde{S}_n \sim \text{Binomial}(n, \theta)$, where \tilde{S}_n denotes the number of

²If $B(m_M^*; n, \theta) = 1 - \sigma$, then both m_M^* , given by (16), and $m_M^* - 1$ maximize \hat{u}^C . Because, in the case in which two adjacent quotas are both optimal, the optimal quota is defined as the larger of the two, the optimal quota is m_M^* .

strong workers, by the law of large numbers,

$$\tilde{S}_n/n \xrightarrow{a.s.} \theta. \quad (\text{A-33})$$

Thus, asymptotically, when contestants are non-strategic, the optimal fraction of retained workers equals θ , i.e., $\lim_{n \rightarrow \infty} m_M^*(n)/n = \theta$.

Let $\gamma = m/n$ be the retention rate. We can rewrite p_o^C , using equation (8), as follows:

$$p_o^C = \frac{\mathbb{E}[\max[\gamma n - \tilde{S}_n, 0]]}{n(1 - \theta)} = \frac{1}{1 - \theta} \mathbb{E} \left[\max \left[\gamma - \frac{\tilde{S}_n}{n}, 0 \right] \right]. \quad (\text{A-34})$$

Equations (A-33) and (A-34) imply that

$$p_o^C \rightarrow \frac{\max[\gamma - \theta, 0]}{1 - \theta} \quad \text{as } n \rightarrow \infty. \quad (\text{A-35})$$

We can rewrite p_o^G , using equation (6), as

$$p_o^G = \gamma \left(\frac{\mu_W}{\theta \mu_S + (1 - \theta) \mu_W} \right). \quad (\text{A-36})$$

Equations (A-35) and (A-36) imply that

$$\text{as } n \rightarrow \infty, \quad p_o^C > (=)(<) p_o^G \quad \text{if } \gamma > (=)(<) \theta + (1 - \theta)(\mu_W/\mu_S). \quad (\text{A-37})$$

Thus, when $\gamma = \theta$, $p_o^C < p_o^G$ as $n \rightarrow \infty$. This implies, by Lemma 2, that asymptotically, challenge equilibria exist when the retention rate equals θ . Note that, as has been shown, asymptotically, the optimal retention rate when workers are non-strategic equals θ . Thus, given that, in retention contests, $\theta \geq \sigma$, implying the satisfaction of condition (17), Proposition 1 implies that, asymptotically, when workers are strategic, the optimal retention rate must equal the cutoff rate such that challenge equilibria exist when the retention rate is below this cutoff while concession equilibria exist when the retention rate is above this cutoff. Hence, by Lemma 2, asymptotically, when workers are strategic, the optimal retention rate equals the one such that, at this rate, $p_o^C = p_o^G$. Thus, by (A-37), asymptotically, the optimal retention rate equals $\theta + (1 - \theta)(\mu_W/\mu_S)$ when workers are strategic. Part (ii) thus follows. \square

Proof of Proposition 2. In a symmetric equilibrium, at the effort stage, each type- t contestant, $t = S, W$, chooses the same capacity, μ_t . Note that, in any symmetric equilibrium, it must be the case that $\mu_t > 0$. This is because, if one type chose zero capacity, then given that, with a positive probability, all contestants are of this type, the probability that all contestants choose zero capacity and, thereby, have zero performance would be positive. In this case, any contestant of this type would be strictly better off choosing $\varepsilon > 0$ capacity and a performance level equal to $\varepsilon > 0$. The cost of choosing $\varepsilon > 0$ capacity can be made arbitrarily small by shrinking ε to zero while, for all positive ε , no matter how small, having $\varepsilon > 0$ performance would generate a gain that is bounded below by a positive number.

Let $P(\cdot; \mu_W, \mu_S)$ be the probability-of-winning function in a symmetric equilibrium of the subgame starting from the risk-taking stage, when, at the effort stage, all type- t contestants choose $\mu_t > 0$. The following two results are useful for the proof.

Result A-1. On the equilibrium path, the capacity, μ_t , chosen by each type- t contestant, $t = S, W$, and the probability-of-winning function, $P(\cdot; \mu_W, \mu_S)$, satisfy that

$$vP'(\mu_t; \mu_W, \mu_S) = c'(\mu_t)/a_t, \quad t = S, W, \quad (\text{A-38})$$

and $\mu_S > \mu_W$.

Proof. By essentially the same argument as the one used in Section 3.2 for establishing concavity of P in any symmetric equilibrium when capacities are exogenous, $P(\cdot; \mu_W, \mu_S)$ is concave for any fixed $\mu_S, \mu_W > 0$.³ Because taking no risk is a best reply to a concave P , by choosing capacity μ , a type- $t \in \{S, W\}$ contestant's expected payoff equals $vP(\mu; \mu_W, \mu_S) - (c(\mu)/a_t)$. In a symmetric equilibrium, a type- t contestant's expected payoff, $vP(\mu; \mu_W, \mu_S) - (c(\mu)/a_t)$, must be maximized at $\mu = \mu_t$. Because P is concave while, by assumption, c is strictly convex with $c(0) = c'(0) = 0$, the maximizer, μ_t , must satisfy the first-order condition given by (A-38).

To show that $\mu_S > \mu_W$ in any symmetric equilibrium, note that, if, to the contrary, $\mu_S \leq \mu_W$, then convexity of c would imply that $c'(\mu_W) \geq c'(\mu_S)$. Then, the fact that $a_W < a_S$ would imply that $c'(\mu_W)/a_W > c'(\mu_S)/a_S$. Thus, by (A-38), it would have to be the case that $P'(\mu_W; \mu_W, \mu_S) > P'(\mu_S; \mu_W, \mu_S)$, which, given concavity of P , could happen only if $\mu_W < \mu_S$, contradicting the hypothesis that $\mu_S \leq \mu_W$. \square

Result A-2. If challenge equilibria exist, then over the support of $P(\cdot; \mu_W, \mu_S)$,

$$P'(x; \mu_W, \mu_S) = m/(n(\theta\mu_S + (1-\theta)\mu_W)). \quad (\text{A-39})$$

Proof. By the argument developed after Lemma 1 in Section 3.2, in challenge equilibria, P is linear over its support and meets the origin. Thus, there exists $\alpha > 0$ such that, in challenge equilibria, $P(x) = \alpha x$ over its support. Equations (4) and (6) then imply that $\alpha = m/(n(\theta\mu_S + (1-\theta)\mu_W))$. The result then follows from the fact that, over the support of P , $P' = \alpha$. \square

Results A-1 and A-2 imply that, over the quotas that support challenge equilibria, the pair of equilibrium choices of capacities, μ_S and μ_W , solves

$$(mv)/(n(\theta\mu_S + (1-\theta)\mu_W)) = c'(\mu_S)/a_S \quad (\text{A-40})$$

$$(mv)/(n(\theta\mu_S + (1-\theta)\mu_W)) = c'(\mu_W)/a_W. \quad (\text{A-41})$$

Note that, when $mv = V$ is fixed, the pair of μ_S and μ_W that solves (A-40) and (A-41) does not vary with m .

Now suppose that challenge equilibria exist at $m = m'$. Then, it must be the case that there exists a pair, μ'_S and μ'_W , that solves (A-40) and (A-41), and challenge equilibria are supported by the capacity pair, μ'_S and μ'_W , under the quota, m' . Because, by Lemma 3, fixing capacities,

³Concavity 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 developed in Section 3.2 for establishing concavity of P can be applied. If $\mu_S = \mu_W$, one can treat each contestant as of the same type at the risk-taking stage by treating either $\theta = 0$ or $\theta = 1$. The proof of concavity of P in Section 3.2 does not rely on the value of θ . Thus, concavity of P also holds when $\mu_S = \mu_W$.

lowering the quota inclines weak contestants to challenge strong contestants, and also because, with $mv = V$ fixed, the solution to (A-40) and (A-41) does not vary with m , it must be the case that the capacity pair, μ'_S and μ'_W , solves (A-40) and (A-41) for all $m = 1, \dots, m'$, and over all $m = 1, \dots, m'$, challenge equilibria are supported by the capacity pair, μ'_S and μ'_W . Because the capacity pair, μ'_S and μ'_W , solves (A-40) and (A-41) and supports challenge equilibria for $m = 1, \dots, m'$, by Results A-1 and A-2, the pair, μ'_S and μ'_W , is the equilibrium pair of capacities for $m = 1, \dots, m'$. The proposition thus follows. \square

Proof of Corollary 1. Follows immediately from Proposition 2 and the argument right before the corollary in the main text. \square

Proof of Proposition 3. The proof requires the following technical result.

Result A-3. Define p_o^C and p_o^G as in (8) and (6), respectively. Define p_s^C as

$$p_s^C := \frac{1}{\theta} \left(\frac{m}{n} - (1 - \theta)p_o^C \right). \quad (\text{A-42})$$

Then $p_o^C > (=)(<)p_o^G$ if and only if $\mu_S/\mu_W > (=)(<)p_s^C/p_o^C$.

Proof. Note that

$$\begin{aligned} p_o^C - p_o^G &= p_o^C - \frac{m}{n} \left(\frac{\mu_W}{\theta \mu_S + (1 - \theta) \mu_W} \right) = p_o^C - (\theta p_s^C + (1 - \theta) p_o^C) \left(\frac{\mu_W}{\theta \mu_S + (1 - \theta) \mu_W} \right) \\ &= \frac{\theta \mu_W p_o^C}{\theta \mu_S + (1 - \theta) \mu_W} \left(\frac{\mu_S}{\mu_W} - \frac{p_s^C}{p_o^C} \right), \quad (\text{A-43}) \end{aligned}$$

where the first equality follows from the definition of p_o^G given in (6), the second from the fact that p_s^C , defined in (A-42), satisfies that

$$\theta p_s^C + (1 - \theta) p_o^C = m/n, \quad (\text{A-44})$$

and the last from simplifying the expression. By (A-43), the sign of $p_o^C - p_o^G$ equals the sign of $(\mu_S/\mu_W) - (p_s^C/p_o^C)$. The result thus follows. \square

Result A-3, combined with Lemma 2, implies that in the subgame starting from the risk-taking stage, when in the effort stage, strong contestants all choose capacity μ_S and weak contestants all choose capacity μ_W , $\mu_S > \mu_W > 0$, concession (challenge) equilibria exist if and only if $\mu_S/\mu_W \geq (<)p_s^C/p_o^C$. By Result A-1, in any symmetric equilibrium, in the effort stage, the capacities chosen by the two types satisfy the first-order conditions given by (A-38) and the condition that $\mu_S > \mu_W > 0$. When $c_t(\mu) = \mu^\alpha/a_t$ for $t = S, W$, the first-order conditions given by (A-38) are equivalent to

$$vP'(\mu_t; \mu_W, \mu_S) = \alpha \mu_t^{\alpha-1}/a_t, \quad t = S, W, \quad (\text{A-45})$$

where $P(\cdot; \mu_W, \mu_S)$ denotes the probability-of-winning function in a symmetric equilibrium of the subgame starting from the risk-taking stage, when, in the effort stage, all type- t contestants choose $\mu_t > 0$. The next result allows us to compute P' , which will enable us to further apply equation (A-45).

Result A-4. Suppose that, in the effort stage, each type- $t \in \{S, W\}$ contestant chooses capacity μ_t , and that $\mu_S > \mu_W > 0$. Then the probability-of-winning function, $P(\cdot; \mu_W, \mu_S)$, in any symmetric equilibrium of the subgame starting from the risk-taking stage satisfies the following conditions:

- i. if $p_o^C < p_o^G$ (or equivalently, $\mu_S/\mu_W < p_s^C/p_o^C$), where p_o^C , p_o^G , and p_s^C are given by (8), (6), and (A-42) respectively, then the subgame has challenge equilibria. In all of these challenge equilibria,

$$P(x; \mu_W, \mu_S) = \min \left[\frac{m}{n(\theta\mu_S + (1-\theta)\mu_W)} x, 1 \right]. \quad (\text{A-46})$$

- ii. If $p_o^C \geq p_o^G$ (or equivalently, $\mu_S/\mu_W \geq p_s^C/p_o^C$), then the subgame has a unique concession equilibrium. In this concession equilibrium,

$$P(x; \mu_W, \mu_S) = \begin{cases} \beta_W x & x \in [0, \tilde{x}] \\ \alpha_S + \beta_S x & x \in [\tilde{x}, \hat{x}] \\ 1 & x \geq \hat{x} \end{cases}, \quad (\text{A-47})$$

where $\max \text{Supp}_W = \min \text{Supp}_S = \tilde{x}$, and β_W , \tilde{x} , α_S , β_S , and \hat{x} are determined by contest parameters as follows:

$$\beta_W = p_o^C / \mu_W \quad (\text{A-48})$$

$$\tilde{x} = \tilde{p} \mu_W / p_o^C \quad (\text{A-49})$$

$$\alpha_S = \frac{\tilde{p} (\mu_S - (p_s^C \mu_W / p_o^C))}{\mu_S - (\tilde{p} \mu_W / p_o^C)} \quad (\text{A-50})$$

$$\beta_S = \frac{p_s^C - \tilde{p}}{\mu_S - (\tilde{p} \mu_W / p_o^C)} \quad (\text{A-51})$$

$$\hat{x} = \frac{(1 - \tilde{p})\mu_S - (1 - p_s^C)(\tilde{p} \mu_W / p_o^C)}{p_s^C - \tilde{p}} \quad (\text{A-52})$$

with p_o^C , p_s^C , and \tilde{p} given, respectively, by (8), (A-42), and

$$\tilde{p} = \sum_{i=n-m}^{n-1} \binom{n-1}{i} (1-\theta)^i \theta^{n-1-i}. \quad (\text{A-53})$$

Proof. (i): The existence of challenge and the non-existence of concession equilibria when $p_o^C < p_o^G$ follow from Lemma 2. By Result A-3, $p_o^C < p_o^G$ is equivalent to $\mu_S/\mu_W < p_s^C/p_o^C$. As shown by Figure 2 and the argument around it, in challenge equilibria, there exists $\beta > 0$ such that $P(x; \mu_W, \mu_S) = \min[\beta x, 1]$. Result A-2 implies that

$$\beta = m / (n(\theta\mu_S + (1-\theta)\mu_W)).$$

Part (i) thus follows.

(ii): The existence of concession and the non-existence of challenge equilibria when $p_o^C \geq p_o^G$ follow from Lemma 2. By Result A-3, $p_o^C \geq p_o^G$ is equivalent to $\mu_S/\mu_W \geq p_s^C/p_o^C$. As shown

by Figure 2 and the argument around it, in concession equilibria, P must have the form given by (A-47), with $\max \text{Supp}_W = \min \text{Supp}_S$.

Now we show that the five constants, β_W , \tilde{x} , α_S , β_S , and \hat{x} , must satisfy equations (A-48)–(A-52) in a concession equilibrium. First, continuity of P , combined with (A-47), implies that

$$\beta_W \tilde{x} = \alpha_S + \beta_S \tilde{x} \quad (\text{A-54})$$

$$\alpha_S + \beta_S \hat{x} = 1. \quad (\text{A-55})$$

Next, by (4) and (8), in a concession equilibrium, $P(\mu_W; \mu_W, \mu_S) = p_o^C$. Because, by (8), p_o^C represents a weak contestant's probability of winning in a concession equilibrium, equations (3) and (A-44) imply that p_s^C represents a strong contestant's probability of winning in a concession equilibrium. Thus, by (4), in a concession equilibrium, $P(\mu_S; \mu_W, \mu_S) = p_s^C$. These results, combined with equation (A-47) and the fact that $\max \text{Supp}_W = \min \text{Supp}_S = \tilde{x}$ in a concession equilibrium, imply that

$$\beta_W \mu_W = p_o^C \quad (\text{A-56})$$

$$\alpha_S + \beta_S \mu_S = p_s^C. \quad (\text{A-57})$$

Finally, because, in a concession equilibrium, $\max \text{Supp}_W = \min \text{Supp}_S = \tilde{x}$, and also because no contestant places point mass, for a given contestant, if her performance equals \tilde{x} , she will outperform all weak rivals but be outperformed by all strong rivals. Given that each rival is strong with probability θ and rival types are independent, the given contestant's probability of winning by having performance equal to \tilde{x} in a concession equilibrium must equal \tilde{p} given by (A-53). Thus, in a concession equilibrium, it must be the case that $P(\tilde{x}; \mu_W, \mu_S) = \tilde{p}$, which, by (A-47), implies that

$$\beta_W \tilde{x} = \tilde{p}. \quad (\text{A-58})$$

Equations (A-54)–(A-58) imply equations (A-48)–(A-52), and the result follows. \square

By Result A-4 and the fact that p_s^C/p_o^C does not depend on μ_S or μ_W (the two endogenous variables in the capacity acquisition model), to prove the proposition, it suffices to show that

- a. when $(a_S/a_W)^{1/(\alpha-1)} < p_s^C/p_o^C$, there exist challenge equilibria but no concession equilibria. In these challenge equilibria, in the effort stage, the capacity levels chosen by the two types satisfy that $\mu_S/\mu_W = (a_S/a_W)^{1/(\alpha-1)}$.
- b. When $(a_S/a_W)^{1/(\alpha-1)} \geq p_s^C/p_o^C$, there exists a concession equilibrium but no challenge equilibrium. In this concession equilibrium, in the effort stage, the capacity levels chosen by the two types satisfy that $\mu_S/\mu_W \geq (a_S/a_W)^{1/(\alpha-1)}$.

In what follows, we prove the proposition by establishing (a) and (b). Note that challenge equilibria exist if and only if (i) the choice of μ_S and μ_W satisfies (A-45), where P is given by (A-46), and (ii) $\mu_S/\mu_W < p_s^C/p_o^C$. In fact, the satisfaction of these conditions is equivalent to the satisfaction of the condition that $(a_S/a_W)^{1/(\alpha-1)} < p_s^C/p_o^C$. To see this, note that, when P

is given by (A-46), equation (A-45) is equivalent to

$$\frac{mv}{n(\theta\mu_S + (1-\theta)\mu_W)} = \frac{\alpha\mu_t^{\alpha-1}}{a_t}, \quad t = S, W, \quad (\text{A-59})$$

which further equates to the satisfaction of

$$\mu_W = \left[\frac{mv a_W}{n\alpha \left(\theta \left(\frac{a_S}{a_W} \right)^{\frac{1}{\alpha-1}} + 1 - \theta \right)} \right]^{\frac{1}{\alpha}}, \quad \mu_S = \left[\frac{mv a_S \left(\frac{a_S}{a_W} \right)^{\frac{1}{\alpha-1}}}{n\alpha \left(\theta \left(\frac{a_S}{a_W} \right)^{\frac{1}{\alpha-1}} + 1 - \theta \right)} \right]^{\frac{1}{\alpha}}, \quad (\text{A-60})$$

$$\text{and } \mu_S/\mu_W = (a_S/a_W)^{1/(\alpha-1)}. \quad (\text{A-61})$$

By (A-61), $\mu_S/\mu_W < p_s^C/p_o^C$ if and only if $(a_S/a_W)^{1/(\alpha-1)} < p_s^C/p_o^C$. Thus, challenge equilibria exist if and only if $(a_S/a_W)^{1/(\alpha-1)} < p_s^C/p_o^C$, with the endogenous capacity levels and the endogenous capacity ratio given by (A-60) and (A-61), respectively.

Next, consider concession equilibria. A concession equilibrium exists if and only if (i) the choice of μ_S and μ_W satisfies (A-45), where P is given by (A-47), and (ii) $\mu_S/\mu_W \geq p_s^C/p_o^C$. Note that, when P is given by (A-47), equations (A-45), (A-48), and (A-51) imply that

$$\frac{\alpha\mu_W^{\alpha-1}}{a_W} = \frac{p_o^C}{\mu_W} \quad (\text{A-62})$$

$$\frac{\alpha\mu_S^{\alpha-1}}{a_S} = \frac{p_s^C - \tilde{p}}{\mu_S - (\tilde{p}\mu_W/p_o^C)}. \quad (\text{A-63})$$

In what follows, we show that, when μ_W and μ_S satisfy (A-62) and (A-63), $\mu_S/\mu_W \geq p_s^C/p_o^C$ if and only if $(a_S/a_W)^{1/(\alpha-1)} \geq p_s^C/p_o^C$.

Divide (A-63) by (A-62) on both sides and rearrange the result. This yields

$$\left(\frac{\mu_S}{\mu_W} \right)^{\alpha-1} - \frac{a_S}{a_W} \left(\frac{p_s^C - \tilde{p}}{\frac{\mu_S}{\mu_W} p_o^C - \tilde{p}} \right) = 0. \quad (\text{A-64})$$

Let $r = \mu_S/\mu_W$ and $z = a_S/a_W$. Treat the left-hand side of (A-64) as a function of r and z and denote it by \mathcal{K} , i.e.,

$$\mathcal{K}(r, z) := r^{\alpha-1} - z \left(\frac{p_s^C - \tilde{p}}{r p_o^C - \tilde{p}} \right). \quad (\text{A-65})$$

By (A-65),

$$\mathcal{K} \left(p_s^C/p_o^C, (p_s^C/p_o^C)^{\alpha-1} \right) = 0. \quad (\text{A-66})$$

Note that p_s^C , given by (A-42), is a strong contestant's probability of winning in a concession equilibrium, i.e., p_s^C represents a contestant's probability of winning if the contestant always beats weak rivals and shares the same winning probability with strong rivals. Also note that \tilde{p} , given by (A-53), is a contestant's probability of winning if the contestant always beats weak rivals but is always beaten by strong rivals. Thus, it must be the case that

$$p_s^C > \tilde{p}. \quad (\text{A-67})$$

Equation (A-65) thus implies that

$$\text{for any fixed } z \geq 0, \mathcal{K} \text{ is increasing in } r \text{ for } r \geq p_s^C/p_o^C, \quad (\text{A-68})$$

$$\& \quad \text{for any fixed } r \geq p_s^C/p_o^C, \mathcal{K} \text{ is decreasing in } z \text{ for } z \geq 0. \quad (\text{A-69})$$

Hence,

$$\begin{aligned} \forall r \geq p_s^C/p_o^C \& 0 \leq z < (p_s^C/p_o^C)^{\alpha-1}, \quad \mathcal{K}(r, z) \geq \mathcal{K}(p_s^C/p_o^C, z) \\ & > \mathcal{K}\left(p_s^C/p_o^C, (p_s^C/p_o^C)^{\alpha-1}\right) = 0, \end{aligned} \quad (\text{A-70})$$

where the first inequality follows from (A-68), the second inequality from (A-69), and the equality from (A-66). By definition, $r = \mu_S/\mu_W$ and $z = a_S/a_W$. Thus, equation (A-70) implies that, when $a_S/a_W < (p_s^C/p_o^C)^{\alpha-1}$ (i.e., when $(a_S/a_W)^{1/(\alpha-1)} < p_s^C/p_o^C$), there exists no capacity ratio, μ_S/μ_W , that both solves equation (A-64) and satisfies $\mu_S/\mu_W \geq p_s^C/p_o^C$. Because the existence of a concession equilibrium requires the existence of $\mu_S/\mu_W \geq p_s^C/p_o^C$ that solves (A-64), concession equilibria do not exist if $(a_S/a_W)^{1/(\alpha-1)} < p_s^C/p_o^C$.

Equations (A-66) and (A-69) imply that

$$\text{for any fixed } z \geq (p_s^C/p_o^C)^{\alpha-1}, \mathcal{K}(p_s^C/p_o^C, z) \leq \mathcal{K}(p_s^C/p_o^C, (p_s^C/p_o^C)^{\alpha-1}) = 0. \quad (\text{A-71})$$

Equations (A-65) and (A-67) imply that

$$\text{for any fixed } z > 0, \mathcal{K}(r, z) \rightarrow \infty \text{ as } r \rightarrow \infty. \quad (\text{A-72})$$

Hence, by equations (A-68), (A-71), and (A-72) and the fact that, with z fixed, $r \mapsto \mathcal{K}(r, z)$ is continuous, for any fixed $z \geq (p_s^C/p_o^C)^{\alpha-1}$, there exists a unique $r^o \geq p_s^C/p_o^C$ such that $\mathcal{K}(r^o, z) = 0$. Thus, by the definitions of r , z , and \mathcal{K} , when $a_S/a_W \geq (p_s^C/p_o^C)^{\alpha-1}$ (i.e., when $(a_S/a_W)^{1/(\alpha-1)} \geq p_s^C/p_o^C$), there exists a unique capacity ratio, μ_S/μ_W , that both solves equation (A-64) and satisfies $\mu_S/\mu_W \geq p_s^C/p_o^C$. Because a concession equilibrium exists if there exists a pair of μ_W and μ_S that solves (A-62) and (A-64) and satisfies $\mu_S/\mu_W \geq p_s^C/p_o^C$, and because there clearly exists a unique μ_W that solves (A-62), a concession equilibrium exists if $(a_S/a_W)^{1/(\alpha-1)} \geq p_s^C/p_o^C$.

Therefore, a concession equilibrium exists if and only if $(a_S/a_W)^{1/(\alpha-1)} \geq p_s^C/p_o^C$. \square

Proof of Result 2. Note that, in challenge equilibria, the probability that a selected contestant is weak satisfies

$$\mathbb{P}[W|\text{selected}] = \frac{(1-\theta)p_W}{\theta p_S + (1-\theta)p_W} = \frac{n(1-\theta)p_W}{m} = \frac{(1-\theta)\mu_W}{\theta\mu_S + (1-\theta)\mu_W}, \quad (\text{A-73})$$

where the first equality follows from Bayes rule, the second from equation (3), and the last from equation (6). By equation (A-73), as long as both the n -contestant/ m -winner contest and the n -contestant/ m' -winner contest have challenge equilibria, both contests will produce the same $\mathbb{P}[W|\text{selected}]$ and, hence, the same $\mathbb{P}[S|\text{selected}] = 1 - \mathbb{P}[W|\text{selected}]$. In this case, using the lottery selection policy will not change the expected number of weak selected contestants or the expected number of strong selected contestants. Hence, it will not affect designer welfare specified by (15). \square

B. Construction of a symmetric equilibrium

Result B-1. If $p_o^C \geq p_o^G$, where p_o^C and p_o^G are given by (8) and (6) respectively, there exists a unique concession equilibrium. In this equilibrium, the probability-of-winning function, P , is given by (A-47), and F_W and F_S are given as follows: define β_W , \tilde{x} , α_S , β_S , \hat{x} by equations (A-48)–(A-52), respectively. Then define

$$\phi(y) := \frac{1}{\beta_W} \sum_{i=n-m}^{n-1} \binom{n-1}{i} [(1-\theta)y]^i [1-(1-\theta)y]^{n-1-i}, \quad y \in [0, 1] \quad (\text{B-1})$$

$$\zeta(y) := \frac{1}{\beta_S} \left(\sum_{i=n-m}^{n-1} \binom{n-1}{i} [1-\theta+\theta y]^i [\theta(1-y)]^{n-1-i} - \alpha_S \right), \quad y \in [0, 1]. \quad (\text{B-2})$$

$\phi : [0, 1] \rightarrow [0, \tilde{x}]$ and $\zeta : [0, 1] \rightarrow [\tilde{x}, \hat{x}]$ are both increasing, smooth, and continuous. Thus, their inverse functions, $\phi^{-1} : [0, \tilde{x}] \rightarrow [0, 1]$ and $\zeta^{-1} : [\tilde{x}, \hat{x}] \rightarrow [0, 1]$ exist. In the concession equilibrium, $\text{Supp}_W = [0, \tilde{x}]$ and $\text{Supp}_F = [\tilde{x}, \hat{x}]$, and over the corresponding support, F_W and F_S are given by

$$F_W(x) = \phi^{-1}(x), \quad x \in [0, \tilde{x}]; \quad F_S(x) = \zeta^{-1}(x), \quad x \in [\tilde{x}, \hat{x}]. \quad (\text{B-3})$$

Proof. Result A-4(ii) shows the unique form of the probability-of-winning function, P , in a concession equilibrium. Note that, in a concession equilibrium, P must be concave and increasing over its support. Thus, to construct a concession equilibrium when $p_o^C \geq p_o^G$, we first show that, when $p_o^C \geq p_o^G$, P , given in Result A-4(ii), is concave and increasing over its support. By Result A-4(ii), this is equivalent to showing that $\beta_W \geq \beta_S > 0$, where β_W and β_S are given by (A-48) and (A-51) respectively.

Note that, by Result A-3, when $p_o^C \geq p_o^G$, $\mu_S \geq p_s^C \mu_W / p_o^C$. In this case,

$$\mu_S - (\tilde{p} \mu_W / p_o^C) \geq (p_s^C - \tilde{p}) \mu_W / p_o^C > 0, \quad (\text{B-4})$$

where the last inequality follows from (A-67). Equations (A-51), (A-67), and (B-4) imply that $\beta_S > 0$ and also that

$$\beta_S = \frac{p_s^C - \tilde{p}}{\mu_S - (\tilde{p} \mu_W / p_o^C)} \leq \frac{p_s^C - \tilde{p}}{(p_s^C - \tilde{p}) \mu_W / p_o^C} = \frac{p_o^C}{\mu_W} = \beta_W,$$

where the last equality follows from (A-48). Thus, when $p_o^C \geq p_o^G$, $\beta_W \geq \beta_S > 0$, implying that P , given in Result A-4(ii), is concave and increasing over its support.

Next, we show that F_W and F_S , constructed in (B-3), jointly produce the form of P shown in Result A-4(ii) and are CDFs that satisfy their capacity constraints. Note that the performance distribution chosen by a contestant of unknown type is given by

$$F(x) = \theta F_S(x) + (1 - \theta) F_W(x). \quad (\text{B-5})$$

In any symmetric equilibrium, no one places point mass. Thus, if a contestant has performance equal to x , her probability of besting any given rival of unknown type equals $F(x)$. To win a place, the contestant has to best at least $(n - m)$ out of her $(n - 1)$ rivals, whose types are

unknown to her and whose performances are independent. Thus, in any symmetric equilibrium, a contestant's probability-of-winning function, P , has a relation to F given by

$$P(x) = \sum_{i=n-m}^{n-1} \binom{n-1}{i} F(x)^i (1-F(x))^{n-1-i}. \quad (\text{B-6})$$

By construction, $\text{Supp}_W = [0, \tilde{x}]$ and $\text{Supp}_S = [\tilde{x}, \hat{x}]$. Thus, by (B-5), $F(x) = (1-\theta)F_W(x)$ for $x \in [0, \tilde{x}]$, and $F(x) = 1-\theta + \theta F_S(x)$ for $x \in [\tilde{x}, \hat{x}]$. Hence, by equations (B-1), (B-2), and (B-6),

$$P(x) = \begin{cases} \beta_W \phi \circ F_W(x) & x \in [0, \tilde{x}] \\ \alpha_S + \beta_S \zeta \circ F_S(x) & x \in [\tilde{x}, \hat{x}] \\ 1 & x \geq \hat{x} \end{cases}. \quad (\text{B-7})$$

When $F_W(x) = \phi^{-1}(x)$ for $x \in [0, \tilde{x}]$ and $F_S(x) = \zeta^{-1}(x)$ for $x \in [\tilde{x}, \hat{x}]$, P , given by (B-7), coincides with P given in Result A-4(ii). Thus, F_W and F_S , constructed in (B-3), jointly produce the form of P shown in Result A-4(ii).

Now we verify that F_W and F_S , constructed in (B-3), satisfy their capacity constraints. First, consider F_W . Let $\hat{\mu}_W$ be the mean of F_W constructed in (B-3). Note that

$$\begin{aligned} \hat{\mu}_W &= \int_{0-}^{\tilde{x}} x dF_W(x) = \int_{0-}^1 F_W^{-1}(y) dy = \int_{0-}^1 \phi(y) dy \\ &= \frac{1}{\beta_W} \int_{0-}^1 \sum_{i=n-m}^{n-1} \binom{n-1}{i} [(1-\theta)y]^i [1-(1-\theta)y]^{n-1-i} dy, \end{aligned} \quad (\text{B-8})$$

where the third equality follows from the construction of F_W in (B-3) and the last from (B-1). Also note that

$$\int_{0-}^1 \sum_{i=n-m}^{n-1} \binom{n-1}{i} [(1-\theta)y]^i [1-(1-\theta)y]^{n-1-i} dy = p_o^C, \quad (\text{B-9})$$

where p_o^C is given in (8). To see why (B-9) holds, note that the left-hand side of (B-9) can be interpreted as a weak contestant's probability of winning if she concedes to all strong rivals and both she and her weak rivals play a uniform performance distribution on $[0, 1]$. In this hypothetical contest, given that the given weak contestant has no chance of besting strong rivals and shares the same probability of winning with weak rivals, the given contestant's probability of winning must equal p_o^C , a weak contestant's probability of winning in a concession equilibrium. Equations (B-8) and (B-9) imply that $\hat{\mu}_W = p_o^C / \beta_W$. Thus, by (A-48), we must have $\hat{\mu}_W = \mu_W$ and, hence, the construction of F_W satisfies W 's capacity constraint. Next, by a similar argument, we can also verify that F_S , constructed in (B-3), satisfies S 's capacity constraint. We omit the detailed proof.

The above analysis verifies that, when $p_o^C \geq p_o^G$, there exist F_W and F_S , constructed in (B-3), that satisfy their capacity constraints and jointly produce an increasing and concave P in the form of (A-47). The concavity of P ensures that the constructed strategies are best replies and thus sustain an equilibrium. \square

Result B-2. Suppose $p_o^C < p_o^G$, where p_o^C and p_o^G are given by (8) and (6) respectively. Then challenge equilibria exist. In all of these challenge equilibria, the probability-of-winning function, P , is given by (A-46). Moreover, there exist positive constants, x^o , ρ_W , and ρ_S , where $x^o < n(\theta\mu_S + (1-\theta)\mu_W)/m$ and $\rho_W > \rho_S$, and distributions, F_W^o and F_S^o , with F_W^o supported by $[0, x^o]$ and F_S^o supported by $[x^o, n(\theta\mu_S + (1-\theta)\mu_W)/m]$, such that, in one of these challenge equilibria, each weak contestant plays F_W^o with probability ρ_W and plays F_S^o with probability $1 - \rho_W$ and each strong contestant plays F_W^o with probability ρ_S and plays F_S^o with probability $1 - \rho_S$ (the construction of such a challenge equilibrium is provided in the proof).

Proof. By Result A-4, in challenge equilibria, the probability-of-winning function, P , is the CDF of a uniform distribution given by (A-46). By Result A-3, $p_o^C < p_o^G$ is equivalent to $\mu_S/\mu_W < p_s^C/p_o^C$, where p_s^C is given by (A-42). Note that neither p_o^C nor p_s^C depend on μ_W or μ_S . To prove Result B-2, it suffices to construct a pair of CDFs, F_S and F_W , that produce a uniform P given by (A-46) and satisfy the specific characterization given in Result B-2 and their capacity constraints under the condition that $\mu_S/\mu_W < p_s^C/p_o^C$. Below we provide this construction.

Note that, when $\mu_S/\mu_W < p_s^C/p_o^C$, there always exists a unique pair of μ_S^o and μ_W^o such that

$$\mu_S^o > \mu_S > \mu_W > \mu_W^o > 0 \quad (\text{B-10})$$

$$\theta\mu_S^o + (1-\theta)\mu_W^o = \theta\mu_S + (1-\theta)\mu_W \quad (\text{B-11})$$

$$\mu_S^o/\mu_W^o = p_s^C/p_o^C. \quad (\text{B-12})$$

Consider an auxiliary contest where weak contestants have capacity μ_W^o and strong contestants have capacity μ_S^o . Because μ_W^o and μ_S^o satisfy (B-12), by Results A-3 and B-1, this auxiliary contest has a concession equilibrium and in this concession equilibrium, $P(x; \mu_W^o, \mu_S^o)$ is given by (A-47). Inspection of (A-46) and (A-47) reveals that, when μ_W^o and μ_S^o satisfy (B-12), we can equivalently express $P(x; \mu_W^o, \mu_S^o)$ using (A-46). Thus, by (B-5) and (B-6), if F_W^o and F_S^o represent, respectively, weak and strong contestants' equilibrium performance distributions in this auxiliary contest (see Result B-1 for the construction of F_W^o and F_S^o), F_W^o and F_S^o satisfy that

$$\sum_{i=n-m}^{n-1} \binom{n-1}{i} (\theta F_S^o(x) + (1-\theta)F_W^o(x))^i (1 - \theta F_S^o(x) - (1-\theta)F_W^o(x))^{n-1-i} = \min \left[\frac{m}{n(\theta\mu_S^o + (1-\theta)\mu_W^o)} x, 1 \right]. \quad (\text{B-13})$$

Now we verify that, when $\mu_S/\mu_W < p_s^C/p_o^C$, there exists a challenge equilibrium in which strong and weak contestants play as follows:

$$\text{W-strategy} = \begin{cases} F_W^o & \text{w. p. } \rho_W \\ F_S^o & \text{w. p. } 1 - \rho_W \end{cases} \quad \text{S-strategy} = \begin{cases} F_W^o & \text{w. p. } \rho_S \\ F_S^o & \text{w. p. } 1 - \rho_S \end{cases}, \quad (\text{B-14})$$

where

$$\rho_W := \frac{\mu_S^o - \mu_W}{\mu_S^o - \mu_W^o} \quad \& \quad \rho_S := \frac{\mu_S^o - \mu_S}{\mu_S^o - \mu_W^o}. \quad (\text{B-15})$$

First, note that, when $\mu_S/\mu_W < p_S^C/p_O^C$, (B-10) holds. In this case, by (B-15),

$$0 < \rho_S < \rho_W < 1. \quad (\text{B-16})$$

Next, note that F_W^O and F_S^O represent the equilibrium performance distributions played in the auxiliary contest, where weak contestants' capacity equals μ_W^O and strong contestants' capacity equals μ_S^O . Thus, the mean of F_W^O equals μ_W^O and the mean of F_S^O equals μ_S^O . Hence, by playing the strategies presented in (B-14), weak contestants' mean performance equals $\rho_W \mu_W^O + (1 - \rho_W) \mu_S^O$, which, given the definition of ρ_W in (B-15), equals μ_W , and strong contestants' mean performance equals $\rho_S \mu_W^O + (1 - \rho_S) \mu_S^O$, which, given the definition of ρ_S in (B-15), equals μ_S . Thus, the strategies presented in (B-14) satisfy the capacity constraints. By (B-16), the strategies in (B-14) give weak contestants a positive probability of besting strong contestants.

Finally, we show that the strategies presented in (B-14) jointly produce a uniform P given in (A-46). Note that, when the two types use the strategies presented in (B-14), the performance distribution played by a contestant of unknown type, F , satisfies

$$\begin{aligned} F(x) &= \theta F_S(x) + (1 - \theta) F_W(x) \\ &= (\theta \rho_S + (1 - \theta) \rho_W) F_W^O(x) + (1 - \theta \rho_S - (1 - \theta) \rho_W) F_S^O(x). \end{aligned} \quad (\text{B-17})$$

Also note that

$$\begin{aligned} \theta \rho_S + (1 - \theta) \rho_W &= \theta \left(\frac{\mu_S^O - \mu_S}{\mu_S^O - \mu_W^O} \right) + (1 - \theta) \left(\frac{\mu_S^O - \mu_W}{\mu_S^O - \mu_W^O} \right) \\ &= \frac{\mu_S^O - (\theta \mu_S + (1 - \theta) \mu_W)}{\mu_S^O - \mu_W^O} = \frac{\mu_S^O - (\theta \mu_S^O + (1 - \theta) \mu_W^O)}{\mu_S^O - \mu_W^O} = 1 - \theta, \end{aligned} \quad (\text{B-18})$$

where the first equality follows from (B-15) and the second last from (B-11). Equations (B-17) and (B-18) imply that

$$F(x) = \theta F_S(x) + (1 - \theta) F_W(x) = \theta F_S^O(x) + (1 - \theta) F_W^O(x). \quad (\text{B-19})$$

Let $P(\cdot; \mu_W, \mu_S)$ be the probability-of-winning function produced by F_W and F_S . Equations (B-6), (B-13), and (B-19) imply that

$$P(x; \mu_W, \mu_S) = \min \left[\frac{m}{n(\theta \mu_S^O + (1 - \theta) \mu_W^O)} x, 1 \right].$$

This further implies, given (B-11), that

$$P(x; \mu_W, \mu_S) = \min \left[\frac{m}{n(\theta \mu_S + (1 - \theta) \mu_W)} x, 1 \right].$$

Thus, the strategies presented in (B-14) jointly produce a uniform P given in (A-46). The result thus follows. \square

C. Exogenous upper bound on performance

Lemma C-1. *Imposing an exogenous upper bound, \bar{x} , on performance, $\bar{x} \geq \mu_S$, does not change any result in Lemma 2.*

Proof. Let \bar{x} be the exogenous upper bound on performance. Suppose that $\bar{x} \geq \mu_S$. Then any symmetric equilibrium has the following properties:

No point mass on zero performance This follows from the same contradiction argument used in the verification of Claim 1 for showing no point mass on zero performance when there is no exogenous upper bound on performance.

No point mass on any performance level on $(0, \bar{x})$ This follows from applying the contradiction argument used in the verification of Claim 1 for showing no point mass on any positive performance level in the case of no exogenous upper bound on performance to any performance level on $(0, \bar{x})$.

P is continuous over $[0, \bar{x})$. This follows from the fact that, as established above, no contestant places point mass on any $x < \bar{x}$.

Thus, only \bar{x} can possibly be a discontinuity point for P . Because contestants are not allowed to place weight on performance above \bar{x} and because P is continuous over $[0, \bar{x})$, the set of optimal fair gambles is closed. Thus, the support of a contestant's equilibrium performance distribution is contained in the set of optimal fair gambles for the contestant. This fact, combined with the argument used around Figure 1 in the main text, implies that the collinearity result stated in Lemma 1, which holds in the case of no exogenous upper bound on performance, also holds when there is an exogenous upper bound, $\bar{x} \geq \mu_S$, on performance.

Result C-1. Suppose that there is an exogenous upper bound, $\bar{x} \geq \mu_S$, on performance. In any symmetric equilibrium, for each type- $t \in \{S, W\}$ contestant, a fair gamble between x' and x'' , $x', x'' \in \text{Supp}_t$, is a best reply, and all performance/probability-of-winning pairs, $(x, P(x))$, such that $x \in \text{Supp}_t$ are collinear.

The lower bound of the support of P is zero. By equation (1), which holds regardless of whether performance is exogenously bounded or not, there exists at least one type, say t , such that $\min \text{Supp}_t = \min \text{Supp } P$. Because, as established above, no contestant places point mass on any $x < \bar{x}$, and because $\mu_W < \mu_S \leq \bar{x}$, weak contestants place no point mass on μ_W . Thus, given that a weak contestant's expected performance equals μ_W , it must be the case that $\min \text{Supp}_W < \mu_W$. This implies, by equation (1), that $\min \text{Supp } P < \mu_W$. Hence, given that $\min \text{Supp}_t = \min \text{Supp } P$ and $\mu_W < \mu_S$, it must be that the case that $\min \text{Supp}_t < \mu_t$. This implies, given that the mean of F_t equals μ_t , the existence of $x^o > \mu_t$ such that $x^o \in \text{Supp}_t$. The result that $\min \text{Supp } P = 0$ then follows from the same contradiction argument used in the verification of Claim 1 for showing zero as the lower bound of the support of P in the case of no exogenous upper bound on performance.

$\text{Supp } P \setminus \{\bar{x}\}$ is a **connected set**. Otherwise, there would exist $0 \leq x' < x'' < \bar{x}$ such that $[x', x''] \cap \text{Supp } P = \{x', x''\}$. In this case, by equation (1), there would exist at least one type, say t , such that $(x', x''] \cap \text{Supp}_t = \{x''\}$. Because, by hypothesis, $x'' < \bar{x}$, and because, as has been established, no contestant places point mass on any $x \in [0, \bar{x})$, $(x', x''] \cap \text{Supp}_t = \{x''\}$ would imply the existence of $\varepsilon > 0$ such that $[x'', x'' + \varepsilon] \in \text{Supp}_t$. The connectedness of $\text{Supp } P \setminus \{\bar{x}\}$ then follows from a similar contradiction argument used in the verification of Claim 1 for showing the connectedness of $\text{Supp } P$ in the case of no exogenous upper bound on performance.

These properties of P , combined with Result C-1, allow us to show the following result:

Result C-2. In any challenge equilibrium, all pairs $(x, P(x))$ such that $x \in \text{Supp } P$ must be collinear.

Proof. First, note that, in any challenge equilibrium, it must be the case that $\text{Supp}_W \cap \text{Supp}_S \neq \emptyset$. This is because, by equation (1) and the facts that $\text{Supp } P \setminus \{\bar{x}\}$ is a connected set and both Supp_W and Supp_S are closed, if $\text{Supp}_W \cap \text{Supp}_S = \emptyset$, it had to be that one type places all weight on \bar{x} and the other type places all weight over some interval lying strictly below \bar{x} . This would imply that the former type always bests the latter type, contradicting the fact that, in a challenge equilibrium, each type has a positive probability of besting the other type.

Next, given that $\text{Supp}_W \cap \text{Supp}_S \neq \emptyset$, there are only two cases to consider: (1) the case in which $\text{Supp}_W \cap \text{Supp}_S$ contains more than one point and (2) the case in which $\text{Supp}_W \cap \text{Supp}_S$ consists of a single point. We first consider case (1). Suppose that, in a challenge equilibrium, there exist two distinct points x' and x'' such that $x', x'' \in \text{Supp}_W \cap \text{Supp}_S$. Then by Result C-1, all pairs $(x, P(x))$ such that $x \in \text{Supp}_t$, $t = S, W$, are collinear with $(x', P(x'))$ and $(x'', P(x''))$. In this case, by equation (1), all pairs $(x, P(x))$ such that $x \in \text{Supp } P$ must be collinear.

Now consider case (2). Suppose that, in a challenge equilibrium, $\text{Supp}_W \cap \text{Supp}_S = x^0$. In case (2), there are two subcases to consider: (a) $x^0 = \bar{x}$, and (b) $x^0 < \bar{x}$. First consider the subcase in which $x^0 = \bar{x}$ (i.e., $\text{Supp}_W \cap \text{Supp}_S = \bar{x}$). Note that, by equation (1) and the connectedness of $\text{Supp } P \setminus \{\bar{x}\}$, $\text{Supp}_W \cup \text{Supp}_S \setminus \{\bar{x}\}$ is connected. Because, by hypothesis, $\text{Supp}_W \cap \text{Supp}_S = \bar{x}$, it must be the case that $\text{Supp}_W \cap \text{Supp}_S \setminus \{\bar{x}\} = \emptyset$. Thus, given that $\text{Supp}_W \cup \text{Supp}_S \setminus \{\bar{x}\}$ is connected and both Supp_W and Supp_S are closed, it must be the case that either $\text{Supp}_W \setminus \{\bar{x}\} = \emptyset$ or $\text{Supp}_S \setminus \{\bar{x}\} = \emptyset$. Because $\text{Supp}_W \setminus \{\bar{x}\} = \emptyset$ would violate the weak type's capacity constraint, it must be the case that $\text{Supp}_S \setminus \{\bar{x}\} = \emptyset$, implying that $\text{Supp}_S = \bar{x}$ (clearly, this subcase can happen only when $\bar{x} = \mu_S$). Given $\text{Supp}_S = \bar{x}$, for this equilibrium to be a challenge equilibrium, it must be that $\bar{x} \in \text{Supp}_W$, implying that $\text{Supp}_S \subset \text{Supp}_W$. Thus, by equation (1), $\text{Supp } P = \text{Supp}_W$. Hence, given that, by Result C-1, all pairs $(x, P(x))$ such that $x \in \text{Supp}_W$ are collinear, all pairs $(x, P(x))$ such that $x \in \text{Supp } P$ must be collinear.

Next, consider the subcase in which $x^0 < \bar{x}$. In this subcase, it must be that at least one type places point mass on \bar{x} . This is because, otherwise, $\text{Supp } P$ would have to be connected. This would imply, given $\text{Supp}_W \cap \text{Supp}_S = x^0 < \bar{x}$ and the fact that no contestant places point mass on any $x < \bar{x}$, that the equilibrium under consideration is not a challenge equilibrium but

a concession equilibrium where $\max \text{Supp}_W = \min \text{Supp}_S$, a contradiction.

Thus, at least one type, say t , places point mass on \bar{x} . Then the hypothesis that $\text{Supp}_W \cap \text{Supp}_S = x^o < \bar{x}$ implies that the other type, denoted by $-t$, must not place point mass on \bar{x} . Then, given the facts that (a) $\text{Supp } P \setminus \{\bar{x}\}$ is connected with $\min \text{Supp } P = 0$, (b) no contestant places point mass on any $x < \bar{x}$, and (c) in a challenge equilibrium, each type has a positive probability of besting the other type, there must exist $x' \in (x^o, \bar{x})$ such that

$$\text{Supp}_t = [0, x^o] \cup \{\bar{x}\} \quad \& \quad \text{Supp}_{-t} = [x^o, x']. \quad (\text{C-1})$$

Then by Result C-1,

$$\text{all pairs } (x, P(x)) \text{ such that } x \in [0, x^o] \cup \{\bar{x}\} \text{ are collinear} \quad (\text{C-2})$$

$$\& \quad \text{all pairs } (x, P(x)) \text{ such that } x \in [x^o, x'] \text{ are collinear.} \quad (\text{C-3})$$

By (C-2) and (C-3), over $[0, x']$, P is piecewise linear, with x^o being the only possible kink point. Note that the slope of P over $(0, x^o)$ must be no less than the slope of P over (x^o, x') . This is because, otherwise, P would be convex and nonlinear over $[0, x']$, in which case, by Jensen's inequality, type $-t$ would be strictly better off deviating to a fair gamble between 0 and x' , a contradiction. Note also that the slope of P over $(0, x^o)$ must be no greater than the slope of P over (x^o, x') . This is because, by (C-2), the slope of P over $(0, x^o)$ equals the slope of the line connecting the two points, $(x^o, P(x^o))$ and $(\bar{x}, P(\bar{x}))$. Thus, if this slope were greater than the slope of P over (x^o, x') , type $-t$ would be strictly better off deviating to a fair gamble between x^o and \bar{x} , a contradiction. Therefore, the slope of P over $(0, x^o)$ has to equal the slope of P over (x^o, x') . This implies, by equations (1), (C-1), (C-2), and (C-3), that all pairs $(x, P(x))$ such that $x \in \text{Supp } P$ are collinear.

The above analysis exhausts all possible configurations of challenge equilibria and the result thus follows. \square

Result C-2 and the fact that $\min \text{Supp } P = 0$ imply that, in any challenge equilibrium (if exists), it is a best reply for a weak contestant to play a “mimicking strategy,” i.e., a strategy that gives her a strong type's performance distribution with probability μ_W/μ_S and zero performance with the complementary probability. Because no contestant places point mass on 0, we have $P(0) = 0$. Thus, in a challenge equilibrium, $p_W = (\mu_W/\mu_S) p_S$. Then, by equation (3), in a challenge equilibrium, a weak contestant's probability of winning equals p_o^C given by (6). Note that a weak contestant's probability of winning in a concession equilibrium still equals p_o^C given by (8). The argument developed in Section 3.2 for showing equation (9) still holds. Thus, equation (9) still holds, i.e., it is still the case that, in any challenge equilibrium, $p_W > p_o^C$. Note also that, in any concession equilibrium, a weak contestant must have no incentive to deviate to the “mimicking strategy.” Because, by playing the “mimicking strategy,” a weak contestant has a probability of winning equal to $(\mu_W/\mu_S) p_S$, the no-profitable-deviation condition implies that, in any concession equilibrium, $p_W \geq (\mu_W/\mu_S) p_S$. This condition, combined with equation (3), implies that equation (7) still holds, i.e., it is still the case that, in any concession

equilibrium, $p_W \geq p_o^G$.

Because equations (6)–(9) still hold, the fact that every symmetric equilibrium is either a concession or a challenge equilibrium implies that, if $p_o^C \geq (<) p_o^G$, only concession (challenge) equilibria exist. The lemma thus follows. \square

D. Pool expansion by including less promising candidates

In Section 5, we studied the effect of risk-taking on the optimal size of the contestant pool. We showed that, with the selection quota fixed, if the contestant pool is sufficiently large, adding more contestants who are as likely to exhibit ability as the contestants in the original pool does not affect the expected number of strong selected contestants. 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 makes selection less meritocratic if the external candidates are less likely to exhibit ability than the contestants in the original pool.⁴

Result D-1. Suppose that the designer can only expand the contestant pool by including external candidates whose prior quality (measured by the probability of being strong, θ) is lower than the internal candidates'. If the contest between only the internal candidates has challenge equilibria, pool expansion reduces designer welfare (10).

Result D-1 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 supports challenge equilibria. Result D-1 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.

Proof of Result D-1. Let θ (θ') be the probability of being strong for each 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 challenge equilibria.

We show that, in any symmetric equilibrium of the expanded contest, the expanded contest has lower winner quality, measured by the probability that a selected contestant is strong, 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 it is feasible for a weak internal candidate to choose a performance distribution that mimics a strong internal candidate's performance distribution with probability μ_W/μ_S and produces zero performance with the complementary probability, it must be the case that

$$\hat{p}_W \geq (\mu_W/\mu_S) \hat{p}_S. \quad (\text{D-1})$$

Analogously,

$$\hat{p}'_W \geq (\mu_W/\mu_S) \hat{p}'_S. \quad (\text{D-2})$$

⁴We still focus our analysis on symmetric Nash equilibria (“symmetric” in the sense that each type- $t \in \{S, W\}$ internal candidate plays the same strategy and each type- $t \in \{S, W\}$ external candidate plays the same strategy).

Equations (D-1) and (D-2) imply that

$$n\theta\hat{p}_S + n'\theta'\hat{p}'_S \leq n\theta\hat{p}_W(\mu_S/\mu_W) + n'\theta'\hat{p}'_W(\mu_S/\mu_W). \quad (\text{D-3})$$

Let $\hat{\Pi}(n, n')$ be winner quality after adding n' contestants to the original contest with n contestants. Note that

$$\begin{aligned} \hat{\Pi}(n, n') &= \frac{n\theta\hat{p}_S + n'\theta'\hat{p}'_S}{n\theta\hat{p}_S + n'\theta'\hat{p}'_S + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W} \\ &\leq \frac{n\theta\hat{p}_W(\mu_S/\mu_W) + n'\theta'\hat{p}'_W(\mu_S/\mu_W)}{n\theta\hat{p}_W(\mu_S/\mu_W) + n'\theta'\hat{p}'_W(\mu_S/\mu_W) + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W}, \end{aligned} \quad (\text{D-4})$$

where the equality follows from the fact that winner quality equals the expected number of strong winners divided by the sum of the expected number of strong winners and the expected number of weak winners, and the inequality follows from (D-3) and the fact that, for any fixed $b > 0$, $f(a) = a/(a+b)$ is increasing in a for $a > 0$. Let $\Pi(n)$ be winner quality in the original contest. Note that, if the original contest has challenge equilibria, then in the original contest, the probability that a weak contestant's probability of winning equals p_o^G given by (6). Thus, by Bayes rule, the probability that a winner is weak equals $(1-\theta)p_o^G/(m/n)$. Hence, given that winner quality (i.e., the probability that a winner is strong) equals one minus the probability that a winner is weak, if the original contest has challenge equilibria,

$$\Pi(n) = 1 - \frac{(1-\theta)p_o^G}{m/n} = 1 - \frac{(1-\theta)\mu_W}{\theta\mu_S + (1-\theta)\mu_W} = \frac{\theta(\mu_S/\mu_W)}{\theta(\mu_S/\mu_W) + 1 - \theta}. \quad (\text{D-5})$$

where the second equality follows from equation (6).

By equations (D-4) and (D-5), if the original contest has challenge equilibria, then for any $\theta' < \theta$ and $n, n' > 0$,

$$\begin{aligned} &\hat{\Pi}(n, n') - \Pi(n) \\ &\leq \frac{n\theta\hat{p}_W(\mu_S/\mu_W) + n'\theta'\hat{p}'_W(\mu_S/\mu_W)}{n\theta\hat{p}_W(\mu_S/\mu_W) + n'\theta'\hat{p}'_W(\mu_S/\mu_W) + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W} - \frac{\theta(\mu_S/\mu_W)}{\theta(\mu_S/\mu_W) + 1 - \theta} \\ &= \frac{n'\hat{p}'_W(\mu_S/\mu_W)(\theta' - \theta)}{\left(n\theta\hat{p}_W(\mu_S/\mu_W) + n'\theta'\hat{p}'_W(\mu_S/\mu_W) + n(1-\theta)\hat{p}_W + n'(1-\theta')\hat{p}'_W\right)\left(\theta(\mu_S/\mu_W) + 1 - \theta\right)} \\ &< 0. \end{aligned}$$

The result then follows immediately from the fact that, fixing m , designer welfare (10) is maximized by maximizing winner quality. \square

E. Generalized capacity constraints

In the main text, we assumed that each type- $t \in \{S, W\}$ contestant's expected performance must equal the type- t contestant's capacity, μ_t . In this case, risk taking does not affect mean performance. In this appendix, we generalize the capacity constraint by imposing the mean constraint not on performance but on a continuous and increasing function, $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, of performance, where $g(0) = 0$ and $\lim_{x \rightarrow \infty} g(x) = \infty$. In particular, we assume that a type- t contestant's choice of her random performance, X_t , must satisfy that

$$\mathbb{E}[g(X_t)] = g(\mu_t), \quad t = S, W. \quad (\text{E-1})$$

Under this specification, choosing a fixed performance level equal to μ_t is feasible to a type- t contestant. Thus, μ_t represents a type- t contestant's performance if she takes no risk. If g is linear, we revert to our previous model, which imposes a mean constraint on performance. If g is strictly convex, then attaining high performance consumes disproportionately more capacity than attaining low performance. In this case, spreading out performance will require a reduction in mean performance in order to satisfy the generalized capacity constraint (E-1). Thus, strictly convex g captures situations in which increasing performance riskiness reduces expected performance.

For example, in Section 2, we assumed that a weak student, by taking no risk, receives 50 marks with certainty. Under our original specification of the capacity constraint, which requires a weak student's mean performance to be fixed at 50, the performance distribution that gives a weak student 100 marks with half probability and zero marks with half probability is feasible to the weak type. In contrast, if $g(x) = x^2$, then this performance distribution will violate the generalized capacity constraint, (E-1), and thus will no longer be feasible to the weak type.⁵ In fact, if $g(x) = x^2$, then to take extremely high risk by placing all weight on zero and 100 marks, a weak student's chance of attaining 100 marks is only 25%, implying an expected mark equaling only 25 (despite the fact that, by playing safe, a weak student receives 50 marks).⁶

It is also worth noting that contest situations in which contestants are restricted to risk-taking strategies with limited riskiness (e.g., variance) can be modeled under a generalized capacity constraint even without introducing an exogenous upper bound on performance. For example, when $g(x) = x^2$, the generalized capacity constraint implies that $\mathbb{E}[X_t^2]$ equals a constant. Given the fact that $\text{Var}[X_t] \leq \mathbb{E}[X_t^2]$, this constraint thus implies an upper bound on the variance of X_t .

Generalizing the capacity constraint, using (E-1), in fact does not alter any of our qualitative conclusions on how changes of competitiveness affect meritocracy. To see this, note that, because $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is continuous and increasing, there is a 1-1 correspondence between performance, x , and *transformed performance*, $y = g(x)$. Because the transformation,

⁵This violation follows from the fact that $(1/2) \times 0^2 + (1/2) \times 100^2 = 5000 > 2500 = 50^2$.

⁶The satisfaction of the constraint (E-1) follows from the fact that $(3/4) \times 0^2 + (1/4) \times 100^2 = 2500 = 50^2$.

g , is increasing, the ranking of transformed performance equals the ranking of performance. Thus, selection based on performance ranks is the same as selection based on transformed-performance ranks. The constraint (E-1) can be interpreted as a mean constraint imposed on transformed performance, and the space of transformed performance is $[0, \infty)$. Hence, if we think of contestants as competing by choosing random transformed-performance, such a competition is exactly the risk-taking contest we analyzed in the main text with a type- $t \in \{S, W\}$ contestant's capacity equaling $g(\mu_t)$. Therefore, selection outcomes under the generalized capacity constraint (E-1) are the same as selection outcomes under our original capacity constraint with capacity being $g(\mu_t)$ for a type- $t \in \{S, W\}$ contestant. With both types' capacities fixed, how changes of contest structure affect meritocracy follow from our analysis made in the main text.⁷

⁷In the main text, where we imposed the mean constraint on performance, the map from performance to the probability of winning in a symmetric equilibrium is given by $x \mapsto P(x)$. In contrast, this map, in our generalized capacity-constraint framework, does not characterize the relation between performance and the equilibrium probability of winning but characterizes the relation between *transformed performance* and the equilibrium probability of winning; the map from performance to the equilibrium probability of winning is instead given by $x \mapsto (P \circ g)(x)$.

F. Nonlinear designer welfare specification and under-selection

This appendix provides an example in which designer welfare has a nonlinear specification and contestant risk taking inclines the designer to “under-select,” i.e., select fewer but a positive number of contestants than what would be optimal in the absence of risk taking.

Assume that the designer cannot vary contest size, n , but chooses the quota, m , to maximize

$$u = \mathbb{E}[\text{\#Strong selected contestants} - (\text{\#Weak selected contestants})^2]. \quad (\text{F-1})$$

This welfare specification differs from the symmetric linear specification, (10), in the following way: in (10), the designer’s cost of selecting weak contestants is linear in the number of weak selected contestants. In contrast, in (F-1), this cost is quadratic and, hence, the cost of selecting many weak contestants is disproportionately larger than the cost of selecting few weak contestants. Consequently, under the nonlinear specification, (F-1), the designer cares about not only the expected number of weak selected contestants but also the variability of the number of weak selected contestants. The latter consideration is absent under the linear specification, (10). As we show below, this consideration can possibly render under-selection optimal.

Suppose that there are $n = 3$ contestants. Each contestant is strong with probability $\theta = 3/5$ and weak with probability $1 - \theta = 2/5$. Contestant types are independent. We first compute the optimal quota in the absence of risk taking. Because this computation does not require the knowledge of capacity levels, μ_S and μ_W , we postpone making assumptions on μ_S and μ_W until we come to the derivation of the optimal quota for the risk-taking contest.

Non-strategic contestants Let $q_{i,n-i}$ be the probability that there are i strong contestants and $n - i$ weak contestants in the contestant pool. Note that

$$q_{i,n-i} = \binom{n}{i} \theta^i (1 - \theta)^{n-i}, \quad i = 0, \dots, n. \quad (\text{F-2})$$

When contestants are non-strategic, strong contestants are prioritized for selection. In this case, selecting m contestants from a contestant pool with i strong contestants and $n - i$ weak contestants will result in $\min[i, m]$ strong contestants selected and $\max[0, m - i]$ weak contestants selected. Thus, under the designer welfare specification provided by (F-1), we have

$$u(m) = \sum_{i=0}^n q_{i,n-i} (\min[i, m] - (\max[0, m - i])^2). \quad (\text{F-3})$$

Equations (F-2) and (F-3), combined with the assumptions that $n = 3$ and $\theta = 3/5$, imply that $u(0) = 0$, $u(1) = 109/125$, $u(2) = 26/25$, and $u(3) = -9/25$. Thus, when contestants are non-strategic, the optimal quota is 2.

Strategic contestants Now consider the case in which contestants are strategic. Note that, if a contest has a challenge equilibrium, it will have multiple challenge equilibria. While all of these challenge equilibria produce the same expected number of strong selected contestants and the same expected number of weak selected contestants and, hence, produce the same designer welfare under our linear welfare specification, (10), these challenge equilibria produce different

distributions of the number of weak selected contestants (with the same mean). Thus, given that, under the nonlinear specification, (F-1), designer welfare depends on the variability of the number of weak selected contestants, these challenge equilibria can produce different levels of designer welfare under the nonlinear specification, (F-1). For this reason, it is difficult to make a general analysis of the effect of risk taking on the optimal quota when designer welfare takes the form of (F-1).

However, the purpose of our discussion here is not to make a general analysis but to show, by constructing an example, that there exists a nonlinear designer welfare specification that can possibly render under-selection optimal. To make this example simple, we consider a special case of our risk-taking contest by assuming that μ_W is very close to μ_S . Our equilibrium analysis implies that, when μ_W is sufficiently close to μ_S , weak contestants will always challenge strong contestants for any quota $m = 1, \dots, n-1$, and as $\mu_W/\mu_S \rightarrow 1$, weak and strong contestants will have the same probability of being selected. Thus, when μ_W is sufficiently close to μ_S , designer welfare will be sufficiently close to the one from random selection. Below we compute designer welfare from random selection for different quotas.

With some abuse of notation, let $p_{i,n-i,j}(m)$ be the probability of selecting $j \leq m$ weak contestants from a contestant pool of i strong contestants and $n-i$ weak contestants through randomly selecting m contestants. Note that, for any $0 \leq j \leq m$,

$$p_{i,n-i,j}(m) = \begin{cases} \frac{\binom{n-i}{j} \binom{i}{m-j}}{\binom{n}{m}} & \text{if } m-i \leq j \leq n-i \\ 0 & \text{otherwise} \end{cases}. \quad (\text{F-4})$$

Also note that, under random selection, the expected number of strong selected contestants given quota m equals $m\theta$.

Thus, under random selection, designer welfare specified in (F-1) can be expressed as

$$u(m) = m\theta - \sum_{i=0}^n q_{i,n-i} \left(\sum_{j=0}^m j^2 p_{i,n-i,j}(m) \right). \quad (\text{F-5})$$

Equations (F-2), (F-4), and (F-5), combined with the assumptions that $n = 3$ and $\theta = 3/5$, imply that under random selection, $u(0) = 0$, $u(1) = 1/5$, $u(2) = 2/25$, and $u(3) = -9/25$. Thus, the optimal quota under random selection is 1, which is smaller than the optimal quota when strong contestants are prioritized for selection.

Because, when contestants are strategic, selection will be sufficiently close to random selection if μ_W/μ_S is sufficiently close to 1, the above analysis implies that, under the nonlinear designer welfare specification, (F-1), under-selection is optimal when $n = 3$, $\theta = 3/5$, and μ_W/μ_S is sufficiently close to 1.⁸

⁸It is easy to verify that, if designer welfare follows the linear specification, (10), then the optimal quota when strong contestants are prioritized for selection is 2, while the optimal quota under random selection is 3 (full selection). Thus, applying the same parametric assumptions to the linear specification, (10), does not lead to the optimality of under-selection.

G. A more general analysis of the contest size effect when capacity is endogenous

In Section 5, we argued, based on our baseline model, that contestant risk taking caps a meritocratic designer's gains from expanding the contestant pool. In Section 6.2, we showed that this finding is robust to the introduction of capacity acquisition if the effort cost functions for acquiring capacity are power functions. In this appendix, we extend this analysis to more general effort cost functions. Throughout this appendix, we assume that the cost of choosing the capacity level, μ , for a type- $t \in \{S, W\}$ contestant is $c_t(\mu) = c(\mu)/a_t$, $a_S > a_W > 0$, and $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is three times differentiable, increasing, and strictly convex, with $c(0) = 0$. We will consider cost functions that have either geometrically convex or geometrically concave marginal cost functions. If c' is (strictly) geometrically convex, $x \mapsto xc''(x)/c'(x)$ is nondecreasing (increasing). If c' is (strictly) geometrically concave, $x \mapsto xc''(x)/c'(x)$ is nonincreasing (decreasing). When effort cost functions are power functions, c' is also a power function and is thus geometrically linear (i.e., both geometrically convex and geometrically concave).

In what follows, we argue that the optimality of a limited contestant pool is robust to the introduction of capacity acquisition if c' is geometrically concave but not if c' is strictly geometrically convex. The key to this argument is the next result.

Lemma G-1. Fixing the quota, m , if challenge equilibria exist at contest size n' , then

- i. if c' is (strictly) geometrically concave, challenge equilibria also exist at any size $n > n'$ and the capacity ratio, μ_S/μ_W , at any size $n > n'$ is (smaller) no greater than at n' ;
- ii. if c' is strictly geometrically convex and if challenge equilibria also exist at size $n'' > n'$, the capacity ratio, μ_S/μ_W , at size n'' is greater than at n' .

We postpone the proof of Lemma G-1 to the end of this appendix. Part (i) of the lemma implies that, if c' is geometrically concave, once contest size is sufficiently large such that challenge equilibria are triggered, challenge equilibria will also be supported at any larger contest size and the capacity ratio, μ_S/μ_W , will be weakly smaller at any larger contest size. By equation (A-73), reduced capacity ratio increases the probability that a selected contestant is weak, $\mathbb{P}[W|\text{selected}]$, in challenge equilibria. As has been argued in Sections 5 and 6.1, with the quota, m , fixed, maximizing designer welfare, (15), is equivalent to maximizing the expected number of strong selected contestants. When the quota is fixed, the latter further equates to minimizing $\mathbb{P}[W|\text{selected}]$. Thus, part (i) implies that, if c' is geometrically concave, once contest size is sufficiently large such that weak contestants challenge strong contestants, a meritocratic designer cannot gain by further increasing contest size.

In contrast, if c' is strictly geometrically convex, then even when contest size supports challenge equilibria, the designer can still further meritocracy by increasing contest size. The reason is as follows. Suppose that c' is strictly geometrically convex and that challenge equilibria exist at contest size n' . Consider any larger size, $n'' > n'$. If n'' also supports challenge

equilibria, then by part (ii) of Lemma G-1, the capacity ratio, μ_s/μ_w , will be greater at size n'' than at size n' . By (A-73), increased capacity ratio reduces $\mathbb{P}[W|\text{selected}]$ in challenge equilibria. Because, with the quota, m , fixed, maximizing designer welfare, (15), is equivalent to minimizing $\mathbb{P}[W|\text{selected}]$, designer welfare is higher at size n'' than at size n' . Now consider the case in which the larger size, n'' , supports concession equilibria. Recall that selection when weak contestants concede to strong contestants is equivalent to selection when contestants are non-strategic. Because increasing contest size increases designer welfare when contestants are non-strategic, and also because, for any fixed quota, m , and size, n , designer welfare when contestants are non-strategic serves as an upper bound on designer welfare when contestants are strategic, if n'' supports concession equilibria, then designer welfare at n'' must be higher than at any size $n < n''$. Thus, given that $n' < n''$, designer welfare is higher at size n'' than at size n' .

Proof of Lemma G-1. The proof requires the following technical result.

Result G-1. Suppose that function $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is three times differentiable, increasing, and strictly convex, with $c(0) = 0$. Let $k > 1$ be a constant and define $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ as $g(x) = (c')^{-1} \circ (k c'(x))$. Then $x \mapsto g(x)/x$ is nonincreasing (decreasing) if c' is (strictly) geometrically convex, and $x \mapsto g(x)/x$ is nondecreasing (increasing) if c' is (strictly) geometrically concave.

Proof. Note that $g(x) = (c')^{-1} \circ (k c'(x))$ implies that

$$\frac{c' \circ g(x)}{c'(x)} = k, \quad x > 0. \quad (\text{G-1})$$

Also note that, for all $x > 0$,

$$\begin{aligned} \left(\frac{g(x)}{x} \right)' &\stackrel{\text{sign}}{=} \frac{k c''(x) x}{c'' \circ g(x)} - g(x) \stackrel{\text{sign}}{=} \frac{x c''(x) c' \circ g(x)}{c'(x) c'' \circ g(x)} - g(x) \\ &= \frac{c' \circ g(x)}{c'' \circ g(x)} \left(\frac{x c''(x)}{c'(x)} - \frac{g(x) c'' \circ g(x)}{c' \circ g(x)} \right) \stackrel{\text{sign}}{=} \frac{x c''(x)}{c'(x)} - \frac{g(x) c'' \circ g(x)}{c' \circ g(x)}, \end{aligned} \quad (\text{G-2})$$

where the second equality follows from (G-1). To understand the last equality, simply note that, by hypothesis, c is increasing and strictly convex. Thus, given that $g(x) > 0$ for all $x > 0$, $c' \circ g(x) > 0$ and $c'' \circ g(x) > 0$ for all $x > 0$. The last equality thus follows.

Because c is strictly convex, $x \mapsto c'(x)$ is increasing. Thus, by equation (G-1) and the hypothesis that $k > 1$, it must be that

$$g(x) > x, \quad x > 0. \quad (\text{G-3})$$

When c' is (strictly) geometrically convex, $x \mapsto x c''(x)/c'(x)$ is nondecreasing (increasing). In this case, by (G-3),

$$\frac{x c''(x)}{c'(x)} \leq (<) \frac{g(x) c'' \circ g(x)}{c' \circ g(x)}, \quad x > 0. \quad (\text{G-4})$$

Equations (G-2) and (G-4) imply that, when c' is (strictly) geometrically convex, $(g(x)/x)' \leq (<) 0$ for $x > 0$, implying that $x \mapsto g(x)/x$ is nonincreasing (decreasing) for $x > 0$.

When c' is (strictly) geometrically concave, $x \mapsto xc''(x)/c'(x)$ is nonincreasing (decreasing). In this case, by a similar argument, $(g(x)/x)' \geq (>)0$ for $x > 0$, implying that $x \mapsto g(x)/x$ is nondecreasing (increasing) for $x > 0$. \square

Given Result G-1, we first establish part (i) of the lemma. Let v be the individual prize value. By Results A-1, A-2, and A-4, challenge equilibria exist at contest size, n , if and only if there exists a pair of $\mu_S > 0$ and $\mu_W > 0$ that solves the following two first-order conditions,

$$\frac{c'(\mu_S)}{a_S} = \frac{vm}{n(\theta\mu_S + (1-\theta)\mu_W)} \quad (\text{G-5})$$

$$\frac{c'(\mu_W)}{a_W} = \frac{vm}{n(\theta\mu_S + (1-\theta)\mu_W)}, \quad (\text{G-6})$$

and satisfies $\mu_S/\mu_W < p_s^C(n)/p_o^C(n)$, where p_s^C and p_o^C are defined in (A-42) and (8), respectively, and treated as functions of n (note that p_s^C and p_o^C do not vary with μ_S or μ_W).

Equations (G-5) and (G-6) imply that

$$\frac{c'(\mu_S)}{c'(\mu_W)} = \frac{a_S}{a_W},$$

which further implies that

$$\mu_S = (c')^{-1} \circ (a_S/a_W c'(\mu_W)). \quad (\text{G-7})$$

Plugging (G-7) into (G-6) shows that there exists a pair of $\mu_S > 0$ and $\mu_W > 0$ that solves equations (G-5) and (G-6) if and only if there exists $\mu_W > 0$ that solves

$$\frac{c'(\mu_W)}{a_W} = \frac{vm}{n(\theta(c')^{-1} \circ (a_S/a_W c'(\mu_W)) + (1-\theta)\mu_W)}. \quad (\text{G-8})$$

Suppose that challenge equilibria exist at $n = n'$. Then our above analysis implies that there exists $\mu^o > 0$ such that equation (G-8) holds for $\mu_W = \mu^o$ and $n = n'$; moreover, the condition $\mu_S/\mu_W < p_s^C(n)/p_o^C(n)$ is satisfied when $n = n'$, $\mu_W = \mu^o$, and $\mu_S = (c')^{-1} \circ (a_S/a_W c'(\mu^o))$.

Now consider any $n'' > n'$. We first argue that, if c' is geometrically concave, then when $n = n''$ and $\mu_W = (n'/n'')\mu^o$, the left hand side of (G-8) is less than its right hand side. To see this, note that

$$\begin{aligned} \frac{c'((n'/n'')\mu^o)}{a_W} &< \frac{c'(\mu^o)}{a_W} = \frac{vm}{n'(\theta(c')^{-1} \circ (a_S/a_W c'(\mu^o)) + (1-\theta)\mu^o)} \\ &= \frac{vm}{n'\mu^o \left(\theta \frac{(c')^{-1} \circ (a_S/a_W c'(\mu^o))}{\mu^o} + (1-\theta) \right)} \leq \frac{vm}{n'\mu^o \left(\theta \frac{(c')^{-1} \circ (a_S/a_W c'((n'/n'')\mu^o))}{(n'/n'')\mu^o} + (1-\theta) \right)} \\ &= \frac{vm}{n''(\theta(c')^{-1} \circ (a_S/a_W c'((n'/n'')\mu^o)) + (1-\theta)((n'/n'')\mu^o))}, \quad (\text{G-9}) \end{aligned}$$

where, in the first line, the inequality follows from the facts that c' is increasing and $n'/n'' \in (0, 1)$, and the equality follows from our above discussion that (G-8) holds for $n = n'$ and $\mu_W = \mu^o$. The inequality in the second line follows from Result G-1, the hypothesis that c' is geometrically concave, and the facts that $n'/n'' \in (0, 1)$ and $a_S/a_W > 1$.

Equation (G-9) implies that the left hand side of (G-8) is less than its right hand side when $n = n''$ and $\mu_W = (n'/n'')\mu^o$ if c' is geometrically concave. Because (G-8) holds for $n = n'$ and

$\mu_W = \mu^o$ and because $n'' > n'$, the left hand side of (G-8) is greater than its right hand side when $n = n''$ and $\mu_W = \mu^o$. Because c' is increasing, the left hand side of (G-8) is increasing in μ_W while its right hand side is decreasing in μ_W . Thus, given that both sides of (G-8) are continuous in μ_W , for any given $n = n'' > n'$, there must exist $\hat{\mu} \in ((n'/n'')\mu^o, \mu^o)$ that uniquely solves (G-8) when c' is geometrically concave.

Therefore, when challenge equilibria exist at n' , for any given $n = n'' > n'$, the pair, $\mu_W = \hat{\mu}$ and $\mu_S = (c')^{-1} \circ (a_S/a_W c'(\hat{\mu}))$, solves equations (G-5) and (G-6) if c' is geometrically concave. To show that challenge equilibria also exist at any $n'' > n'$, it suffices to show that this pair satisfies the challenge-equilibrium condition that $\mu_S/\mu_W < p_s^C(n'')/p_o^C(n'')$. We know from Lemma 3 that, with the capacity ratio, μ_S/μ_W , and the quota, m , fixed, if challenge equilibria exist at n' , challenge equilibria also exist at any $n'' > n'$. Thus, given that the capacity ratio at n' equals $(c')^{-1} \circ (a_S/a_W c'(\mu^o))/\mu^o$ and that challenge equilibria exist at n' , it must be the case that

$$(c')^{-1} \circ (a_S/a_W c'(\mu^o))/\mu^o < p_s^C(n'')/p_o^C(n''), \quad n'' > n'. \quad (\text{G-10})$$

Because $\hat{\mu} < \mu^o$, by Result G-1, when c' is geometrically concave,

$$(c')^{-1} \circ (a_S/a_W c'(\hat{\mu}))/\hat{\mu} \leq (c')^{-1} \circ (a_S/a_W c'(\mu^o))/\mu^o. \quad (\text{G-11})$$

Equations (G-10) and (G-11) imply that

$$(c')^{-1} \circ (a_S/a_W c'(\hat{\mu}))/\hat{\mu} < p_s^C(n'')/p_o^C(n'').$$

Thus, the pair, $\mu_W = \hat{\mu}$ and $\mu_S = (c')^{-1} \circ (a_S/a_W c'(\hat{\mu}))$, both solves equations (G-5) and (G-6) for any $n = n'' > n'$ and satisfies the challenge-equilibrium condition that $\mu_S/\mu_W < p_s^C(n'')/p_o^C(n'')$. Thus, if challenge equilibria exist at n' and if c' is geometrically concave, challenge equilibria also exist at any $n'' > n'$.

To argue that the capacity ratio, μ_S/μ_W , at any $n'' > n'$ is (smaller) no greater than at n' if c' is (strictly) geometrically concave, simply note that, by (G-7), in challenge equilibria, the capacity ratio satisfies

$$\mu_S/\mu_W = (c')^{-1} \circ (a_S/a_W c'(\mu_W))/\mu_W. \quad (\text{G-12})$$

By Result G-1, the right hand side of (G-12) is (increasing) nondecreasing in μ_W if c' is (strictly) geometrically concave. Because $\hat{\mu} < \mu^o$, meaning that μ_W is smaller at any $n'' > n'$ than at n' , it must be the case that the right hand side of (G-12) at any $n'' > n'$ is (smaller) no greater than at n' if c' is (strictly) geometrically concave. Thus, by (G-12), the capacity ratio, μ_S/μ_W , at any $n'' > n'$ is (smaller) no greater than at n' if c' is (strictly) geometrically concave. This completes the proof of part (i).

To establish part (ii), suppose that challenge equilibria exist at both n' and $n'' > n'$. Because, in challenge equilibria, μ_W solves (G-8), and also because the left hand side of (G-8) is increasing in μ_W while its right hand side is decreasing in μ_W , the value of μ_W that solves (G-8) when $n = n'' > n'$ must be smaller than the value of μ_W that solves (G-8) when $n = n'$. Because, in challenge equilibria, the capacity ratio satisfies (G-12), and also because, by Result G-1, the

right hand side of (G-12) is decreasing in μ_W if c' is strictly geometrically convex, by (G-12), the capacity ratio, μ_S/μ_W , is greater at $n'' > n'$ than at n' if c' is strictly geometrically convex. Part (ii) thus follows. \square

H. More than two types of contestants

In our model analyzed in the main text, the contestants are of two possible types. In such a setting, a symmetric equilibrium is one of two possible types, concession, in which weak contestants concede victory to strong contestants, and challenge, in which weak contestants have a positive probability of besting strong contestants. In contrast, if there are more than two possible types of contestants, there will be more than two possible equilibrium configurations. For example, if the contestants are of three possible types, high, medium, and low, then there can exist equilibrium configurations in which low and medium types both concede to high-type contestants while low-type contestants challenge medium-type contestants. Such a partial-challenge-partial-concession equilibrium configuration does not exist when contestant types are binary.

In this appendix, we give a discussion of how our results derived in the main text might change if there are an arbitrary finite number of contestant types. Proposition H-1 below shows that the qualitative nature of our contest-size result derived in Section 5 extends to this more general type-distribution case. Proposition H-2 shows that our contest-quota result derived in Section 4 also extends when contest size is sufficiently large and the designer gains only by selecting the strongest type. Finally, we use an example to show that, in contrast to the quota-inflation result derived in Section 4, when the designer gains also by selecting some intermediate types, selective standards can sometimes be optimal.

In what follows, we assume that there are a finite number of contestant types, $t = 1, \dots, T$. Each contestant of type t has capacity μ_t . Assume, without loss of generality, that $\mu_1 > \dots > \mu_T$. Thus, type 1 represents the strongest type and type T represents the weakest type. The probability that a contestant is of type t equals θ_t , where $\theta_t \in (0, 1)$ and $\sum_{t=1}^T \theta_t = 1$. We keep the rest of the assumptions made in Section 3.

Robustness of the contest-size result

Applying essentially the same argument in the proofs of Claim 1 and Lemma 1 shows that

Lemma H-1. In any symmetric equilibrium, the probability-of-winning function, P , is (a) continuous, (b) has an interval support, $[0, \hat{x}]$, where $\hat{x} < \infty$ is endogenous, and (c) is concave over its support. Moreover, for each type- $t \in \{1, \dots, T\}$ contestant, a fair gamble between x' and x'' , $x', x'' \in \text{Supp}_t$, is a best reply, and all performance/probability-of-winning pairs, $(x, P(x))$, such that $x \in \text{Supp}_t$ are collinear.

Lemma H-1, coupled with an argument used for deriving the two equilibrium configurations in Section 3.2, implies that, in any symmetric equilibrium, the probability-of-winning function, P , is piecewise linear over its interval support, $[0, \hat{x}]$. This result, coupled with the collinearity condition stated in Lemma H-1, allows to us to show that

Lemma H-2. In any symmetric equilibrium in which each contestant, regardless of her type, has a positive probability of besting the strongest type (i.e., type 1), the probability-of-winning function, P , is linear over its interval support, $[0, \hat{x}]$, and meets the origin.

Proof. By Lemma H-1, P is continuous and, hence, $P(0) = 0$. Suppose, by contradiction, that P is nonlinear over its interval support, $[0, \hat{x}]$. Then given the piecewise linearity of P , P has at least one kink point over the interior of its support. Let $x^o \in (0, \hat{x})$ be the highest kink point. The piecewise linearity of P then implies that P is linear over $[x^o, \hat{x}]$. Because P is kinked at x^o , the concavity of P given by Lemma H-1 implies that no pair $(x, P(x))$, $x < x^o$, can be collinear with the line segment of P over $[x^o, \hat{x}]$. Thus, by the collinearity condition given in Lemma H-1 and the fact that contestants place no point mass in any symmetric equilibrium (this fact is implied by the continuity of P), in any symmetric equilibrium, any type of contestant who places weight over $[x^o, \hat{x}]$ places no weight over $[0, x^o]$, and any type of contestant who places weight over $[0, x^o]$ places no weight over $[x^o, \hat{x}]$. Thus, all the contestant types can be divided into two groups, an “A” group placing weight only over $[x^o, \hat{x}]$ and a “B” group placing weight only over $[0, x^o]$. Clearly, the B group concedes victory to the A group. Note that the strongest type must belong to the A group. Otherwise, the performance of the strongest type would be stochastically dominated by the performance of all types in the A group, which is only possible if the strongest type’s mean performance, i.e., capacity, is less than the capacity of all types in the A group. But this is not possible because, by assumption, the strongest type has the greatest capacity. Given that the strongest type belongs to the A group, the B group concedes to the strongest type, contradicting the hypothesis that each contestant, regardless of her type, has a positive probability of besting the strongest type. The lemma thus follows. \square

Note that, if P is linear over its interval support, $[0, \hat{x}]$, and meets the origin, each type’s equilibrium probability of winning will be proportional to the type’s capacity. Thus, by Lemma H-2 and an argument similar to the one for equation (A-73), we obtain that

Corollary H-1. In any symmetric equilibrium in which each contestant, regardless of her type, has a positive probability of besting the strongest type (i.e., type 1), the probability that a selected contestant is of type t , is given by

$$\mathbb{P}[\text{Type } t | \text{Selected}] = \frac{\theta_t \mu_t}{\sum_{j=1}^T \theta_j \mu_j}.$$

To analyze the effect of contest size on designer welfare in settings with T contestant types, we adopt the following designer welfare function, u , where

$$u = \mathbb{E} \left[\sum_{t=1}^T \lambda_t \# \text{Type-}t \text{ selected contestants} \right], \quad \lambda_1 > 0 > \lambda_T \text{ \& } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_T. \quad (\text{H-1})$$

The welfare specification given by equation (H-1) is a simple extension of the asymmetric welfare specification, equation (15), we used in Section 6.1 to the case of T types of contestants. The condition $\lambda_1 > 0 > \lambda_T$ implies that the designer gains utiles by selecting the strongest type

(type 1) and loses utiles by selecting the weakest type (type T). The condition $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_T$ implies that the designer is always weakly better off by selecting stronger types.

The next proposition shows that the qualitative nature of our contest-size result, Theorem 2, derived in a two-type setting extends to a more general, T -type setting.

Proposition H-1. Suppose the selection quota, m , is fixed but the designer can vary contest size, n .

- i. There exists n^c such that, for all $n > n^c$, in any symmetric equilibrium, each contestant, regardless of her type, has a positive probability of besting the strongest type (type 1).
- ii. Moreover, for any $n' > n'' > n^c$, designer welfare at $n = n'$ equals designer welfare at $n = n''$.

Proof. (i): Consider any type t . Let p_t be the equilibrium probability of winning for contestants of type t . Because a contestant of type t always has the option of choosing a high-risk strategy that gives her a strongest type's performance distribution with probability μ_t/μ_1 and zero performance with probability $1 - (\mu_t/\mu_1)$, it must be the case that

$$p_t \geq (\mu_t/\mu_1)p_1 \geq (\mu_t/\mu_1)(m/n), \quad (\text{H-2})$$

where the last inequality follows from the fact that the equilibrium probability of winning for the strongest type is no less than the average winning probability, m/n .

If a contestant of type t concedes to the strongest type (type 1), then the contestant of type t can possibly win only if the number of type-1 rivals, which we denote by \tilde{S}_{n-1} , is less than the quota, m . Note that $\tilde{S}_{n-1} \sim \text{Binomial}(n-1, \theta_1)$. Thus, if \hat{p}_t denotes the supremum of a type- t contestant's probability of winning by conceding to the strongest type, it must be the case that

$$\hat{p}_t \leq \mathbb{P}[\tilde{S}_{n-1} < m], \quad (\text{H-3})$$

where $\tilde{S}_{n-1} \sim \text{Binomial}(n-1, \theta_1)$.

Note that, for any fixed $i \geq 0$,

$$0 \leq n \mathbb{P}[\tilde{S}_{n-1} = i] = n \binom{n-1}{i} \theta_1^i (1 - \theta_1)^{n-1-i} < c n^{i+1} (1 - \theta_1)^n, \quad (\text{H-4})$$

where $c > 0$ is a constant. Theorem 3.20(d) in Rudin (1976) shows that $\lim_{n \rightarrow \infty} n^a p^n = 0$ for all $a > 0$ and $p \in [0, 1)$. Thus, the last expression in (H-4) goes to 0 as $n \rightarrow \infty$. Thus, applying the Squeeze Theorem to (H-4) shows that $\lim_{n \rightarrow \infty} n \mathbb{P}[\tilde{S}_{n-1} = i] = 0$ for any fixed $i \geq 0$. Hence, for any fixed m ,

$$\lim_{n \rightarrow \infty} n \mathbb{P}[\tilde{S}_{n-1} < m] = \sum_{i=0}^{m-1} \left(\lim_{n \rightarrow \infty} n \mathbb{P}[\tilde{S}_{n-1} = i] \right) = 0. \quad (\text{H-5})$$

Equation (H-5) implies that, with $m > 0$ fixed, for n sufficiently large, $n \mathbb{P}[\tilde{S}_{n-1} < m] < (\mu_t/\mu_1)m$ and, hence, $\mathbb{P}[\tilde{S}_{n-1} < m] < (\mu_t/\mu_1)(m/n)$. This result, combined with equations (H-2) and (H-3), implies that, with m fixed, for n sufficiently large, $p_t > \hat{p}_t$. Thus, when n is sufficiently large, for any contestant, conceding to the strongest type is a strictly dominated strategy. Part (i) thus follows.

(ii): By part (i) and Corollary H-1, at any $n > n^c$, designer welfare, specified in (H-1), satisfies

$$u = m \left(\sum_{t=1}^T \lambda_t \frac{\theta_t \mu_t}{\sum_{j=1}^T \theta_j \mu_j} \right).$$

Part (ii) then follows from the fact that contest size, n , does not enter the above expression for u . \square

Robustness of the selection-quota result

The next proposition shows that, when the designer gains utiles only by selecting the strongest type (and loses utiles by selecting the other types), under-selection is never optimal for large contests.

Proposition H-2. Suppose that the designer's welfare function follows equation (H-1) with $\lambda_1 > 0 > \lambda_2 \geq \dots \geq \lambda_T$. Suppose that the designer can vary only the selection quota. If m_M^* represents the optimal quota with non-strategic contestants and m^* represents the optimal quota with strategic contestants, then when contest size, n , is sufficiently large, either $m^* \geq m_M^*$ or $m^* = 0$.

Proof. Because, by the law of large numbers, the fraction of contestants of the strongest type converges almost surely to θ_1 , the hypothesis that the designer gains only by selecting contestants of the strongest type implies that $\lim_{n \rightarrow \infty} m_M^*/n = \theta_1$. Because the equilibrium probability of winning for contestants of the strongest type must be no less than m/n , if the designer sets the quota at $m = m_M^*$, then in the limit (i.e., as $n \rightarrow \infty$), the strongest type's equilibrium probability of winning must be no less than θ_1 . In this case, by equation (H-2), in the limit, the equilibrium probability of winning for each contestant of type t must be no less than $\theta_1(\mu_t/\mu_1) > 0$. However, if a contestant concedes to the strongest type, then in the limit, her probability of winning will be zero when $m = m_M^*$. Thus, for n sufficiently large, if the designer sets the quota at $m = m_M^*$, each contestant, regardless of her type, will challenge the strongest type. Increasing competitiveness by shrinking the quota will only induce more risk taking. Thus, for n sufficiently large, if the designer sets the quota at $m \in \{1, \dots, m_M^*\}$, each contestant, regardless of her type, will challenge the strongest type. Then, by Corollary H-1, for n sufficiently large, designer welfare, specified in equation (H-1), satisfies

$$u(m) = m \left(\sum_{t=1}^T \lambda_t \frac{\theta_t \mu_t}{\sum_{j=1}^T \theta_j \mu_j} \right), \quad 1 \leq m \leq m_M^*. \quad (\text{H-6})$$

If the whole term inside the brackets in the above equation is negative, then for n sufficiently large, setting a zero quota, which gives the designer $u(0) = 0$, strictly dominates setting any $m \in \{1, \dots, m_M^*\}$. If this term is positive, then u given in equation (H-6) is increasing in m . In this case, for n sufficiently large, setting $m = m_M^*$ strictly dominates setting any $m < m_M^*$. If this term is zero, then $u(m) = 0$ for $m = 0, 1, \dots, m_M^*$. In this case, given that we have assumed

in the main text that, if the designer is indifferent between two quotas, he chooses the larger quota, the designer will choose m_M^* over $m < m_M^*$. Thus, in all of these cases, the designer does not choose any $m = 1, \dots, m_M^* - 1$. The proposition thus follows. \square

Optimality of selective standards: An example In what follows, we use a simple example to show that, when the designer gains utiles not only by selecting the strongest type but also by selecting some intermediate types, under-selection can sometimes be optimal. Suppose $T = 3$ (i.e., the contestants are of three possible types). For expositional convenience, we call these three types, high (H), medium (M), and low (L). Suppose that there are $n = 3$ contestants, and a contestant has an equal probability of being any of the three possible types. The designer sets the quota, m , to maximize his welfare function, u , given as follows:

$$u = \mathbb{E}[\#H \text{ selected contestants} + (1/10)\#M \text{ selected contestants} - \#L \text{ selected contestants}]. \quad (\text{H-7})$$

Thus, the designer gains one utile by selecting a high-type contestant, gains one tenth of a utile by selecting a medium-type contestant, and loses one utile by selecting a low-type contestant.

If the contestants are non-strategic, the high type will have the highest priority for winning a place, followed by the medium type. In this case, it can be shown that $u(0) = 0$, $u(1) = 187/270$, $u(2) = 7/9$, and $u(3) = 1/10$ (we omit the derivation details). Thus, if the contestants are non-strategic, the optimal quota will be 2.

Now consider strategic contestants. Suppose that the capacity for the high, medium, and low type, denoted by μ_H , μ_M , and μ_L , respectively, is given by $\mu_H = 4$, $\mu_M = 1$, and $\mu_L = 1 - \varepsilon$. We consider the case where $\varepsilon > 0$ is infinitesimal. We approximate contestant behavior by their behavior in the limiting case where $\varepsilon = 0$. In this limiting case, the medium and the low types have the same capacity and thus, we can apply our results derived in Section 3.2 to characterize equilibria, treating the high type as the “strong type” and the medium and the low types as the “weak type;” the probability that a contestant is of the “strong type” (“weak type”) equals $1/3$ ($2/3$) and the capacity for the “strong type” (“weak type”) equals 4 (1).

Thus, by equations (6) and (8), in the limiting case, if $m = 1$, then $p_o^G = 1/6 > 4/27 = p_o^C$, whereas if $m = 2$, then $p_o^G = 1/3 < 14/27 = p_o^C$. Hence, by Lemma 2, in the limiting case, if $m = 1$, there exist equilibria in which contestants of the “weak type” (i.e., the medium and the low types) challenge contestants of the “strong type” (i.e., the high type), whereas if $m = 2$, there exist equilibria in which contestants of the “weak type” concede to contestants of the “strong type.” Lemma 2 further implies that, in the limiting case, if $m = 1$, the equilibrium probability of winning for contestants of the medium and the low types both equal $1/6$ and the equilibrium probability of winning for contestants of the high type equals $2/3$, whereas if $m = 2$, the equilibrium probability of winning for contestants of the medium and the low types both equal $14/27$ and the equilibrium probability of winning for contestants of the high type equals $26/27$. Thus, by Bayes’ rule, in the limiting case, if $m = 1$, then the selected contestant will be of the high type with probability $2/3$, of the medium type with probability $1/6$, and of

the low type with probability $1/6$, implying that the designer's welfare, given in equation (H-7), satisfies $u(1) = (2/3) + (1/10) \times (1/6) - (1/6) = 31/60$. Similarly, in the limiting case, if $m = 2$, then a selected contestant will be of the high type with probability $13/27$, of the medium type with probability $7/27$, and of the low type with probability $7/27$, implying that $u(2) = 2 \times [(13/27) + (1/10) \times (7/27) - (7/27)] = 67/135$. Thus, $u(1) = 31/60 > 67/135 = u(2)$. Also note that $u(0) = 0$ and $u(3) = 1/10$. Therefore, in the limiting case, when the contestants are strategic, the optimal quota will be 1, which is smaller than the optimal quota in the absence of risk taking. This implies that, when the capacity difference between the medium type and the low type is sufficiently small (i.e., when ε is sufficiently close to 0), under-selection is optimal.

Intuitively, when contestants are non-strategic, the medium type will have priority over the low type for selection. In this case, the designer's utility gain from selecting medium-type contestants encourages the designer to set a large quota. However, in the example, when contestants are strategic, given the tiny capacity difference between the medium type and the low type, a low-type contestant will have almost the same probability of being selected as a medium-type contestant. Because, in the example, the designer's utility gain from selecting a medium-type contestant is much smaller than the utility loss from selecting a low-type contestant, setting a large quota, which is optimal when contestants are non-strategic, will no longer be optimal when contestants are strategic. Thus, when contestants are strategic, either a small quota or a zero quota is optimal. Because, in the example, the high type has much higher capacity than the medium and the low types and each contestant has a decent probability of being of the high type, the designer has a high probability of selecting a high-type contestant under a small quota. Thus, in the example, it is optimal for the designer to set a small quota rather than a zero quota. Consequently, selective standards are optimal in this example.

I. No aggregate uncertainty with respect to contestant types

Our quota-inflation result, Theorem 1, is established in a setting where each contestant's type is an i.i.d. draw from the same type distribution. In such a setting, the total number of strong contestants in the contestant pool is random. In this appendix, we show that the qualitative nature of our quota-inflation result extends to cases where the total number of strong contestants is fixed and is common knowledge.

To this end, throughout this appendix, we assume that the total number of strong contestants is a constant, s , $0 < s < n$, and is common knowledge. Under this assumption, the total number of weak contestants is $n - s$. We assume that, while the designer knows the total number of strong contestants and the total number of weak contestants, he does not know the type of any given contestant and, from his point of view, all of the contestants are ex ante identical. We keep the rest of the assumptions of our baseline model made in Section 3.

Given that the designer knows the number of strong contestants, s , when the contestants are non-strategic, it is obvious that the optimal quota equals s . In what follows, we aim to show that, when the contestants are strategic, the optima quota is either no less than s or equal to 0 (i.e., risk taking inclines the designer either to inflate the quota or not to run a contest).

We still base our discussion on symmetric equilibria, where contestants of the same type play the same strategy, i.e., each type- t contestant chooses performance distribution F_t with support Supp_t , $t \in \{S, W\}$. Let $\bar{x}_t \equiv \max \text{Supp}_t$ and $\underline{x}_t \equiv \min \text{Supp}_t$ be the upper bound and the lower bound of Supp_t , respectively.

A contestant's probability-of-winning function maps the contestant's realized performance, x , to her probability of being selected and is determined endogenously by her rivals' strategies. Note that, in this modified model, contestants of different types face different distributions of rival types: each strong contestant faces $s - 1$ strong rivals and $n - s$ weak rivals, whereas each weak contestant faces s strong rivals and $n - s - 1$ weak rivals. Consequently, the probability-of-winning function faced by strong contestants is different from the probability-of-winning function faced by weak contestants. Let $P_t : \mathbb{R}_+ \rightarrow [0, 1]$ be the probability-of-winning function faced by each type- $t \in \{S, W\}$ contestant in a symmetric equilibrium. Let p_t be a type- $t \in \{S, W\}$ contestant's equilibrium probability of winning.

To simplify our discussion, in the subsequent analysis, we assume that

Assumption I-1. The total number of strong contestants, s , is smaller than the total number of weak contestants, $n - s$, i.e., $s < n/2$.

Assumption I-1 will be used only in the proof of Lemma I-1 below to show that, if the selection quota, m , equals the total number of strong contestants, s , then in any symmetric equilibrium, a weak contestant has zero probability of winning if her performance equals the lower bound of the support of a weak type's performance distribution (i.e., $P_W(\underline{x}_W) = 0$) and no weak contestant places point mass on the upper bound of the support of a weak type's

performance distribution.⁹

Our key result is given by Proposition I-1, presented at the end of this appendix. To obtain this proposition, we require a couple of lemmas. The first one is the following:

Lemma I-1. Suppose $m = s$. Then, in any symmetric equilibrium, $\bar{x}_W \geq \mu_S$, $P_W(\underline{x}_W) = 0$, and no weak contestant places point mass on \bar{x}_W .

Proof. First, note that, because weak contestants always have the option of challenging strong contestants by adopting a high-risk strategy, in any symmetric equilibrium, weak contestants must have a positive probability of winning. When $m = s$, for the $n - s$ weak contestants to have a positive probability of winning, the s strong contestants must have a positive probability of losing. However, if $\bar{x}_W < \mu_S$, then given $m = s$, all of the s strong contestants would ensure winning by playing safe, i.e., by choosing a degenerate distribution at their capacity, μ_S , a contradiction. The result that $\bar{x}_W \geq \mu_S$ thus follows.

Next, note that, because $\underline{x}_W \leq \mu_W < \mu_S \leq \bar{x}_S$, a weak contestant cannot ensure outperforming a strong contestant if the weak contestant's performance equals \underline{x}_W . Thus, given that, when $m = s$, a weak contestant cannot ensure winning if she cannot ensure outperforming any strong contestant, it must be the case that $P_W(\underline{x}_W) < 1$. Assumption I-1 implies that the number of weak contestants is at least 2 and, for a weak contestant to win at $m = s$, she has to outperform at least some of her weak rivals. By definition, \underline{x}_W is the lower bound of Supp_W . Thus, if, by contradiction, $P_W(\underline{x}_W) > 0$, it would be the case that weak contestants place point mass on \underline{x}_W . In this case, given $P_W(\underline{x}_W) < 1$, \underline{x}_W would be a discontinuity point for P_W , and $P_W(\underline{x}_W +) > P_W(\underline{x}_W)$. A weak contestant would then be strictly better off adopting a distribution that places mass on $\underline{x}_W + \varepsilon$, for ε sufficiently small, rather than on \underline{x}_W (a detailed argument is similar to the one used for showing the continuity of P in our baseline model; see the verification of Claim 1 in Online Appendix A). Thus, by placing mass on \underline{x}_W , weak contestants could not be optimizing, a contradiction. This contradiction implies that $P_W(\underline{x}_W) = 0$.

Finally, we show that no weak contestant places point mass on \bar{x}_W . Because, as has been established, $\bar{x}_W \geq \mu_S$, and because $\mu_S > \mu_W$, the capacity constraints imply that $F_S(\bar{x}_W) > 0$ and $F_W(\bar{x}_W) > 0$. Thus, $P_W(\bar{x}_W) > 0$. If, by contradiction, weak contestants placed point mass on \bar{x}_W , there would exist a positive probability that all of the $n - s$ weak contestants tie at \bar{x}_W . In this case, given that ties are broken by fair randomization, each weak contestant would have a positive probability of being bested by all of her $n - s - 1$ weak rivals. Because, when $m = s < n/2$, a weak contestant cannot win if she bests none of her weak rivals, in the positive-probability event where all of the $n - s$ weak contestants tie at \bar{x}_W , no weak contestant could ensure winning. Thus, if weak contestants placed point mass on \bar{x}_W , we would have $P_W(\bar{x}_W) < 1$. However, given that $0 < P_W(\bar{x}_W) < 1$, by an argument similar to the one used above for showing that no weak contestant places point mass on \underline{x}_W , we can show that, by

⁹We conjecture that Lemma I-1 holds even in the absence of Assumption I-1.

placing mass on \bar{x}_W , weak contestants could not be optimizing, a contradiction. The result that no weak contestant places point mass on \bar{x}_W thus follows. \square

Lemma I-1 allows us to establish the next result, which shows that, when $m = s$, in any symmetric equilibrium, the probability-of-winning ratio between strong and weak contestants must be no less than their capacity ratio.

Lemma I-2. Suppose $m = s$. Then, in any symmetric equilibrium, $p_S/p_W \geq \mu_S/\mu_W$.

Proof. By Lemma I-1, when $m = s$, $P_W(\underline{x}_W) = 0$. In this case, $P_W(\underline{x}_W+) = P_W(\underline{x}_W) = 0$. Also by Lemma I-1, when $m = s$, no weak contestant places point mass on \bar{x}_W . Thus, given that, by definition, \bar{x}_W is the upper bound of Supp_W , there must exist $x' < \bar{x}_W$ such that weak contestants randomize continuously over $[x', \bar{x}_W]$. Hence, given that $P_W(\underline{x}_W+) = P_W(\underline{x}_W) = 0$, the definitions of \underline{x}_W and \bar{x}_W imply that a fair gamble between \underline{x}_W and \bar{x}_W must be a best reply for a weak contestant. Because a weak contestant's mean performance equals μ_W , if she chooses this fair gamble, her performance will equal \bar{x}_W with probability $(\mu_W - \underline{x}_W)/(\bar{x}_W - \underline{x}_W)$ and equal \underline{x}_W with the complementary probability. Because, by playing a best reply, a weak contestant will obtain her equilibrium probability of winning, and also because $P_W(\underline{x}_W) = 0$, it must be the case that, when $m = s$,

$$p_W = \frac{\mu_W - \underline{x}_W}{\bar{x}_W - \underline{x}_W} P_W(\bar{x}_W). \quad (\text{I-1})$$

Now consider a strong contestant's probability of winning by choosing a fair gamble between \underline{x}_W and \bar{x}_W . Note that $\underline{x}_W \leq \mu_W < \mu_S$ and, by Lemma I-1, $\bar{x}_W \geq \mu_S$. Thus, $\underline{x}_W < \mu_S \leq \bar{x}_W$. Hence, there exists a fair gamble between \underline{x}_W and \bar{x}_W for a strong contestant.¹⁰ Because a strong contestant's mean performance equals μ_S , if she chooses this fair gamble, her performance will equal \bar{x}_W with probability $(\mu_S - \underline{x}_W)/(\bar{x}_W - \underline{x}_W)$ and equal \underline{x}_W with the complementary probability. A strong contestant's equilibrium probability of winning must be no less than her probability of winning by playing such a fair gamble. Thus, when $m = s$,

$$p_S \geq \frac{\mu_S - \underline{x}_W}{\bar{x}_W - \underline{x}_W} P_S(\bar{x}_W) + \left(1 - \frac{\mu_S - \underline{x}_W}{\bar{x}_W - \underline{x}_W}\right) P_S(\underline{x}_W) \geq \frac{\mu_S - \underline{x}_W}{\bar{x}_W - \underline{x}_W} P_S(\bar{x}_W). \quad (\text{I-2})$$

Because, by Lemma I-1, when $m = s$, no weak contestant places point mass on \bar{x}_W , and also because \bar{x}_W is defined as the upper bound of Supp_W , it must be that a contestant outperforms all of her weak rivals with certainty if her performance equals \bar{x}_W . Let q be a contestant's probability of besting any given strong rival by having performance equal to \bar{x}_W . Because a strong contestant faces $s - 1$ strong rivals and $n - s$ weak rivals, if her performance equals \bar{x}_W , she will best $n - s$ rivals with certainty and best $s - 1$ rivals each with probability q . Because a weak contestant faces s strong rivals and $n - s - 1$ weak rivals, if her performance equals \bar{x}_W , she will best $n - s - 1$ rivals with certainty and best s rivals each with probability q . Thus,

$$P_S(\bar{x}_W) \geq P_W(\bar{x}_W). \quad (\text{I-3})$$

¹⁰If $\bar{x}_W = \mu_S$, then such a fair gamble places all the weight on \bar{x}_W .

Equation (I-3), combined with (I-1) and (I-2), implies that

$$\frac{p_S}{p_W} \geq \frac{\mu_S - \underline{x}_W}{\mu_W - \underline{x}_W} \geq \frac{\mu_S}{\mu_W}, \quad (\text{I-4})$$

where the last inequality follows from the fact that $\underline{x}_W \mapsto (\mu_S - \underline{x}_W)/(\mu_W - \underline{x}_W)$ is increasing. The lemma then follows immediately from (I-4). \square

The next lemma contrasts with Lemma I-2. It shows that, when $0 < m < s$, in any symmetric equilibrium, the probability-of-winning ratio between strong and weak contestants must be no greater than their capacity ratio.

Lemma I-3. Suppose $0 < m < s$. Then, in any symmetric equilibrium, $p_S/p_W \leq \mu_S/\mu_W$.

Proof. When $0 < m < s$, the number of strong contestants is at least 2, and for a strong contestant to win, she has to outperform at least some of her strong rivals. This result, coupled with an argument similar to the one used in the proof of Lemma I-1 for showing that $P_W(\underline{x}_W) = 0$ when $m = s$, implies that, when $0 < m < s$,

$$P_S(\underline{x}_S) = 0. \quad (\text{I-5})$$

Moreover, by an argument similar to the one used in the proof of Lemma I-1 for showing that, when $m = s$, no weak contestant places point mass on \bar{x}_W , it must be the case that, when $0 < m < s$, no strong contestant places point mass on \bar{x}_S . Then, by an argument analogous to the one used in the proof of Lemma I-2 for showing that, when $m = s$, a fair gamble between \underline{x}_W and \bar{x}_W is a best reply for a weak contestant, it must be the case that, when $0 < m < s$, a fair gamble between \underline{x}_S and \bar{x}_S is a best reply for a strong contestant. Because a strong contestant's mean performance equals μ_S , if she chooses this fair gamble, her performance will equal \bar{x}_S with probability $(\mu_S - \underline{x}_S)/(\bar{x}_S - \underline{x}_S)$ and equal \underline{x}_S with the complementary probability. Because, by playing a best reply, a strong contestant will obtain her equilibrium probability of winning, and also because, by (I-5), $P_S(\underline{x}_S) = 0$, it must be the case that, when $0 < m < s$,

$$p_S = \frac{\mu_S - \underline{x}_S}{\bar{x}_S - \underline{x}_S} P_S(\bar{x}_S). \quad (\text{I-6})$$

Because $\bar{x}_S \geq \mu_S$ and $\mu_S > \mu_W$, $\bar{x}_S > \mu_W$. Thus, there exists a fair gamble between 0 and \bar{x}_S for a weak contestant. Because a weak contestant's mean performance equals μ_W , if she chooses this fair gamble, her performance will equal \bar{x}_S with probability μ_W/\bar{x}_S and equal 0 with the complementary probability. A weak contestant's equilibrium probability of winning must be no less than her probability of winning by playing such a fair gamble. Thus, when $0 < m < s$,

$$p_W \geq \frac{\mu_W}{\bar{x}_S} P_W(\bar{x}_S). \quad (\text{I-7})$$

By an argument completely analogous to the one used in the proof of Lemma I-2 for establishing equation (I-3), we can show that

$$P_W(\bar{x}_S) \geq P_S(\bar{x}_S). \quad (\text{I-8})$$

Equation (I-8), combined with (I-6) and (I-7), implies that

$$\frac{p_S}{p_W} \leq \left(\frac{\mu_S - \underline{x}_S}{\bar{x}_S - \underline{x}_S} \right) \frac{\bar{x}_S}{\mu_W} \leq \left(\frac{\mu_S}{\bar{x}_S} \right) \frac{\bar{x}_S}{\mu_W} = \frac{\mu_S}{\mu_W}, \quad (\text{I-9})$$

where the second inequality follows from the fact that, given $\mu_S \leq \bar{x}_S$, $\underline{x}_S \mapsto (\mu_S - \underline{x}_S)/(\bar{x}_S - \underline{x}_S)$ is nonincreasing. The lemma then follows immediately from (I-9). \square

The above lemmas allow us to establish our key result, given as follows:¹¹

Proposition I-1. If m^* represents the optimal quota when contestants are strategic, then either $m^* \geq s$ or $m^* = 0$.

Proof. The proposition holds trivially for $s = 1$; henceforth, suppose $s > 1$. Suppose that designer welfare, u , is given by the symmetric specification, equation (10). The case in which designer welfare is given by the asymmetric specification, (15), follows from a similar argument. Note that

$$u(m) = s p_S - (n - s) p_W. \quad (\text{I-10})$$

Because the total number of selected contestants must equal the quota, m ,

$$s p_S + (n - s) p_W = m. \quad (\text{I-11})$$

Equations (I-10) and (I-11) imply that

$$u(m) = m - 2(n - s) p_W. \quad (\text{I-12})$$

Equation (I-11), combined with Lemmas I-2 and I-3, implies that

$$\begin{aligned} p_W &\geq \frac{m}{n - s + s(\mu_S/\mu_W)} \quad \text{if } 0 < m < s, \\ \& \quad p_W &\leq \frac{m}{n - s + s(\mu_S/\mu_W)} \quad \text{if } m = s. \end{aligned}$$

Thus, by (I-12),

$$u(m) \leq m \left(1 - \frac{2(n - s)}{n - s + s(\mu_S/\mu_W)} \right) \quad \text{if } 0 < m < s, \quad (\text{I-13})$$

$$\& \quad u(m) \geq m \left(1 - \frac{2(n - s)}{n - s + s(\mu_S/\mu_W)} \right) \quad \text{if } m = s. \quad (\text{I-14})$$

If $1 - \frac{2(n - s)}{n - s + s(\mu_S/\mu_W)} < 0$, by (I-13), $u(m) < 0$ for all $0 < m < s$. In this case, $u(0) = 0 > u(m)$ for all $0 < m < s$ and, hence, choosing any $m = 1, \dots, s - 1$ is strictly dominated by choosing $m = 0$.

If $1 - \frac{2(n - s)}{n - s + s(\mu_S/\mu_W)} = 0$, by (I-13) and (I-14), $u(m) \leq 0 \leq u(s)$ for all $0 < m < s$. In this case, choosing $m = s$ weakly dominates choosing any $m = 1, \dots, s - 1$. Because, by assumption, when the designer is indifferent between two quotas, he chooses the larger quota, the designer will choose $m = s$ over $m = 1, \dots, s - 1$.

¹¹Construction of a symmetric equilibrium is available from the authors upon request.

If $1 - \frac{2(n-s)}{n-s+s(\mu_S/\mu_W)} > 0$, then by (I-13) and (I-14),

$$u(s) \geq s \left(1 - \frac{2(n-s)}{n-s+s(\mu_S/\mu_W)} \right) > m \left(1 - \frac{2(n-s)}{n-s+s(\mu_S/\mu_W)} \right) \geq u(m), \quad \text{for } 0 < m < s.$$

In this case, choosing $m = s$ strictly dominates choosing any $m = 1, \dots, s-1$.

Therefore, a meritocratic designer will never choose any $m = 1, \dots, s-1$. The proposition thus follows. \square

References for Appendices

Rudin, Walter, *Principles of Mathematical Analysis*, McGraw-Hill, 1976.