On the Optimal Design of Biased Contests*

Qiang Fu† Zenan Wu‡

February 21, 2020

Abstract

This paper explores the optimal design of biased contests. A designer imposes an identity-dependent treatment on contestants, which varies the balance of the playing field. A generalized lottery contest typically yields no closed-form equilibrium solutions, which nullifies the usual implicit programming approach to optimal contest design and limits analysis to restricted settings. We propose an alternative approach that allows us to circumvent this difficulty and characterize the optimum in a general setting under a wide array of objective functions without solving for the equilibrium explicitly. Our technique applies to a broad array of contest design problems, and the analysis it enables generates novel insights into incentive provision in contests and their optimal design. For instance, we demonstrate that the conventional wisdom of leveling the playing field, which is obtained in limited settings in previous studies, does not generally hold.

Keywords: Contest Design; Optimal Biases; Tullock Contest.

JEL Classification Codes: C72; D72.

*We are grateful to Coeditor Dilip Mookherjee and four anonymous reviewers for constructive and insightful comments on earlier versions of this paper. We thank Kyung Hwan Baik, Stefano Barbieri, Ulrich Doraszelski, Mikhail Drugov, Hanming Fang, Jörg Franke, Alex Gershkov, Simon Grant, Eun Jeong Heo, Ju Hu, Soo Jin Kim, Kai Konrad, Takashi Kunimoto, Wolfgang Leininger, Jianpei Li, Bin Liu, Qingmin Liu, Tore Nilssen, Paul Schweinzer, Zheng Song, Xi Weng, Jingyi Xue, Lixin Ye, Jun Yu, and Junjie Zhou, and seminar/conference participants at the Chinese University of Hong Kong (Shenzhen), Singapore Management University, Sun Yat-sen University, Sungkyunkwan University, University of International Business and Economics, the 4th Annual Conference on Contests: Theory and Evidence, the 2018 International Conference on Economic Theory and Applications (Chengdu), and the CUHK-ShanghaiTech Forum on Chinese Economy for helpful discussions, suggestions, and comments. We thank Shanglyu Deng and Yuxuan Zhu for excellent research assistance. Wu thanks the National Natural Science Foundation of China (No. 71803003) and the seed fund of the School of Economics, Peking University, for financial support. Any errors are our own.

†Department of Strategy and Policy, National University of Singapore, 15 Kent Ridge Drive, Singapore, 119245. Email: bizfq@nus.edu.sg

‡School of Economics, Peking University, Beijing, China, 100871. Email: zenan@pku.edu.cn

1 Introduction

Contests are widely administered in practice to mobilize productive effort. For instance, workers strive to be promoted to higher rungs on hierarchical ladders inside a firm (see, for instance, Rosen, 1986). Governments, firms, and even wealthy individuals sponsor innovation contests to promote research efforts (see Che and Gale, 2003). In a contest, contenders expend costly effort to vie for limited prizes and are rewarded based on their relative performance instead of absolute output metrics.

The ubiquity of contest-like competitive activities has triggered broad interest in their strategic substance and the optimal design of competitive schemes that spur incentive provision.[1] This paper explores a classic question: How should a designer bias the competition to boost the performance of a contest? Contestants' behaviors sensitively depend on their relative competitiveness, which can often be determined endogenously by the choice of contest rules. A designer can impose identity-dependent preferential treatments on contestants—tailored to their individual characteristics—to vary contestants' relative standing. Consider, for instance, government policies that favor small and medium-sized enterprises (SMEs) in public procurement to support local entrepreneurship (Che and Gale, 2003; Epstein, Mealem, and Nitzan, 2011) and colleges that allocate bonus points to minority applicants (Fu, 2006; Franke, 2012).

The literature broadly embraces the notion that a more level playing field fuels competition.[2,3] The conventional wisdom, however, is obtained in restricted settings—e.g., two players, stylized contest technologies, and limited objective functions—due to technical challenges. This paper develops a novel optimization approach that allows us to circumvent the analytical difficulty and identify the key properties of the optimum in a general context. The analysis yields novel implications that illuminate the nature of incentive provision in contests and refute the conventional wisdom.

Nature of the Generalized Optimization Problem The conventional wisdom of leveling the playing field is underpinned primarily by the rationale that favoring the underdog boosts his incentive, which further deters the favorite from slacking off. This logic, however, rests on contestants' nonmonotone best responses in bilateral strategic relation (Lazear and Rosen, 1981; Dixit, 1987). Involving more than two players fundamentally alters the nature of the strategic interaction in a contest and its optimal design.

[1] See Fu and Wu (2019a) for a recent survey of theoretical studies of contests.

[2] See the recent survey of Chowdhury, Esteve-González, and Mukherjee (2019) on biased contests.

[3] Two notable exceptions are provided by Fu, Lu, and Lu (2012) and Drugov and Ryvkin (2017). The former show that a performance-maximizing administrator may allocate more productive resources to an ex ante stronger firm. The latter show that it can be optimal to bias an otherwise symmetric contest. Both studies focus on two-player settings.

First, setting optimal identity-dependent preferential treatments in a two-player setting is a unidimensional problem, because favoring one equivalently handicaps the other. With more than two contestants, the strategic interactions are no longer reciprocal or direct. Contestants are reflexively entangled, which expands the channels through which a treatment could manipulate their behavior.

Imagine a contest with three players indexed by 1, 2, and 3. Suppose that a favorable bias is imposed on player 3. This directly boosts his own incentive, which compels the other two to respond. The favor given to player 3 also affects the strategic interaction between players 1 and 2: Player 1's response to the more competitive player 3 forces player 2 to adjust his behavior, and vice versa. This compounds the incentive effect of the bias on player 3; its overall effect must sum up contestants' responses over all of the links.

Second, a two-player setting narrows the scope of the optimal biased contest design problem. With more than two contestants, setting biases not only manipulates the balance of the playing field, but also selects preferred contestants: Handicapping a player can force him to exit, which is possible only if at least three contenders are present.

The conventional wisdom—which is obtained from restricted settings—deserves to be examined more generally. However, the analysis entails substantial complications. Optimal contest design results in a mathematical program with equilibrium constraints (MPEC) and typically requires an implicit programming approach. One has to solve for the equilibrium bidding strategies for any given parameterized contest rule, insert the solution into the objective function, and search for the optimal rule (e.g., Franke, Kanzow, Leininger, and Schwartz, 2013). The approach loses its bite in an asymmetric n-player contest, as in general it yields no closed-form equilibrium solution.

We propose an alternative optimization approach that allows us to characterize the optimum without solving explicitly for the equilibrium. Next, we provide a snapshot of the approach and its underlying logic.

Optimization Approach We adopt the framework of generalized lottery contests to model a noisy winner-take-all contest in which a higher effort does not ensure a win. Suppose that the contest involves $n \geq 2$ players who differ in their prize valuations. For a given effort profile $\boldsymbol{x} \equiv (x_1, \ldots, x_n)$, one wins with a probability

$$p_i(\boldsymbol{x}) = \frac{f_i(x_i)}{\sum_{j=1}^{n} f_j(x_j)},$$

where $f_i(\cdot)$ maps one's effort outlays onto his effective output and is conventionally called the *impact function* of contestant $i \in \{1, \ldots, n\}$. We focus on the two most popularly adopted instruments for identity-dependent preferential treatments in the literature: The impact function takes the form

$$f_i(x_i; \alpha_i, \beta_i) = \alpha_i \cdot h(x_i) + \beta_i,$$

where α_i is a *multiplicative bias* and β_i is *an additive headstart.* The designer imposes a contest rule $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \equiv \big((\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n)\big)$, with $\alpha_i, \beta_i \geq 0$, which depicts how each contestant is favored or handicapped vis-à-vis his opponents.

Despite the lack of a closed-form solution, a unique equilibrium exists under mild regularity conditions. The equilibrium condition alludes to a correspondence, which provides a system of equations; each equation expresses an individual's equilibrium effort as a function of his own *equilibrium winning odds* and *prize valuation.* The correspondence thus literally disaggregates the strategic interaction between contestants into a series of individual decision problems. The contest rule $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ does not appear in the equation and is encapsulated in each contestant's equilibrium winning probability. The design objective can be rewritten accordingly as a function of equilibrium winning probability distribution. Instead of optimizing over the choice of contest rule, we let the designer directly assign winning probabilities among contestants to maximize the reformulated objective function, which reduces the optimization problem to a simple programming that allocates probability mass among contestants based on their prize valuations. Finally, we demonstrate that any winning probability distribution can be induced by a contest rule in equilibrium, which closes the loop.

Implications and Applications In this paper, we set up a general objective function that addresses a wide spectrum of concerns in contest design. Our analysis yields rich implications that reveal general properties of optimal biased contests.

First, we show that allowing for headstarts $\boldsymbol{\beta}$—in addition to the freedom to set biases $\boldsymbol{\alpha}$—cannot further improve the performance of the contest. It is thus without loss of generality to focus solely on the optimal choice of biases $\boldsymbol{\alpha}$.

Second, we establish a general exclusion principle. The literature has debated whether certain players should be excluded from the competition (e.g., Baye, Kovenock, and de Vries, 1993; and Fang, 2002). In contrast to previous studies that allow for outright exclusion, we consider implicit exclusion by setting biases. Under mild conditions, we show that the optimal exclusion is monotone in the sense that exclusion always starts from the the weakest.

Third, we apply our approach to the classical effort-maximizing problem. To maximize total effort, the optimum must involve at least three active contestants whenever possible. A two-player contest is thus suboptimal and a knife-edge case. Further, the optimum precludes a "superstar," in that an individual contestant's winning odds must fall below 1/2. We then proceed to the maximization of the expected winner's effort and show that the optimum keeps only the two top-ranked contestants active.

Fourth, our approach allows us to reexamine the conventional wisdom of leveling the playing field. The literature has centered on two fundamental questions: (i) Should contestants' winning odds be equalized (i.e., leveling the playing field in terms of ex post equilibrium outcomes)? (ii) Should the contest rule favor weaker contestants vis-à-vis their stronger opponents (i.e., leveling the playing field in terms of ex ante contest rules)? Our analysis overturns the conventional wisdom. We show that equalized winning odds are an artifact of bilateral competitions. With three or more contestants, the strongest player may turn out to be the least likely winner; contestants' equilibrium winning probabilities can even be nonmonotone with respect to the rankings of their prize valuations.[4] Further, we demonstrate that the contest rule may even upset the balance of the contest by favoring stronger contestants when more than two contestants are involved; the optimal biases can be nonmonotone, in the sense that a middle-ranked contestant is the most privileged.

The rest of the paper proceeds as follows. Section 2 describes the contest model and the optimization problem. Section 3 develops our optimization approach and characterizes the optimal contests. Section 4 reexamines the conventional wisdom of leveling the playing field, and Section 5 concludes. Appendix A lays out the microfoundations of the underlying contest model. Appendix B collects proofs that are not provided in the main text.

2 Setup and Preliminaries

In this section, we present the fundamentals of the underlying contest game.

2.1 Generalized Lottery Contests

There are $n \geq 2$ risk-neutral contestants competing for a prize. The prize bears a value $v_i > 0$ for each contestant $i \in \mathcal{N} \equiv \{1, \ldots, n\}$, with $v_1 \geq \ldots \geq v_n > 0$, which is common knowledge. A contestant's prize valuation measures his strength, as a higher valuation motivates effort. Contestants simultaneously submit their effort entries $x_i \geq 0$ to vie for the prize, which incur a cost of $c(x_i)$.

We consider a generalized lottery contest with a ratio-form contest success function: For a given effort profile $\boldsymbol{x} \equiv (x_{1,}, \ldots, x_n)$, a contestant i wins with a probability

$$p_i(\boldsymbol{x}) = \begin{cases} \dfrac{f_i(x_i)}{\sum_{j=1}^n f_j(x_j)} & \text{if} \quad \sum_{j=1}^n f_j(x_j) > 0, \\ \dfrac{1}{n} & \text{if} \quad \sum_{j=1}^n f_j(x_j) = 0, \end{cases} \tag{1}$$

[4] In a standard lottery contest with $h(x_i) = x_i$, Franke, Kanzow, Leininger, and Schwartz (2013) show in a numerical example that the optimal biased contest rule favors ex ante weaker contestants but does not fully level the playing field, in the sense that an ex ante stronger contestant wins with a larger probability.

where the function $f_i(\cdot)$, labeled the impact function in the contest literature, converts one's effort into his effective output and satisfies $f_i(x_i) \geq 0$ for all $x_i \geq 0$. A contestant $i \in \mathcal{N}$ is excluded from the contest if $f_i(x_i) = 0$ for all $x_i \geq 0$. In the extreme case in which only one contestant has an increasing impact function, while the others' is a zero constant, we assume that he wins automatically.[5] Appendix A presents two rationales for the model's microeconomic underpinning: (i) a noisy-ranking approach adapted from the discrete-choice model (Clark and Riis, 1996; Jia, 2008); and (ii) a research tournament analogy (Loury, 1979; Dasgupta and Stiglitz, 1980; Fullerton and McAfee, 1999; Baye and Hoppe, 2003).

Given $\boldsymbol{x} \equiv (x_1, \ldots, x_n)$ and (1), contestant i's expected payoff can be written as

$$\pi_i(\boldsymbol{x}) := p_i(\boldsymbol{x}) \cdot v_i - c(x_i).$$

Our paper encapsulates contestants' heterogeneity into the difference in their prize valuations. The model depicts a context in which the prize is nonmonetary and contestants value it differently. It should be noted that our analysis accommodates an alternative setup that allows for heterogeneity in effort costs. To see this, suppose that the prize carries a common monetary value—which we normalize to unity—while contestants differ in their abilities. Following Moldovanu and Sela (2001, 2006) and Moldovanu, Sela, and Shi (2007), a contestant i's effort cost takes the form $c_i(x_i) = c(x_i)/d_i$, with $d_1 \geq \ldots \geq d_n > 0$. The parameter d_i measures one's ability: A more competent contestant is endowed with a larger d_i and bears a lower effort cost. Each contestant chooses effort x_i to maximize the expected payoff $p_i(\boldsymbol{x}) - c(x_i)/d_i$, which is equivalent to maximizing $p_i(\boldsymbol{x}) \cdot d_i - c(x_i)$. The game is isomorphic to that in our baseline setting, and the parameter d_i plays the same role as v_i. The analysis in the baseline setting naturally extends.[6]

2.2 Regularity Condition and Equilibrium Property

The set of impact functions $\left\{f_i(\cdot)\right\}_{i=1}^n$, together with contestants' valuations $\boldsymbol{v} \equiv (v_1, \ldots, v_n)$ and the effort cost function $c(\cdot)$, defines a simultaneous-move contest game. We impose the following regularity condition.

Definition 1 ***(Regular Concave Contests)*** *A contest* $\left(\boldsymbol{v}, \left\{f_i(\cdot)\right\}_{i=1}^n, c(\cdot)\right)$ *is called a regular concave contest if (i) the impact function for contestant* $i \in \mathcal{N}$ *is either a nonnegative constant or a twice-differentiable function, with* $f_i(x_i) \geq 0$, $f_i'(x_i) > 0$, *and* $f_i''(x_i) \leq 0$ *for all* $x_i \geq 0$*; and (ii) the effort cost function satisfies* $c(0) = 0$, $c'(x_i) > 0$, *and* $c''(x_i) \geq 0$ *for all* $x_i > 0$.

[5] This assumption is imposed to guarantee the existence of a pure-strategy Nash equilibrium.

[6] In an online appendix, we analyze an extended setting in which the heterogeneity in effort cost functions is modeled more generally.

The above definition simply requires the usual concave impact functions and a convex effort cost function, which ensure a concave payoff function in effort and are widely adopted in the literature. Szidarovszky and Okuguchi (1997) and Cornes and Hartley (2005) prove the existence and uniqueness of the equilibrium in the above contest game with $f_i(0) = 0$ for all $i \in \mathcal{N}$. Their results cannot be applied directly to contests that allow for headstarts, i.e., $f_i(0) > 0$ for some $i \in \mathcal{N}$. The following theorem relaxes the zero-headstart assumption.

Theorem 1 ***(Existence and Uniqueness of Equilibrium)*** *There exists a unique pure-strategy Nash equilibrium in a regular concave contest game* $\left(\boldsymbol{v}, \left\{f_i(\cdot)\right\}_{i=1}^n, c(\cdot)\right)$.

Our study focuses on the above-defined concave contests for two reasons. First, when impact functions are convex, a pure-strategy equilibrium does not often exist. Although mixed-strategy equilibria exist, they generally are not unique and their properties remain elusive in the literature (e.g., Ewerhart, 2015, 2017). Second, the condition alludes to the usual production technology with nonincreasing marginal output.

2.3 Design Instruments and Contest Objectives

Theorem 1 allows us to set up the contest design problem in a two-stage structure. First, the designer sets the contest rule and announces it publicly; second, contestants exert effort simultaneously to vie for the prize. We first discuss the instruments available to the designer and then elaborate on the properties and implications of the objective function.

2.3.1 Design Instruments

We follow the tradition in the literature and mainly focus on two types of instruments to model identity-dependent preferential treatment: (i) multiplicative biases—i.e., weights on contestants' effective output—and (ii) additive headstarts. To put this formally, the impact function takes the form

$$f_i\left(x_i; \alpha_i, \beta_i\right) = \alpha_i \cdot h\left(x_i\right) + \beta_i. \tag{2}$$

The function $h(\cdot)$ is exogenously given as the fundamental *contest technology*;[7] the identity-dependent treatment imposed on each contestant $i \in \mathcal{N}$ is given by a tuple (α_i, β_i), with $\alpha_i, \beta_i \geq 0$.[8] The contest technology $h(\cdot)$ is assumed to have the following properties.

[7] In an online appendix, we analyze an extended setting in which contestants are endowed with heterogeneous contest technologies $h_i(\cdot)$.

[8] Drugov and Ryvkin (2017) study a two-player contest with headstart in which contestant 1 wins with a probability $p_1 = (x_1 + \beta)/(x_1 + x_2)$, and contestant 2 wins with a probability $1 - p_1$. This two-player contest is equivalent to a lottery contest in which contestants 1 and 2 are endowed with an identity-dependent headstart of β and $-\beta$, respectively.

Assumption 1 ***(Concave Contest Technology)*** *$h(\cdot)$ is twice differentiable, with $h(0) = 0$, $h'(x) > 0$, and $h''(x) \leq 0$ for all $x > 0$.*[9]

Both the multiplicative bias, α_i, and the additive headstart, β_i, are popularly adopted in the literature to model preferential treatments. Fu (2006); Franke (2012); Franke, Kanzow, Leininger, and Schwartz (2013, 2014); and Epstein, Mealem, and Nitzan (2011) focus on the former, while Clark and Riis (2000); Konrad (2002); Siegel (2009, 2014); Kirkegaard (2012); and Li and Yu (2012) consider the latter. Franke, Leininger, and Wasser (2018) allow for both. Both instruments vary a contestant's (deterministic) output, but through starkly different channels: α_i scales a contestant's output up or down for any given effort, while β_i directly adds to it regardless of his effort. The contrast inspires interesting comparisons, which generate useful implications for contest design.

2.3.2 A General Objective Function

The designer chooses $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ to maximize an objective function $\Lambda(\cdot)$, which is a function of the effort profile $\boldsymbol{x} \equiv (x_1, \ldots, x_n)$; the profile of winning probabilities $\boldsymbol{p} \equiv (p_1, \ldots, p_n)$; and the profile of prize valuations $\boldsymbol{v} \equiv (v_1, \ldots, v_n)$. We impose the following regularity condition on $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$.

Assumption 2 ***(Objective Function)*** *Fixing $\boldsymbol{p} \equiv (p_1, \ldots, p_n)$ and $\boldsymbol{v} \equiv (v_1, \ldots, v_n)$, $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$ is weakly increasing in x_i for all $i \in \mathcal{N}$.*

The assumption simply requires that contestants' efforts accrue to the benefit of the contest designer: For a given winning probability distribution $\boldsymbol{p}$, an increase in a contestant's effort does not reduce the designer's payoff.

The objective function $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$ encompasses a wide array of scenarios. First consider the following:

$$\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) := \sum_{i=1}^{n} x_i + \psi \sum_{i=1}^{n} p_i v_i - \gamma \sum_{i=1}^{n} \left(p_i - \frac{\sum_{j=1}^{n} p_j}{n} \right)^2, \tag{3}$$

with $\psi \geq 0$ and $\gamma \geq 0$. The function obviously satisfies Assumption 2.

When the weights ψ and γ both reduce to zero, the above expression boils down to $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = \sum_{i=1}^{n} x_i$, the popularly studied objective of total effort maximization. The objective function (3) allows the designer to have a direct preference for contestants' winning probability distribution. The term $\sum_{i=1}^{n} \left(p_i - (\sum_{j=1}^{n} p_j)/n\right)^2$ is the variance of the winning

[9] With $\alpha_i, \beta_i \geq 0$, Assumption 1 ensures that the game satisfies the requirements of Definition 1, and Theorem 1 applies, by which a unique pure-strategy equilibrium exists.

probabilities. With $\gamma > 0$, the designer prefers a less predictable outcome. For instance, in sports competitions, spectators often not only appreciate contenders' efforts, but also demand more suspense about the eventual winner (see Chan, Courty, and Hao, 2008; and Ely, Frankel, and Kamenica, 2015).[10] The contest objective also accommodates the pursuit of selection efficiency (see Meyer, 1991; Hvide and Kristiansen, 2003; Ryvkin and Ortmann, 2008; and Fang and Noe, 2018): The additional component $\sum_{i=1}^{n} p_i v_i$ strictly increases when a contestant of a higher valuation is able to win more often, which also provides an example of how contestants' prize valuations could directly affect the designer's payoff.[11]

In many competitive events, however, only the winner's effort is relevant to the organizer's interest. Suppose that the contest designer only cares about the expected winner's effort. The objective function can be written as

$$\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = \sum_{i=1}^{n} p_i x_i, \tag{4}$$

which clearly satisfies Assumption 2. This objective function has gained increasing attention in the literature (e.g., Moldovanu and Sela, 2006; Serena, 2017; and Barbieri and Serena, 2019). A CEO succession race motivates candidates to develop their managerial skills when carrying out assigned tasks: Large public firms—e.g., GE and HP—often have difficulty retaining losing candidates, which would lead them to focus only on the acquisition of human capital from the winner (Fu and Wu, 2019b).[12]

[10]Such a preference is also assumed by Fort and Quirk (1995), Szymanski (2003), and Runkel (2006) in two-player settings.

[11]The contest designer may care about both effort supply and contestants' welfare (e.g., Epstein, Mealem, and Nitzan, 2011). Recall that a contestant i has an expected payoff $\pi_i = p_i v_i - x_i$ with linear effort cost functions. This preference can formally be expressed as $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) := \phi \sum_{i=1}^{n} \pi_i + (1-\phi)\sum_{i=1}^{n} x_i = \phi \sum_{i=1}^{n} p_i v_i + (1-2\phi)\sum_{i=1}^{n} x_i$. Assumption 2 is satisfied if and only if $\phi \leq \frac{1}{2}$, in which case this objective function boils down to a case of the objective function (3). Higher efforts, however, would cause net disutility to the designer if her preference over contestants' welfare is excessively strong—i.e., $\phi > \frac{1}{2}$—which defies Assumption 2.

[12]It is useful to point out that the expected winner's effort may differ subtly from the expected winner's performance. As previously noted, a generalized lottery contest can either be underpinned by a noisy tournament adapted from a discrete-choice model or a research tournament. Contestants' output or performance is a random variable that increases with their efforts. Fu and Wu (2019b) consider a succession race in which a firm selects a CEO based on observed output, but candidates' efforts add to their human capital, which leads to objective (4) when the firm only cares about the successor's skill. However, when the designer benefits from the winner's noisy output or performance—e.g., a procurement tournament or an architectural design competition—the objective function will be formulated alternatively, depending on the underlying noisy production process. In a noisy tournament, it is given by $\sum_{i=1}^{n} p_i f_i(x_i)$. In a research tournament à la Fullerton and McAfee (1999), it is $\sum_{i=1}^{n} f_i(x_i)$.

3 Optimal Contest Design: Analysis

Given the existence and uniqueness of a pure-strategy equilibrium in the contest game for arbitrary $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, the optimal contest design problem yields a typical mathematical program with equilibrium constraints (MPEC): Contestants' equilibrium effort profile, $\boldsymbol{x}$, is endogenously determined in the equilibrium as a function of $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, and the designer chooses $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ for the following optimization problem:

$$
\begin{aligned}
& \max_{\{\boldsymbol{x},\boldsymbol{\alpha},\boldsymbol{\beta}\}} \Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) \\
\text{subject to } x_i &= \arg\max_{x_i \geq 0} \pi_i(\boldsymbol{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \\
p_i(\boldsymbol{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}) &= \begin{cases} \dfrac{f_i(x_i; \alpha_i, \beta_i)}{\sum_{j=1}^n f_j(x_j; \alpha_j, \beta_j)} & \text{if } \sum_{j=1}^n f_j(x_j; \alpha_j, \beta_j) > 0, \\[2ex] \dfrac{1}{n} & \text{if } \sum_{j=1}^n f_j(x_j; \alpha_j, \beta_j) = 0. \end{cases}
\end{aligned}
$$

The conventional approach requires an equilibrium solution of effort profile $\boldsymbol{x}$ for an arbitrary $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, which is, in general, unavailable. We take a detour to bypass the difficulty, and the approach can be described as follows:

i. We resort to the first-order conditions for the unique equilibrium of a contest game under an arbitrary contest rule $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, and show that the optimum can always be achieved by a contest rule with zero headstart. This allows us to focus on only the optimal choice of $\boldsymbol{\alpha}$.

ii. We establish a correspondence between contestants' equilibrium effort profile $\boldsymbol{x}$ and equilibrium winning probability distribution $\boldsymbol{p}$.

iii. Based on the correspondence noted above, we rewrite the objective as a function of the winning probability distribution. Instead of searching directly for the optimal contest rule, we let the designer assign equilibrium winning probabilities to contestants. We then solve for the probability distribution that maximizes the objective function.

iv. Finally, we identify the contest rule that induces the desirable winning probability distribution in equilibrium.

In the unique equilibrium of a contest game, the first-order condition $\partial \pi_i(\boldsymbol{x})/\partial x_i = 0$ must be satisfied for an active contestant $i \in \mathcal{N}$. With the impact functions specified in expression (2), the condition can be rewritten as

$$\frac{\sum_{j\neq i}\left[\alpha_j h(x_j)+\beta_j\right]}{\left\{\sum_{j=1}^{n}\left[\alpha_j h(x_j)+\beta_j\right]\right\}^2}\cdot h'(x_i)=\frac{1}{\alpha_i v_i}\cdot c'(x_i),\text{ for } x_i>0.$$

Similarly, the following inequality holds if contestant i remains inactive in equilibrium:

$$\frac{\sum_{j\neq i}\left[\alpha_j h(x_j)+\beta_j\right]}{\left\{\sum_{j=1}^{n}\left[\alpha_j h(x_j)+\beta_j\right]\right\}^2}\cdot h'(x_i)\leq\frac{1}{\alpha_i v_i}\cdot c'(x_i),\text{ for } x_i=0.$$

The above equilibrium conditions, together with the winning probability $p_i(\boldsymbol{x})$ specified in Equation (1), imply immediately that

$$p_i(1-p_i)v_i=c'(x_i)\cdot\frac{\alpha_i h(x_i)+\beta_i}{\alpha_i h'(x_i)},\text{ for } x_i>0,^{13} \tag{5}$$

and

$$p_i(1-p_i)v_i\leq c'(x_i)\cdot\frac{\alpha_i h(x_i)+\beta_i}{\alpha_i h'(x_i)},\text{ for } x_i=0. \tag{6}$$

3.1 Suboptimality of Additive Headstart

We now demonstrate that multiplicative biases outperform additive headstarts. Specifically, we show that fixing an arbitrary contest rule with positive headstarts, we can always construct an alternative contest rule with zero headstart that induces the same equilibrium winning probability distribution but strictly higher effort.

A sketch proof is provided below. Denote by $(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*)\equiv\big((\alpha_1^*,\ldots,\alpha_n^*),(\beta_1^*,\ldots,\beta_n^*)\big)$ the optimal contest rule that maximizes $\Lambda(\boldsymbol{x},\boldsymbol{p},\boldsymbol{v})$; the corresponding equilibrium effort profile and winning probabilities are denoted by $\boldsymbol{x}^*\equiv(x_1^*,\ldots,x_n^*)$ and $\boldsymbol{p}^*\equiv(p_1^*,\ldots,p_n^*)$, respectively. Suppose that $\beta_t^*>0$ for some $t\in\mathcal{N}$ in the optimum. We focus on an arbitrary active contestant t, i.e., $x_t^*>0$, as the logic naturally extends to inactive ones with $x_t^*=0$. Recall the equilibrium condition

$$p_t^*(1-p_t^*)v_t=c'(x_t^*)\cdot\frac{\alpha_t^* h(x_t^*)+\beta_t^*}{\alpha_t^* h'(x_t^*)}.$$

[13] We need $\alpha_i>0$ for the right-hand side to be well defined, which clearly holds. In fact, if $\alpha_i=0$, it is straightforward to see that $x_i=0$ is a strictly dominant strategy for player i due to the fact that costly effort has zero impact on player i's winning probability.

Denote by $x^\dagger$ the unique solution to the following equation:

$$c'(x_t^*) \cdot \frac{\alpha_t^* h(x_t^*) + \beta_t^*}{\alpha_t^* h'(x_t^*)} = c'(x^\dagger) \cdot \frac{h(x^\dagger)}{h'(x^\dagger)}.\quad (7)$$[14]

Simple analysis would verify that $x^\dagger > x_t^*$, given $\beta_t^* > 0$. Consider an alternative contest rule with $\widetilde{\boldsymbol{\alpha}} \equiv (\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_n)$ and $\widetilde{\boldsymbol{\beta}} \equiv (\widetilde{\beta}_1, \ldots, \widetilde{\beta}_n)$, such that

$$\left(\widetilde{\alpha}_i, \widetilde{\beta}_i\right) := \begin{cases} \left(\frac{\alpha_t^* h(x_t^*) + \beta_t^*}{h(x^\dagger)}, 0\right) & \text{for} \quad i = t, \\ (\alpha_i^*, \beta_i^*) & \text{for} \quad i \neq t. \end{cases}$$

In words, all contestants are awarded the same identity-dependent treatment as before except for contestant t. The new contest rule removes the headstart for contestant t. Simple algebra verifies that the equilibrium effort profile under the new contest rule $(\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\beta}})$—which we denote by $\widetilde{\boldsymbol{x}}^* \equiv (\widetilde{x}_1^*, \ldots, \widetilde{x}_n^*)$—is given by

$$\widetilde{x}_i^* = \begin{cases} x^\dagger & \text{for} \quad i = t, \\ x_i^* & \text{for} \quad i \neq t. \end{cases}$$

The new contest rule outperforms under Assumption 2. It induces the same winning probability distribution, because $\widetilde{\alpha}_t \cdot h(x^\dagger) + \widetilde{\beta}_t = \alpha_t^* \cdot h(x_t^*) + \beta_t^*$ by our construction, while the effort of contestant t strictly increases because $x^\dagger > x_t^*$ by Equation (7).[15] This argument leads to the following.

Theorem 2 ***(Suboptimality of Headstart)*** *Suppose that Assumptions 1 and 2 are satisfied. The optimum can always be achieved by choosing multiplicative biases $\boldsymbol{\alpha}$ only and setting headstarts $\boldsymbol{\beta}$ to zero.*

It is thus without loss of generality to abstract away headstart and focus on multiplicative biases when searching for the optimal biased contests, i.e., assuming $f_i(x_i; \alpha_i, \beta_i) = \alpha_i \cdot h(x_i)$, with $\beta_i = 0$ for all $i \in \mathcal{N}$.[16] Franke, Leininger, and Wasser (2018, Proposition 3.6) obtain

[14]The existence and uniqueness of the solution $x^\dagger$ follows from the facts that $c'(x) \cdot h(x)/h'(x)$ is strictly increasing in x, $\lim_{x \searrow 0} c'(x) \cdot h(x)/h'(x) = 0$, and $\lim_{x \nearrow \infty} c'(x) \cdot h(x)/h'(x) = \infty$.

[15]A closer inspection of Equation (7) indicates that $x^\dagger > x_t^*$ may not hold if the headstart β_t is allowed to be negative, in which case the comparison depends on the properties of $c'(\cdot)$, $h(\cdot)$, and $h'(\cdot)$. Drugov and Ryvkin (2017) allow for negative headstart (see Footnote 8) and show that a deviation from zero headstart can locally improve the performance of the contest, depending on the sign of $c'''(\cdot)$. They focus on the local property of the objective function with respect to the design instrument. It is noteworthy that negative headstart could nullify the contest success function (1) and cause irregularity to the contest game when examining the global property of the objective function. We therefore focus on a setting of $\boldsymbol{\beta} \geq \mathbf{0}$.

[16]Headstarts, however, can be preferred to multiplicative biases by a total-effort-maximizing contest designer in all-pay auctions. See Li and Yu (2012) and Franke, Leininger, and Wasser (2018) for more details.

similar results. Specifically, they show in a standard lottery contest—i.e., $h(x_i) = x_i$—that a positive headstart is suboptimal when the designer aims to maximize total effort. Our analysis generalizes Franke et al. (2018) in two dimensions: First, we allow for a flexible contest technology, and second, the optimization problem addresses a broad objective.

3.2 Reformulated Design Problem

Theorem 2 allows us to derive the fundamental equilibrium correspondence that underpins our optimization approach: With $\beta_i = 0$, the following must hold in an equilibrium:

$$p_i(1-p_i)v_i = c'(x_i) \cdot \frac{h(x_i)}{h'(x_i)}, \forall\, i \in \mathcal{N}. \tag{8}$$

A system of n set-valued functional equations depicts the relation between winning probability distribution $\boldsymbol{p}$ and contestants' effort profile $\boldsymbol{x}$ in equilibrium, with the right-hand side strictly increasing with x_i. In what follows, we call the system of equations the *equilibrium correspondence* of the contest game. The correspondence reminds us of the first-order condition (5) for an active player. However, it also holds for an *inactive* contestant, as $x_i = 0$ is associated with $p_i = 0$. Further, define the inverse of $\log(c'(x) \cdot h(x)/h'(x))$ as $g(\cdot)$. Assumption 1 and the convexity of the effort cost function imply that $g(\cdot)$ is well defined. In particular, $g(\cdot)$ is a strictly increasing function, with $g(-\infty) = 0$ and $g(\infty) = \infty$. The correspondence (8) can be rewritten as

$$x_i = g\Big(\log(p_i\left(1-p_i\right)) + \log(v_i)\Big), \forall\, i \in \mathcal{N}. \tag{9}$$

Two remarks are in order. First, each equation in the system of equations (9) literally delineates a direct and unique relation between x_i and (p_i, v_i) for an individual contestant $i \in \mathcal{N}$. The equilibrium probability p_i can be viewed as a *sufficient statistic* of the equilibrium in the contest: p_i is not exogenously given, but endogenously determined jointly by contestants' equilibrium effort profile $\boldsymbol{x} \equiv (x_1, \ldots, x_n)$ and the treatment $\boldsymbol{\alpha} \equiv (\alpha_1, \ldots, \alpha_n)$. Second, the correspondence (9) unveils the nature of incentive provision in contests. A contestant's effort decision ultimately takes into account two basic factors: (i) value (v_i), i.e., how much he can be rewarded when he wins; and (ii) prospect (p_i), i.e., the expectation about how likely he is to win.

The correspondence (9) opens a new avenue for contest design. The objective function $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$ can be rewritten as $\Lambda\big(\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{v}), \boldsymbol{p}, \boldsymbol{v}\big)$; instead of setting $\boldsymbol{\alpha}$ directly, we treat winning probability distribution $\boldsymbol{p}$ as the design variable and let the designer maximize

$\Lambda\big(\boldsymbol{x}(\boldsymbol{p},\boldsymbol{v}),\boldsymbol{p},\boldsymbol{v}\big)$, subject to (9) and the following feasibility constraints:

$$\sum_{i=1}^{n} p_i = 1, \text{ and } p_i \geq 0, \text{ for all } i \in \mathcal{N}. \tag{10}$$

A maximizer automatically exists for any smooth and continuous objective $\Lambda\big(\boldsymbol{x}(\boldsymbol{p},\boldsymbol{v}),\boldsymbol{p},\boldsymbol{v}\big)$ given that the choice set, defined by (10), is an $(n-1)$-dimensional simplex. The following is established as the last piece of the puzzle.

Theorem 3 ***(Implementing Winning Probabilities by Setting Biases)*** *Fix any equilibrium winning probability distribution $\boldsymbol{p} \equiv (p_1, \ldots, p_n) \in \Delta^{n-1}$.*

i. If $p_j = 1$ for some $j \in \mathcal{N}$, then $\boldsymbol{p} \equiv (p_1, \ldots, p_n)$ can be induced by the following set of biases $\boldsymbol{\alpha}(\boldsymbol{p}) \equiv \big(\alpha_1(\boldsymbol{p}), \ldots, \alpha_n(\boldsymbol{p})\big)$:

$$\alpha_i(\boldsymbol{p}) = \begin{cases} 1 & \text{if } \; i = j, \\ 0 & \text{if } \; i \neq j. \end{cases}$$

ii. If there exist at least two active contestants, then $\boldsymbol{p} \equiv (p_1, \ldots, p_n)$ can be induced by the following set of biases $\boldsymbol{\alpha}(\boldsymbol{p}) \equiv \big(\alpha_1(\boldsymbol{p}), \ldots, \alpha_n(\boldsymbol{p})\big)$:

$$\alpha_i(\boldsymbol{p}) = \begin{cases} \frac{p_i}{h\Big(g\big(\log(p_i(1-p_i))+\log(v_i)\big)\Big)} & \text{if } \; p_i > 0, \\ 0 & \text{if } \; p_i = 0. \end{cases} \tag{11}$$

Theorem 3 formally states that the contest designer can properly construct the set of weights $\boldsymbol{\alpha}$ to induce any equilibrium winning probability distribution.[17] The result closes the loop for the reformulated optimization problem: Upon obtaining the maximizer to $\Lambda\big(\boldsymbol{x}(\boldsymbol{p},\boldsymbol{v}),\boldsymbol{p},\boldsymbol{v}\big)$, the optimal biases $\boldsymbol{\alpha}^* \equiv (\alpha_1^*, \ldots, \alpha_n^*)$ can readily be identified by invoking Theorem 3.

Consider, for example, the widely studied Tullock contest with $h(x_i) = (x_i)^r$ and assume a linear effort cost function $c(x_i) = x_i$. An equation in the correspondence (9) boils down to $x_i = rp_i(1-p_i)v_i$. The above-mentioned objective function (3) can be rewritten as

$$\Lambda\big(\boldsymbol{x}(\boldsymbol{p},\boldsymbol{v}),\boldsymbol{p},\boldsymbol{v}\big) := \sum_{i=1}^{n}\big[rp_i(1-p_i)v_i\big] + \psi\sum_{i=1}^{n} p_i v_i - \gamma \sum_{i=1}^{n}\left(p_i - \frac{\sum_{j=1}^{n} p_j}{n}\right)^2,$$

[17]It should be noted that the biases $\boldsymbol{\alpha}$ that induce each given $\boldsymbol{p}$ are not unique. For instance, the same equilibrium outcome can be induced by multiplying all α_i by some positive factor.

which gives rise to a quadratic programming. Standard technique would obtain a handy closed-form solution to the optimal biases $\boldsymbol{\alpha}$.[18] In contrast, we primarily focus on the general implications of the contest design problem, instead of solving for closed-form solutions in specific settings.

The reformulation enormously simplifies the design problem. By the equilibrium correspondence (9), each contestant chooses his effort as if he responds merely to (p_i, v_i), his own winning odds and prize valuation: The strategic linkages between contestants seemingly dissolve when the winning probability distribution is treated as a design variable. This approach insulates the designer from the distraction of the complex strategic interaction of the contest game; instead, the reformulated optimization problem boils down to a simple programming that allocates probability mass among contestants purely based on the profile of their prize valuations.

3.3 A General Exclusion Principle

Recall that the contest designer, when setting $\boldsymbol{\alpha}$, can effectively exclude a contestant by imposing zero weight on his entry, which discourages him from exerting positive effort. We now explore the hidden dimension of the design problem: Which contestants should be included in the optimal contest?

Define $\tau : \mathcal{N} \to \mathcal{N}$ as a permutation of the set of players $\mathcal{N} \equiv \{1, \ldots, n\}$. In particular, player i is replaced by player $\tau(i)$ in the rearrangement. With slight abuse of notation, let us define $\tau(\boldsymbol{x}) := (x_{\tau(1)}, \ldots, x_{\tau(n)})$, $\tau(\boldsymbol{p}) := (p_{\tau(1)}, \ldots, p_{\tau(n)})$, and $\tau(\boldsymbol{v}) := (v_{\tau(1)}, \ldots, v_{\tau(n)})$. Similarly, let $\tau_{ij}(\boldsymbol{x})$ denote the permutation obtained by swapping contestants i and j. To obtain more mileage, we impose the following condition on $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$.

Assumption 3 *The contest designer's objective $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$ satisfies the following properties:*

i. for all permutations τ of $\mathcal{N}$, $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = \Lambda\big(\tau(\boldsymbol{x}), \tau(\boldsymbol{p}), \tau(\boldsymbol{v})\big)$;

ii. if $(p_i, x_i) = (0, 0)$ for some contestant $i \in \mathcal{N}$, then $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) \leq \Lambda\big(\boldsymbol{x}, \boldsymbol{p}, \tau_{ij}(\boldsymbol{v})\big)$ for all $j \in \mathcal{N}$ such that $v_j < v_i$;

iii. fixing $\boldsymbol{p} \equiv (p_1, \ldots, p_n)$ and $\boldsymbol{v} \equiv (v_1, \ldots, v_n)$, $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$ is strictly increasing in x_i if $p_i > 0$.

Part (i) of the above assumption implies that the designer's preference is anonymous: She does not have ex ante preference over certain players. Part (ii) of the assumption indicates that the prize value for a contestant is more likely to accrue to the designer's benefit when

[18]The application of our optimization approach and the solutions to optimal biases in Tullock contest settings are available from the authors upon request.

he is active. The requirement is automatically satisfied in the simplest case in which the objective function is independent of contestants' prize valuations, e.g., in which the designer maximizes total effort or the expected winner's effort. Part (iii) states that the designer would strictly benefit if an active player exerts more effort.

Part (iii) of Assumption 3 implies Assumption 2:[19] Theorem 2 thus remains in place, and headstarts are suboptimal for contest design under Assumption 3. Assumption 3 is by no means restrictive, as all of the examples discussed in Section 2.3.2 satisfy the requirements. We obtain the following.

Theorem 4 ***(Exclusion Principle)*** *Suppose that Assumptions 1 and 3 are satisfied. If $p_i^* = 0$ for some $i \in \mathcal{N}$ in the optimum, then $p_j^* = 0$ for all $j \in \mathcal{N}$, with $v_j < v_i$.*

By Theorem 4, exclusion in the optimum must be *monotone*: Whenever the designer intends to exclude contestants, she targets the ex ante weakest. This result stands in contrast to those obtained in previous studies. In an all-pay auction, Baye, Kovenock, and de Vries (1993) show that a total-effort-maximizing contest designer may strategically exclude the strongest contestant. In contrast, Fang (2002) demonstrates that the designer does not have a strict incentive to exclude players from a lottery contest—i.e., $h(x_i) = x_i$. Both studies assume total effort maximization and outright exclusion, while we allow for a general objective function and an indirect exclusion approach, i.e., allowing the designer to bias the contest to discourage certain contestants' participation.

The monotone exclusion principle may compel one to conjecture that an ex ante stronger contestant—i.e., one with a larger v_i—would win with a (weakly) higher probability in the optimum. However, this may not hold in general. We will elaborate in Section 4.

3.4 Optimal Contests: Maximizing Total Effort and the Expected Winner's Effort

We now apply our approach to two typical scenarios for contest design. First, we set ψ and γ in the objective function (3) to zero, and consider the situation in which the contest designer aims to maximize aggregate effort, i.e., $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = \sum_{i=1}^n x_i$. Second, we consider the objective function (4), the maximization of the expected winner's effort—i.e., $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = \sum_{i=1}^n p_i x_i$.

Maximizing Total Effort With slight abuse of notation, we denote, respectively, by $\boldsymbol{p}^* \equiv (p_1^*, \ldots, p_n^*)$ and $\boldsymbol{\alpha}^* \equiv (\alpha_1^*, \ldots, \alpha_n^*)$ the total-effort-maximizing winning probabilities

[19] To be more rigorous, we need to impose the condition that $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v})$ is weakly increasing in x_i at $p_i = 0$ for all $i \in \mathcal{N}$.

and the corresponding optimal biases. Consider a two-player contest with $v_1 \geq v_2$. It is well known in the literature that in a Tullock contest setting—i.e., $h(x_i) = (x_i)^r$—the optimum fully balances the playing field, with $p_1^* = p_2^* = \frac{1}{2}$, for all $r \in (0, 1]$. This can be achieved by setting α_2^* to $(v_1/v_2)^r$ with $(v_1/v_2)^r \geq 1$ and normalizing α_1^* to one. By the equilibrium correspondence, the analysis can readily accommodate flexible contest technology $h(\cdot)$ and multiple players. Recall that in the equilibrium,

$$x_i = g\Big(\log\big(p_i(1-p_i)\big) + \log(v_i)\Big), \forall\, i \in \mathcal{N},$$

which indicates that x_i strictly increases with $p_i(1-p_i)$. Note that $p_i\,(1-p_i)$ is nonmonotone in p_i: It first increases and then drops, being maximized uniquely at $p_i = \frac{1}{2}$. To put this intuitively, one gives up when he faces a slim chance of winning, while he also slacks off when he expects an easy win, which underpins the nonmonotone best-response function in a standard contest game (Dixit, 1987). This observation implies immediately that the total-effort-maximizing contest perfectly levels the playing field—i.e., $p_1^* = p_2^* = 1/2$—in a two-player contest, regardless of $h(\cdot)$. This generalizes the conventional wisdom obtained in previous studies. Moreover, the following proposition can be obtained.

Proposition 1 ***(Total-effort-maximizing Contests)*** *Suppose that $n \geq 2$, Assumption 1 is satisfied, and the designer aims to maximize total effort. Then the following statements hold:*

i. The optimal contest allows for at least three active players if possible.

ii. The optimal contest does not allow any contestant to win with a probability more than 1/2, i.e., $p_i^* \leq 1/2$, $\forall\, i \in \mathcal{N}$, with equality if and only if $n = 2$.

The first part of Proposition 1 generalizes Franke, Kanzow, Leininger, and Schwartz (2013, Theorem 4.6), and shows that a head-to-head competition is suboptimal whenever a third contestant is available, regardless of the distribution of prize valuations. Suppose otherwise that in a multiplayer contest only two players are kept active. Optimization requires that they have equal chance to win, as noted above. Recall that x_i strictly increases with $p_i(1-p_i)$, and $p_i(1-p_i)$ is maximized when $p_i = \frac{1}{2}$, with $d\left[p_i(1-p_i)\right]/dp_i\big|_{p_i=1/2} = 0$. With a simple additive objective function $\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = \sum_{i=1}^{n} x_i$, the designer can be strictly better off by adjusting contest rule $\boldsymbol{\alpha}$ to award a third player a very small probability of winning: In the new equilibrium, the third player contributes positive effort; the other two would barely reduce their effort, because the marginal effect on $p_i(1-p_i)$ is negligible.

The second part of Proposition 1 provides a key property of the optimum regarding the winning probability distribution. The optimum precludes a "superstar," in the sense that an individual contestant's winning odds must be strictly less than the sum of the others', i.e., $p_i^* < 1/2$, $\forall\, i \in \mathcal{N}$, whenever the contest involves three or more contestants. It is never optimal to let contestant i win with a probability p_i strictly more than $1/2$. Suppose the contrary. The designer, instead, can induce the same effort from contestant i by assigning $1-p_i$ and elicit more effort from the others by allocating to them the saved probability mass $2p_i - 1$.

It is unclear, in the case of $n \geq 3$, whether the optimal contest completely levels the playing field—i.e., $p_i^* = 1/n$—and whether an ex ante stronger contestant would necessarily be handicapped more, i.e., a larger v_i is associated with a smaller α_i in the optimum. We apply our approach to these classical questions in Section 4 and show that the conventional wisdom does not universally hold.

Maximizing the Expected Winner's Effort Next, we consider the maximization of the expected winner's effort. Unlike maximizing aggregate effort $\sum_{i=1}^n x_i$, the objective function $\sum_{i=1}^n p_i x_i$ is nonadditive in the contestant's effort, because the winning probability p_i is a function of effort profile $\boldsymbol{x}$ and is factored in multiplicatively. Our approach is immune to the nuance. Denote by $\boldsymbol{p}^{**} \equiv (p_1^{**}, \ldots, p_n^{**})$ the winning probabilities in the optimal contest. We obtain the following.

Proposition 2 ***(Optimal Contest that Maximizes the Expected Winner's Effort)*** *Suppose that Assumption 1 is satisfied and the designer aims to maximize the expected winner's effort. Then only the two ex ante strongest contestants would remain active in the optimal contest. Moreover, the ex ante stronger player always wins with a strictly higher probability than the underdog, independent of the shape of $g(\cdot)$. That is, if $v_1 > v_2$, then $p_1^{**} > p_2^{**} > 0$.*[20]

By Proposition 2, the optimal contest must sufficiently preserve individual incentives by including only the two most competitive contestants. The playing field is never fully balanced, as the winning probability assignment is "assortative," i.e., the top dog wins more often. This stands in contrast to the optimum established in Proposition 1 under total effort maximization for the case of $n = 2$.

The result can again be interpreted in light of the correspondence (9). Intuitively, maximizing the weighted sum $\sum_{i=1}^n p_i x_i$ requires that the probability mass be concentrated on the minimal number of the most productive contestants, i.e., the two strongest contestants. Further, suppose otherwise that the two active contestants win with equal chance. The designer can be strictly better off by shifting a small amount of probability mass from p_2 to p_1. Recall that $x_i = g\Big(\log\big(p_i(1-p_i)\big) + \log(v_i)\Big)$. Its impact on $p_i(1-p_i)$ fades away on the margin, while a larger probability is attached to a higher effort: $x_1 > x_2$ because $v_1 > v_2$.

[20]It is straightforward to show that $p_1^{**} = p_2^{**} = 1/2$ if $v_1 = v_2$.

4 Leveling the Playing Field: Reexamined

We now apply our approach to explore a classical question in the contest literature: How should the balance of the playing field be optimally set to maximize total effort when contestants are heterogeneous? The question can be examined in terms of either ex post *outcomes* or ex ante *contest rules*. The former concerns how contestants' winning odds are ranked in the optimum with respect to their innate strength, while the latter explores whether weaker contestants are favored vis-à-vis their stronger opponents. In Section 3.4, we generalize the conventional wisdom in a two-player setting and obtain that the optimum handicaps the stronger and equalizes winning odds regardless of the contest technology $h(\cdot)$. In an n-player lottery contest, Franke, Kanzow, Leininger, and Schwartz (2013) show in a numerical example that the optimal contest is biased in favor of weaker players—i.e., $\alpha_i^* < \alpha_j^*$ for $v_i > v_j$, and $x_i^*, x_j^* > 0$—although the playing field is not fully balanced—i.e., $p_i^* > p_j^*$ for $v_i > v_j$, and $x_i^*, x_j^* > 0$. Our approach allows us to examine this systematically.

4.1 Ranking of Winning Probabilities in the Optimum

Recall the function $g(\cdot)$, which is defined as the inverse of $\log\left(c'(x) \cdot h(x)/h'(x)\right)$. We first obtain the following.

Proposition 3 ***(Winning Probabilities in Total-effort-maximizing Contests)*** *Suppose that Assumption 1 is satisfied and the designer aims to maximize total effort. Consider a contest with $n \geq 3$ players. For two arbitrary active contestants $i, j \in \mathcal{N}$ with $v_i > v_j$, $p_i^* > p_j^*$ if $g(\cdot)$ is a strictly convex function.*

Proposition 3 predicts that for active contestants, a larger prize valuation ensures strictly higher equilibrium winning odds in the optimum when the function $g(\cdot)$ is convex. A convex $g(\cdot)$ is common. For instance, a Tullock contest with $h(x_i) = (x_i)^r$ and a linear effort cost leads to $g(z) = r\exp(z)$, which is evidently strictly convex.

The logic of Proposition 3 is straightforward in light of the fundamental correspondence:

$$x_i = g\left(\log\left(p_i(1-p_i)\right) + \log(v_i)\right), \forall\, i \in \mathcal{N}.$$

Obviously, x_i is supermodular in (p_i, v_i) when $g(\cdot)$ is strictly convex in its arguments: $\partial^2 x_i/\partial p_i \partial v_i$ must be strictly positive because by Proposition 1, $p_i^* < 1/2$ in the optimum. The function $g(\cdot)$ depicts how a contestant's effort choice takes into account prize value and the prospect for his win: One steps up his effort when he expects a more rewarding prize (i.e., increasing v_i) or when he is more confident (i.e., increasing p_i) for $p_i < 1/2$. The supermodularity implies that a brighter prospect for a win incentivizes a contestant more when he also benefits more from the prize. Total effort can be maximized only when the assignment of $\boldsymbol{p}$ with respect to $\boldsymbol{v}$ is assortative, i.e., assigning larger equilibrium winning probability to a contestant of larger prize valuation.

Analogously, the assignment is set to be reversed when the function turns concave. It should be noted that $g(\cdot)$ cannot be globally concave. Recall that the function is the inverse of $\log\big(c'(x)\cdot h(x)/h'(x)\big)$. For a contest technology $h(\cdot)$ that satisfies Assumption 1 and a cost function $c(x)$ with finite $c'(0)$, $\log\big(c'(x)\cdot h(x)/h'(x)\big)$ approaches negative infinity in the neighborhood of zero, which precludes globally concave $g(\cdot)$. An exhaustive comparative static of probability ranking is infeasible, because the property of $g(\cdot)$ remains elusive in general.

We construct a parameterized setting to illustrate the impact of $g(\cdot)$ on the probability series in the optimum. Assume a linear effort cost function $c(x) = x$, and parametrize the contest technology $h(\cdot)$ by a variable $\sigma \in (0, 1]$ as follows:

$$h_\sigma(x) := \exp\left(\int_1^x \frac{1}{\zeta_\sigma^{-1}(t)}dt\right),$$

where $\zeta_\sigma^{-1}(t)$ is the inverse function of $\zeta_\sigma(\cdot)$ given by

$$\zeta_\sigma(z) := \begin{cases} \frac{1}{2}z & \text{if } 0 < z < \sigma, \\ \sigma - \frac{\sigma^2}{2z} & \text{if } \sigma \le z \le 2, \\ \frac{\sigma^2}{8}z + (\sigma - \frac{1}{2}\sigma^2) & \text{if } z > 2. \end{cases}$$

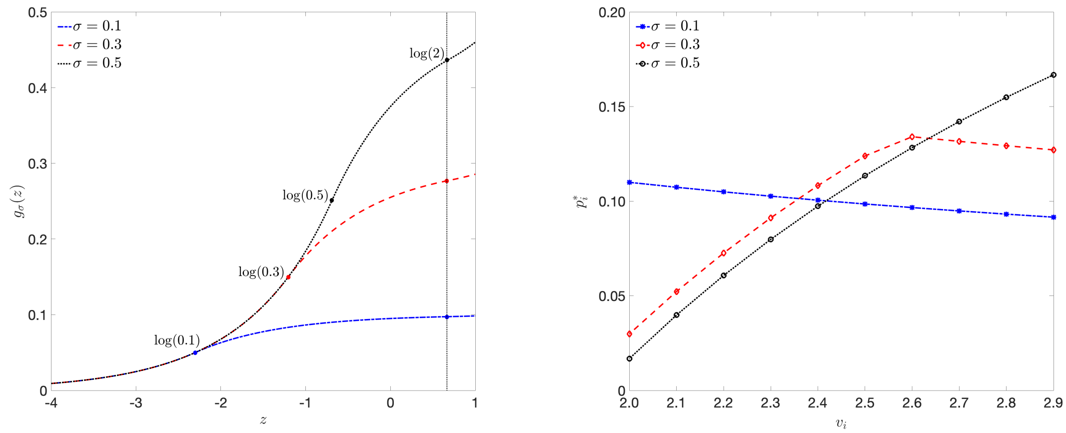
The expression of $g(\cdot)$, which we again index by σ, can be written as

$$g_\sigma(z) = \zeta_\sigma(e^z) = \begin{cases} \frac{1}{2}e^z & \text{if } z < \log\sigma, \\ \sigma - \frac{\sigma^2}{2}e^{-z} & \text{if } \log\sigma \le z \le \log 2, \\ \frac{\sigma^2}{8}e^z + \big(\sigma - \frac{1}{2}\sigma^2\big) & \text{if } z > \log 2. \end{cases}$$ [21]

The function $g_\sigma(z)$ is strictly convex in z for $z < \log\sigma$ and $z > \log 2$, and is strictly concave in z for $\log\sigma \le z \le \log 2$.

Suppose that $n = 10$ and $(v_1, v_2, \ldots, v_{10}) = (2.9, 2.8, \ldots, 2.0)$. With a linear effort cost function $c(x) = x$ and the constructed contest technology $h_\sigma(\cdot)$, contestant i's first-order

[21] Alternatively, the same $g_\sigma(\cdot)$ can be obtained by assuming the Tullock contest technology $h(x) = x^r$, with $r \in (0, 1]$, and an effort cost function $c(x) = r\int_0^x [e^{g_\sigma^{-1}(\omega)}/\omega]d\omega$. Our subsequent analysis would naturally extend to this alternative setting and obtains comparative statics with respect to the property of the cost function. It is straightforward to verify that the constructed effort cost function satisfies $c(0) = 0$, $c'(x) > 0$, and $c''(x) \ge 0$ for all $x > 0$.



(a) $g_\sigma(z)$ under different levels of σ (b) $(p_1^*, \dots, p_{10}^*)$ under different levels of σ

Figure 1: $g_\sigma(z)$ and $(p_1^*, \dots, p_{10}^*)$ under Different Levels of σ.

condition can now be rewritten as

$$p_i(1-p_i)v_i = \frac{h_\sigma(x_i)}{h'_\sigma(x_i)} = \zeta_\sigma^{-1}(x_i) \Rightarrow x_i = \zeta_\sigma\left(p_i(1-p_i)v_i\right).$$

Note that $p_i(1-p_i)v_i < 3/4 < 1$ in the example because $v_i < 3$ for all $i \in \mathcal{N} \equiv \{1, \dots, 10\}$. This indicates that the region $[0, \infty)$ in the support of $g_\sigma(\cdot)$ is irrelevant. The variable σ therefore measures the concavity/convexity of the $g_\sigma(\cdot)$ function in the relevant support $(-\infty, 0)$, as Figure 1(a) depicts: $g_\sigma(\cdot)$ is globally concave in the relevant support as $\sigma \searrow 0$; it is globally convex in the relevant support as $\sigma \nearrow 1$.

The profile of the optimal equilibrium probabilities $(p_1^*, \dots, p_{10}^*)$ for different values of σ are reported as follows:

σ	p_1^*	p_2^*	p_3^*	p_4^*	p_5^*	p_6^*	p_7^*	p_8^*	p_9^*	p_{10}^*
0.1	0.0915	0.0931	0.0948	0.0966	0.0985	0.1005	0.1026	0.1049	0.1073	0.1099
0.3	0.1271	0.1293	0.1316	0.1340	0.1239	0.1082	0.0912	0.0726	0.0522	0.0299
0.5	0.1668	0.1549	0.1421	0.1283	0.1134	0.0973	0.0798	0.0607	0.0398	0.0168

In the case of $\sigma = 0.5$, $p_i^* > p_j^*$ whenever $v_i > v_j$, as predicted by Proposition 3. In contrast, with $\sigma = 0.1$, $g_\sigma(\cdot)$ is concave in the relevant support and the ranking is entirely reversed, which implies that the optimal contest severely handicaps stronger contestants, such that they are less likely to win. The logic that underpins Proposition 3 can be flipped to interpret this observation. With a concave $g(\cdot)$, an increase in v_i reduces the marginal impact of p_i on x_i. A contestant can less effectively be motivated by an improvement in the prospect of a win when he has a higher valuation for the prize. This suggests that a lower winning probability should be assigned to a contestant with a higher prize valuation. The ranking is nonmonotone in the intermediate case of $\sigma = 0.3$. As Figure 1(b) illustrates, p_i^* strictly increases with i first and then decreases, with player 4 being the most probable winner.

4.2 Ranking of Multiplicative Biases in the Optimum

In this part, we examine the optimal contest rule—i.e., the multiplicative biases $\boldsymbol{\alpha}^*$—that maximizes total effort. Assume a Tullock contest with $n \geq 3$, $h(x_i) = (x_i)^r$, $r \in (0,1]$, and a linear effort cost function $c(x_i) = x_i$. The setting streamlines our analysis for two reasons. First, as stated above, the fundamental equilibrium correspondence under a Tullock contest setting can be simplified as

$$x_i = rp_i(1-p_i)v_i, \forall\, i \in \mathcal{N},$$

which allows for a closed-form solution of the optimal bias rule $\boldsymbol{\alpha}^*$, as the optimization problem yields a simple quadratic programming. Second, the total effort of the contest can be rewritten as $\sum_{i=1}^n x_i = r\sum_{i=1}^n p_i(1-p_i)v_i$, which implies immediately that the optimal probability distribution $\boldsymbol{p}^*$, or the winning probability ranking in the optimum, is independent of the parameter r. This allows us to focus on the property of optimal contest rule and enables lucid comparative statics with respect to r. The following fully characterizes the optimum.

Proposition 4 ***(Total-effort-maximizing Tullock Contests)*** *Assume without loss of generality that contestants are ordered such that $v_1 \geq v_2 \geq \ldots \geq v_n > 0$, $h(x_i) = (x_i)^r$, with $r \in (0,1]$, and $c(x_i) = x_i$. Suppose that the contest designer aims to maximize total effort. Then the equilibrium winning probabilities $\boldsymbol{p}^* \equiv (p_1^*, \ldots, p_n^*)$ are given by*

$$p_i^* = \begin{cases} \frac{1}{2}\left(1 - \frac{1}{v_i} \times \frac{\kappa-2}{\sum_{j=1}^{\kappa}\frac{1}{v_j}}\right) & \text{if } \; i \in \{1,\ldots,\kappa\}, \\ 0 & \text{if } \; i \in \mathcal{N} \setminus \{1,\ldots,\kappa\}, \end{cases} \tag{12}$$

where κ is given by

$$\kappa := \max\left\{ m = 2,\ldots,n \,\middle|\, \frac{m-2}{\sum_{j=1}^m \frac{1}{v_j}} < v_m \right\}.$$

Moreover, the corresponding weights, denoted by $\boldsymbol{\alpha}^* \equiv (\alpha_1^*, \ldots, \alpha_n^*)$, that induce $\boldsymbol{p}^* \equiv (p_1^*, \ldots, p_n^*)$ are given by

$$\alpha_i^* = \begin{cases} \frac{(p_i^*)^{1-r}}{\left[(1-p_i^*)v_i\right]^r} & \text{if } \; p_i^* > 0, \\ 0 & \text{if } \; p_i^* = 0. \end{cases}$$

Proposition 4 allows us to rank $\boldsymbol{\alpha}^* \equiv (\alpha_1^*, \ldots, \alpha_n^*)$ with respect to the parameter r.

Proposition 5 ***(Comparative Statics of the Optimal Biases with Respect to*** r***)*** *Assume without loss of generality that contestants are ordered such that* $v_1 \geq v_2 \geq \ldots \geq v_n > 0$*,* $h(x_i) = (x_i)^r$*, with* $r \in (0, 1]$*, and* $c(x_i) = x_i$*. Suppose that the contest designer aims to maximize total effort. Then the following holds:*

i. Suppose that contestants i *and* j *remain active in the total-effort-maximizing contest (i.e.,* $i, j \leq \kappa$*). If* $v_i > v_j$*, then there exists a cutoff* $\overline{r}_{ij} \in (0, 1)$ *such that* $\alpha_i^* \gtreqless \alpha_j^*$ *if* $r \lesseqgtr \overline{r}_{ij}$*.*

ii. Define an upper bound $\overline{r}_{\max} := \max_{\{i<j\leq\kappa\}}\{\overline{r}_{ij}\}$ *and a lower bound* $\overline{r}_{\min} := \min_{\{i<j\leq\kappa\}}\{\overline{r}_{ij}\}$*.* α_m^* *is decreasing in* $m \in \{1, \ldots, \kappa\}$ *when* $r \leq \overline{r}_{\min}$*, and is increasing when* $r \geq \overline{r}_{\max}$*. For* $r \in (\overline{r}_{\min}, \overline{r}_{\max})$*, the optimal biases* $\boldsymbol{\alpha}^*$ *are nonmonotone.*

Proposition 5 indicates that the usual leveling-the-playing-field principle does not hold in general. It first states that for a given pair of active contestants, the optimal bias rule can favor either depending on the size of r. More generally, Proposition 5(ii) identifies two cutoffs. When the contest sufficiently rewards more effort, i.e., $r \geq \overline{r}_{\max}$, a larger weight is assigned to a weaker active player, i.e., one with a lower prize valuation, in which case the conventional wisdom remains. In contrast, when r falls below a lower bound $\overline{r}_{\min}$, the prediction is entirely reversed, and the designer further upsets the balance of the contest in the optimum by favoring stronger contestants, i.e., α_m^* is decreasing in m.[22,23] When r falls in the intermediate range $(\overline{r}_{\min}, \overline{r}_{\max})$, the ranking of α_i^* is no longer monotone.

[22]Ample evidence can be found in practice for the practice of reverse handicapping in favor of ex ante stronger contenders. Consider, for instance, the widespread industry policy that gives unfair advantage to large organizations to promote "national champions" for domestic dominance and international preeminence; e.g., the dirigiste policy in France from 1945 to 1947 and Korea's industrialization programs. Alternatively, the financial fair-play regulation (FFP) in European football (soccer) has been broadly criticized for the anticompetition role it played in perpetuating the dominance of "big clubs": The rule requires that European football clubs balance their books and not spend more than the income they generate, which solidifies an incumbent "big" club's advantage in attracting talent, given the superior revenue it receives based on its past track record. Möller (2012) formally studies the trade-off between competitive balance and incentives in a dynamic contest in which one's early success improves his competence in future. He shows that an optimally designed contest may maximize the heterogeneity between players in terms of productivity along the dynamics.

[23]Soccer is broadly viewed as the least predictable major sporting discipline. Ben-Naim, Vazquez, and Redner (2007) and Anderson and Sally (2013) provide extensive empirical evidence that soccer matches produced "upsets"—i.e., pregame underdogs overcoming favorites—more frequently than other sports, which alludes to a relatively more significant role played by luck in soccer matches vis-à-vis skill or effort. Our result can thus arguably shed light on the European FFP regulation that advantages big clubs (see Footnote 22). This stands in contrast to various measures in the NBA—e.g., the draft lottery and salary cap—that maintain a level playing field. Anderson and Sally, among others, show that the results of basketball matches are the most predictable based on teams' quality (see https://knowledge.wharton.upenn.edu/article/sports-by-the-numbers-predicting-winners-and-losers/).

We construct a numerical example to illustrate the comparative statics. Again, suppose that $n = 10$ and $(v_1, v_2, \ldots, v_{10}) = (2.9, 2.8, \ldots, 2.0)$. To ease comparison with respect to r, we normalize the sum of optimal weights established by Proposition 4 to one and define $\alpha_i' \equiv \alpha_i^*/(\sum_{j=1}^n \alpha_j^*)$ for all $i \in \mathcal{N} \equiv \{1, \ldots, 10\}$.[24] The optimal bias rule for a given r can then be identified:

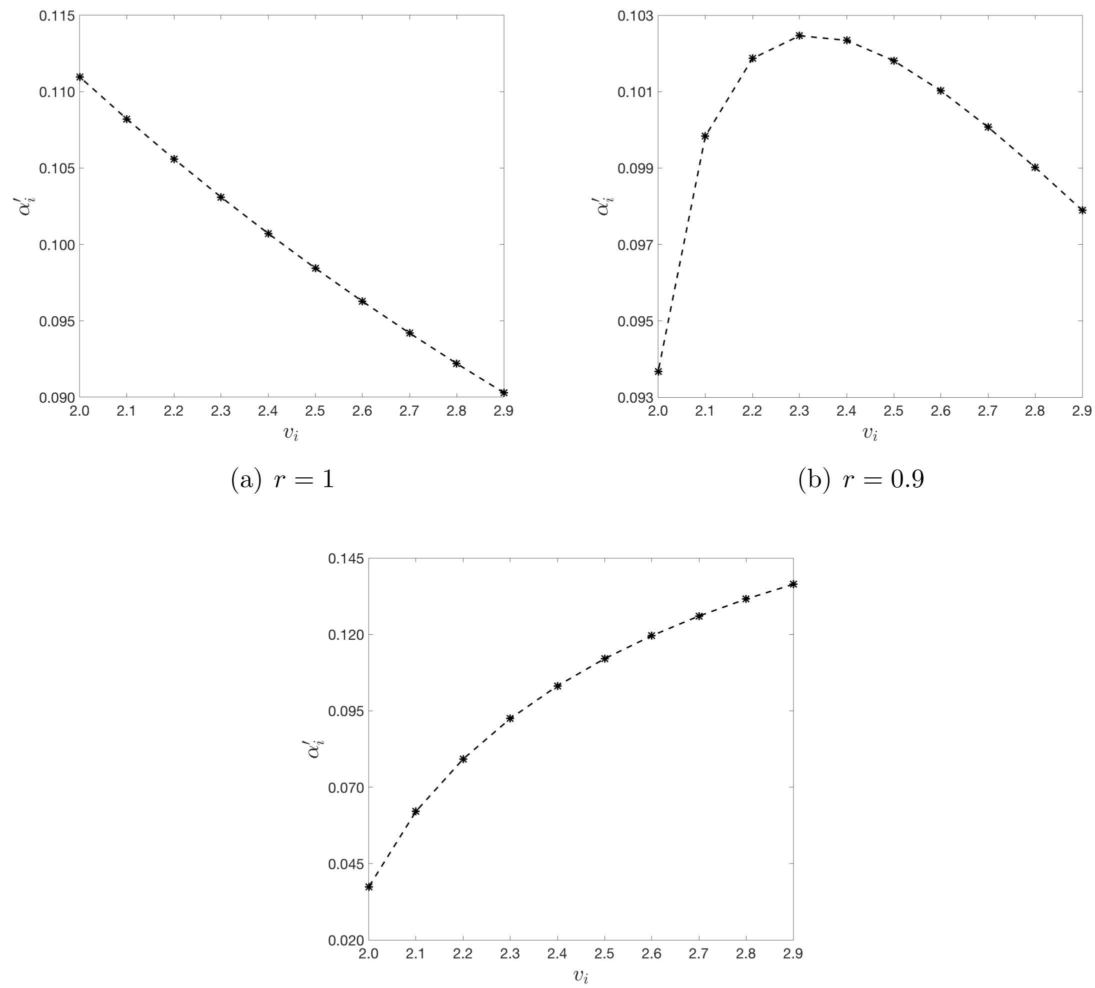
r	α_1'	α_2'	α_3'	α_4'	α_5'	α_6'	α_7'	α_8'	α_9'	α_{10}'
1.0	0.0903	0.0922	0.0942	0.0963	0.0984	0.1007	0.1031	0.1056	0.1082	0.1110
0.9	0.0979	0.0990	0.1001	0.1010	0.1018	0.1023	0.1025	0.1019	0.0998	0.0937
0.4	0.1364	0.1316	0.1260	0.1196	0.1121	0.1032	0.0925	0.0792	0.0621	0.0374

We illustrate the three cases in Figure 2. Monotone rankings of $(\alpha_1', \ldots, \alpha_{10}')$ arise in the case of both a large r ($r = 1$) and a small r ($r = 0.4$): The former exemplifies the conventional wisdom of leveling the playing field, while the latter entirely contradicts that. In the case of intermediate r ($r = 0.9$), contestant 7, with a prize valuation 2.3, is favored the most by the designer [see Figure 2(b)]: The optimal contest levels the playing field for contestants 1-7, but discounts the output of the weakest three. The second panel of Figure 2 depicts the case of nonmonotone ranking. The curve that traces α_m' with respect to contestants' prize valuation v_m is inverted U-shaped.

The optimal bias rule subtly depends on the parameter r. The comparative statics can again be interpreted in light of the fundamental correspondence and our optimization approach. As stated above, $\boldsymbol{p}^*$, the winning probability distribution in the optimum, remains constant regardless of r. Imagine that r decreases. A higher effort—contributed by a stronger contestant—can be less effectively converted into higher winning odds, which narrows the spread in $\boldsymbol{p}^*$ and, in turn, depletes contestants' effort incentives. To counteract this effect and restore the required distribution $\boldsymbol{p}^*$, a stronger contestant must be handicapped less severely because a larger α_i imposed on a stronger contestant enlarges the spread in the distribution of winning probabilities for any given effort profile.

More intuitively, recall the usual rationale for leveling the playing field: Preferential treatment motivates the underdog, which in turn prevents the favorite from slacking off. This logic can be cast into doubt when r decreases. A smaller r diminishes all contestants' incentives. On the one hand, a weaker contestant would respond less sensitively in his effort choice to the extra favor. On the other hand, a smaller r erodes a strong contestant's advantage because his higher effort is less effective for securing larger winning odds, which prevents him from slacking off regardless of the contest rule. When handicapping strong contestants, both the positive incentive effect for underdogs and the disciplinary effect on

[24] The variable α_i' can be interpreted as contestant i's winning probabilities if all contestants exert the same amount of effort.



(a) $r = 1$

(b) $r = 0.9$

(c) $r = 0.4$

Figure 2: Optimal Total-effort-maximizing Bias Rule under Different Levels of r.

the favorite diminish. The optimum could favor favorites more to preserve their momentum.

5 Concluding Remarks

In this paper, we develop a novel optimization approach to study the design of biased contests. A designer imposes identity-dependent preferential treatments on heterogeneous contestants. Based on a fundamental correspondence derived from the equilibrium condition, we characterize the general properties of the optimal contest rule in a substantially generalized setting without solving for the equilibrium explicitly. The analysis enabled by the approach generates useful theoretical implications that contrast starkly with those obtained in the restricted settings considered in previous studies. In particular, we demonstrate that the conventional wisdom of leveling the playing field may not hold in general. The contest rule could favor stronger contestants vis-à-vis their weaker opponents.

Our approach substantially eases the analysis of optimal contest design and can be applied to a broad array of scenarios. Fu and Wu (2019c) extend this approach to the setting of an all-pay auction and reexamine the classical issue of comparing all-pay auctions and lottery contests under general design objectives. The approach can also be applied in dynamic settings. For instance, Fu and Wu (2019b) consider a two-stage contest in which the designer assigns individualized weights to contestants' second-stage effort entries based on their first-stage ranking.

References

Anderson, Christopher and David Sally, *The numbers game: Why everything you know about soccer is wrong*, Penguin, 2013.

Barbieri, Stefano and Marco Serena, "Winners' efforts in team contests," *Working Paper*, 2019.

Baye, Michael R. and Heidrun C. Hoppe, "The strategic equivalence of rent-seeking, innovation, and patent-race games," *Games and Economic Behavior*, 2003, *44* (2), 217–226.

_ , Dan Kovenock, and Casper G. de Vries, "Rigging the lobbying process: An application of the all-pay auction," *American Economic Review*, 1993, *83* (1), 289–294.

Ben-Naim, E., F. Vazquez, and S. Redner, "What is the most competitive sport?," *Journal of the Korean Physical Society*, 2007, *50* (1), 124–126.

Chan, William, Pascal Courty, and Li Hao, "Suspense: Dynamic incentives in sports contests," *Economic Journal*, 2008, *119* (534), 24–46.

Che, Yeon-Koo and Ian L. Gale, "Optimal design of research contests," *American Economic Review*, 2003, *93* (3), 646–671.

Chowdhury, Subhasish M., Patricia Esteve-González, and Anwesha Mukherjee, "Heterogeneity, leveling the playing field, and affirmative action in contests," *Working Paper*, 2019.

Clark, Derek J. and Christian Riis, "On the win probability in rent-seeking games," *Mimeo*, 1996.

_ and _ , "Allocation efficiency in a competitive bribery game," *Journal of Economic Behavior & Organization*, 2000, *42* (1), 109–124.

Cornes, Richard and Roger Hartley, "Asymmetric contests with general technologies," *Economic Theory*, 2005, *26* (4), 923–946.

Dasgupta, Partha and Joseph Stiglitz, "Industrial structure and the nature of innovative activity," *Economic Journal*, 1980, *90* (358), 266–293.

Dixit, Avinash, "Strategic behavior in contests," *American Economic Review*, 1987, *77* (5), 891–898.

Drugov, Mikhail and Dmitry Ryvkin, "Biased contests for symmetric players," *Games and Economic Behavior*, 2017, *103*, 116–144.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica, "Suspense and surprise," *Journal of Political Economy*, 2015, *123* (1), 215–260.

Epstein, Gil S., Yosef Mealem, and Shmuel Nitzan, "Political culture and discrimination in contests," *Journal of Public Economics*, 2011, *95* (1), 88–93.

Ewerhart, Christian, "Mixed equilibria in Tullock contests," *Economic Theory*, 2015, *60* (1), 59–71.

— , "Contests with small noise and the robustness of the all-pay auction," *Games and Economic Behavior*, 2017, *105*, 195–211.

Fang, Dawei and Thomas Noe, "Lowering the bar and limiting the field: The effect of strategic risk-taking on selection contests," *Working Paper*, 2018.

Fang, Hanming, "Lottery versus all-pay auction models of lobbying," *Public Choice*, 2002, *112* (3-4), 351–371.

Fort, Rodney and James Quirk, "Cross-subsidization, incentives, and outcomes in professional team sports leagues," *Journal of Economic Literature*, 1995, *33* (3), 1265–1299.

Franke, Jörg, "Affirmative action in contest games," *European Journal of Political Economy*, 2012, *28* (1), 105–118.

— , Christian Kanzow, Wolfgang Leininger, and Alexandra Schwartz, "Effort maximization in asymmetric contest games with heterogeneous contestants," *Economic Theory*, 2013, *52* (2), 589–630.

— , — , — , and — , "Lottery versus all-pay auction contests: A revenue dominance theorem," *Games and Economic Behavior*, 2014, *83*, 116–126.

— , Wolfgang Leininger, and Cédric Wasser, "Optimal favoritism in all-pay auctions and lottery contests," *European Economic Review*, 2018, *104*, 22–37.

Fu, Qiang, "A theory of affirmative action in college admissions," *Economic Inquiry*, 2006, *44* (3), 420–428.

— and Jingfeng Lu, "Micro foundations of multi-prize lottery contests: A perspective of noisy performance ranking," *Social Choice and Welfare*, 2012, *38* (3), 497–517.

— and Zenan Wu, "Contests: Theory and topics," *Oxford Research Encyclopedia of Economics and Finance*, 2019a.

— and — , "Disclosure and favoritism in hierarchical competitions," *Working Paper*, 2019b.

— and — , "Auctions versus lottery all-pay contests: A general dominance theorem," *Working Paper*, 2019c.

— , Jingfeng Lu, and Yuanzhu Lu, "Incentivizing R&D: Prize or subsidies?," *International Journal of Industrial Organization*, 2012, *30* (1), 67–79.

Fullerton, Richard L. and R. Preston McAfee, "Auctioning entry into tournaments," *Journal of Political Economy*, 1999, *107* (3), 573–605.

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

Jia, Hao, "A stochastic derivation of the ratio form of contest success functions," *Public Choice*, 2008, *135* (3-4), 125–130.

Kirkegaard, René, "Favoritism in asymmetric contests: Head starts and handicaps," *Games and Economic Behavior*, 2012, *76* (1), 226–248.

Konrad, Kai A., "Investment in the absence of property rights; the role of incumbency advantages," *European Economic Review*, 2002, *46* (8), 1521–1537.

Lazear, Edward P. and Sherwin Rosen, "Rank-order tournaments as optimum labor contracts," *Journal of Political Economy*, 1981, *89* (5), 841–864.

Li, Sanxi and Jun Yu, "Contests with endogenous discrimination," *Economics Letters*, 2012, *117* (3), 834–836.

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

Manski, Charles F. and Daniel McFadden, *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press Cambridge, MA, 1981.

McFadden, Daniel, "Conditional logit analysis of qualitative choice behavior," in Zaremka P., ed., *Frontier in Econometrics*, 1973.

_ , "The measurement of urban travel demand," *Journal of Public Economics*, 1974, *3* (4), 303–328.

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.

_ and _ , "Contest architecture," *Journal of Economic Theory*, 2006, *126* (1), 70–96.

_ , _ , and Xianwen Shi, "Contests for status," *Journal of Political Economy*, 2007, *115* (2), 338–363.

Möller, Marc, "Incentives versus competitive balance," *Economics Letters*, 2012, *117* (2), 505–508.

Rosen, Sherwin, "Prizes and incentives in elimination tournaments," *American Economic Review*, 1986, *76* (4), 701–715.

Runkel, Marco, "Optimal contest design, closeness and the contest success function," *Public Choice*, 2006, *129* (1-2), 217–231.

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

Serena, Marco, "Quality contests," *European Journal of Political Economy*, 2017, *46*, 15–25.

Siegel, Ron, "All-pay contests," *Econometrica*, 2009, *77* (1), 71–92.

_ , "Asymmetric contests with head starts and nonmonotonic costs," *American Economic Journal: Microeconomics*, 2014, *6* (3), 59–105.

Szidarovszky, Ferenc and Koji Okuguchi, "On the existence and uniqueness of pure Nash equilibrium in rent-seeking games," *Games and Economic Behavior*, 1997, *18* (1), 135–140.

Szymanski, Stefan, "The economic design of sporting contests," *Journal of Economic Literature*, 2003, *41* (4), 1137–1187.

Appendix A Microfoundation

We interpret the microeconomic substance of the generalized lottery contest model from two perspectives.

Noisy Ranking Clark and Riis (1996) and Jia (2008) show that a generalized lottery contest is underpinned by a unique noisy ranking system. Imagine that contestants are evaluated through a set of noisy signals of their performance ℓ_is. Following the discrete choice framework of McFadden (1973, 1974),[25] the noisy signal ℓ_i is assumed to be described by

$$\log \ell_i = \log f_i(x_i) + \varepsilon_i, \forall\, i \in \mathcal{N},$$

where the deterministic and strictly increasing production function $f_i(\cdot)$: $\mathbb{R}_+ \to \mathbb{R}_+$ measures the output of contestant i's effort x_i,[26] and the additive noise term ε_i reflects the randomness in the production process or the imperfection of the measurement and evaluation process. Idiosyncratic noises $\boldsymbol{\varepsilon} := \{\varepsilon_i, i \in \mathcal{N}\}$ are independently and identically distributed, being drawn from a type I extreme-value (maximum) distribution, with a cumulative distribution function

$$G(\varepsilon_i) = e^{-e^{-\varepsilon_i}}, \varepsilon_i \in (-\infty, +\infty), \forall\, i \in \mathcal{N}.$$

A contestant i prevails if he outperforms all others: This noisy-ranking tournament boils down to a generalized lottery contest, because

$$\Pr\left(\ell_i > \max_{j \neq i} \ell_j\right) = \frac{f_i(x_i)}{\sum_{j=1}^{n} f_j(x_j)}.$$

Isomorphism to R&D Contests Baye and Hoppe (2003) demonstrate the isomorphism between a generalized lottery contest, the research tournament model proposed by Fullerton and McAfee (1999), and the patent race model suggested by Loury (1979) and Dasgupta and Stiglitz (1980). This provides a more intuitive microeconomic underpinning for the model.

To illustrate the equivalence, we focus on the research tournament model of Fullerton and McAfee (1999). A sponsor—who is interested in an innovative technology—invites $n \geq 2$ R&D firms to carry out the project. Firms develop the technology and submit their products to the designer. The entry of the highest quality wins and its developer is awarded a prize, such as a procurement contract. Each firm i's valuation of the prize is given by $v_i > 0$.

Each firm i decides on its own input $x_i \geq 0$ in developing the technology. The quality

[25] The framework of McFadden's discrete choice model is further introduced and studied in various respects by works collected in Manski and McFadden (1981).

[26] Define $\log f_i(x_i) = -\infty$ if $f_i(x_i) = 0$.

q_i of firm i's product is randomly drawn from a distribution with cumulative distribution function $[\Gamma(q_i)]^{f_i(x_i)}$. The function $\Gamma(\cdot)$ is a continuous cumulative distribution function on a support $[\underline{q}, \overline{q}]$, with $\overline{q} > \underline{q}$. By Fullerton and McAfee (1999) and Baye and Hoppe (2003), the term $f_i(x_i)$—which increases with x_i—can intuitively be interpreted as the number of research ideas generated in developing the product and indicates the firm's research capacity: Each research idea allows the firm to produce a prototype, with its quality being drawn from the distribution function $\Gamma(\cdot)$. A firm simply presents its best prototype to the sponsor as its entry, and the quality of its entry thus follows the distribution function $[\Gamma(q_i)]^{f_i(x_i)}$: The more ideas a firm generates, the more likely a higher q_i can be realized, and the more likely the firm can leapfrog its competitors. As pointed out by Baye and Hoppe (2003) and Fu and Lu (2012), a firm i wins the prize with a probability

$$\Pr\left(q_i > \max_{j \neq i} q_j\right) = \frac{f_i(x_i)}{\sum_{j=1}^{n} f_j(x_j)}.$$

A similar equivalence can be established between a generalized lottery contest model and the "first past the post" patent race model of Loury (1979) and Dasgupta and Stiglitz (1980), in which a firm secures a rent if it makes a scientific discovery earlier than its competitors. Fu and Lu (2012) further reveal the underlying statistical linkage between these R&D contests and the generalized lottery contest model (1).

Appendix B Proofs

Proof of Theorem 1

Proof. Note that $x_i = 0$ is a strictly dominant strategy for contestant i if $f_i(\cdot)$ is a constant. Therefore, it suffices to prove the theorem for the case in which $f_i(\cdot)$ satisfies $f_i'(x_i) > 0$, $f_i''(x_i) \leq 0$ and $f_i(0) \geq 0$ for all $i \in \mathcal{N}$.

For notational convenience, define $y_i := f_i(x_i)$, $\delta_i := f_i(0)$, $\widetilde{f}_i(x_i) := f_i(x_i) - \delta_i$, and $\lambda_i(y_i) := c\big(\widetilde{f}_i^{-1}(y_i - \delta_i)\big)/v_i$. It follows immediately that $c(x_i) = \lambda_i(y_i) \cdot v_i$. Moreover, we have that $\lambda_i' > 0$ and $\lambda_i'' \geq 0$. The expected payoff of contestant $i \in \mathcal{N}$ choosing $y_i \geq \delta_i$ is equal to

$$\left[\frac{y_i}{\sum_{j=1}^n y_j} - \lambda_i(y_i)\right] \cdot v_i.$$

It remains to show that there exists a unique equilibrium $\boldsymbol{y}^* \equiv (y_1^*, \ldots, y_n^*)$ that satisfies $y_i^* \geq \delta_i$ for all $i \in \mathcal{N}$. Let $s := \sum_{j=1}^n y_j$ and $\underline{\delta} := \sum_{j=1}^n \delta_j$. It is clear that $s \geq \underline{\delta}$. The first-order condition of the above expected utility with respect to y_i yields the following:

$$\frac{s - y_i}{s^2} - \lambda_i'(y_i) \leq 0, \text{ with equality if } y_i > \delta_i.$$

Fixing s, let us define $y_i(s)$ as the following:

$$y_i(s) := \begin{cases} \delta_i & \text{if } s^2\lambda_i'(\delta_i) - s + \delta_i \geq 0, \\ \text{The unique solution to } s - y_i = s^2\lambda_i'(y_i) & \text{otherwise.} \end{cases} \tag{13}$$

It is straightforward to verify that $y_i(s)$ is well defined and continuous in $s \in [\delta_i, \infty]$. Moreover, we must have that $y_i(s) \in (\delta_i, s)$ if $s^2\lambda_i'(\delta_i) - s + \delta_i < 0$.

Suppose that there exists an interval of s such that $y_i(s) > \delta_i$. It follows immediately from the implicit function theorem that

$$y_i'(s) = \frac{1 - 2s\lambda_i'(y_i)}{1 + s^2\lambda_i''(y_i)} = \frac{2y_i(s) - s}{\left[1 + s^2\lambda_i''(y_i)\right] s}, \tag{14}$$

where the second equality follows from $s - y_i = s^2\lambda_i'(y_i)$. Therefore, $y_i(s)$ is strictly decreasing in this interval if $2y_i < s$ and strictly increasing otherwise. By Equation (13), the latter case occurs if and only if

$$s - \frac{1}{2}s > s^2\lambda_i'\left(\frac{s}{2}\right) \Leftrightarrow 2s\lambda_i'\left(\frac{s}{2}\right) < 1.$$

Note that $2s\lambda_i'\left(\frac{s}{2}\right)$ is strictly increasing in s, which implies that there exists at most one solution to $2s\lambda_i'\left(\frac{s}{2}\right) = 1$. Denote the solution by $\hat{s}_i$ whenever it exists.

Next, we denote the two different real number solutions of $s^2\lambda_i'(\delta_i) - s + \delta_i = 0$ by $s_i^{\dagger}$ and $s_i^{\dagger\dagger}$ respectively, with $s_i^{\dagger} < s_i^{\dagger\dagger}$, whenever they exist. The above analysis, together with the fact that the expression $s^2\lambda_i'(\delta_i) - s + \delta_i$ in Equation (13) is quadratic in s, implies that the function $y_i(s)$ must fall into one of the following four cases:

Case I: There exist no different real number solutions of $s^2\lambda_i'(\delta_i) - s + \delta_i = 0$ for $s \in [\underline{\delta}, \infty]$. Then we must have that $s^2\lambda_i'(\delta_i) - s + \delta_i \geq 0$ for all $s \geq \underline{\delta}$, which in turn implies that $y_i(s) = \delta_i$ for all $s \geq \underline{\delta}$ by Equation (13). To slightly abuse the notation, we let $s_i^{\dagger\dagger} := \underline{\delta}$ for this case.

Case II: $s_i^{\dagger} \leq \underline{\delta} \leq s_i^{\dagger\dagger}$ and $y_i(\underline{\delta}) \leq \frac{1}{2}\underline{\delta}$. Then $y_i(s)$ is strictly decreasing in s for $s \in [\underline{\delta}, s_i^{\dagger\dagger}]$, and $y_i(s) = \delta_i$ for $s \in [s_i^{\dagger\dagger}, \infty]$.

Case III: $s_i^{\dagger} \leq \underline{\delta} \leq s_i^{\dagger\dagger}$ and $y_i(\underline{\delta}) > \frac{1}{2}\underline{\delta}$. It can be verified that $\underline{\delta} < \hat{s}_i < s_i^{\dagger\dagger}$. Therefore, $y_i(s)$ is strictly increasing in s for $s \in [\underline{\delta}, \hat{s}_i]$; is strictly decreasing in s for $s \in [\hat{s}_i, s_i^{\dagger\dagger}]$; and $y_i(s) = \delta_i$ for $s \in [s_i^{\dagger\dagger}, \infty]$.

Case IV: $\underline{\delta} < s_i^{\dagger} < s_i^{\dagger\dagger}$. It can be verified that $s_i^{\dagger} < \hat{s}_i < s_i^{\dagger\dagger}$. Moreover, $y_i(s)$ is strictly increasing in s for $s \in [s_i^{\dagger}, \hat{s}_i]$; is strictly decreasing in s for $s \in [\hat{s}_i, s_i^{\dagger\dagger}]$; and $y_i(s) = \delta_i$ for $s \in [\underline{\delta}, s_i^{\dagger}] \cup [s_i^{\dagger\dagger}, \infty]$.

The four cases are depicted in Figure 3 graphically. For Case I and Case II, we define $s_i := \underline{\delta}$; for Case III and Case IV, we define $s_i := \hat{s}_i \geq \underline{\delta}$. It is straightforward to verify that $y_i(s) > \frac{1}{2}s$ holds if $s < s_i$ for all four cases. Without loss of generality, we order the contestants such that

$$s_1 \geq s_2 \geq \ldots \geq s_n \geq \underline{\delta}.$$

Define $Y(s) := \sum_{i=1}^{n} y_i(s) - s$. It remains to show that $Y(s) = 0$ has a unique positive solution. First, note that no solution exists for $s < s_2$, because

$$Y(s) := \sum_{i=1}^{n} y_i(s) - s \geq y_1(s) + y_2(s) - s > \frac{1}{2}s + \frac{1}{2}s - s = 0, \text{ for } s < s_2.$$

Next, we claim that $Y(s)$ is strictly decreasing in s for $s \geq s_2$. Clearly, $Y(s)$ is strictly decreasing in s for $s \geq s_1$. Moreover, for $s \in [s_2, s_1]$, $Y(s)$ can be rewritten as

$$Y(s) = \underbrace{\sum_{i=2}^{n} y_i(s)}_{\text{first term}} + \underbrace{[y_1(s) - s]}_{\text{second term}}.$$

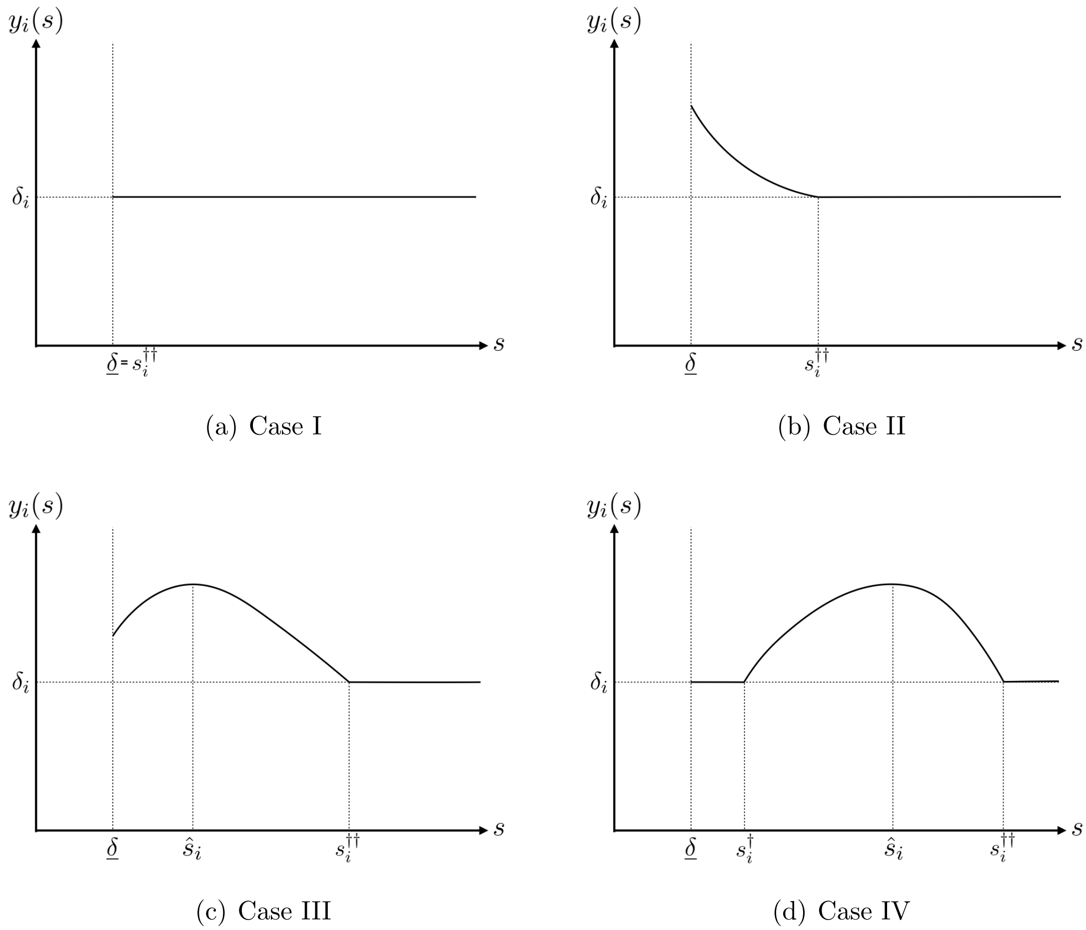


Figure 3: $y_i(s)$.

Because $s \geq s_2 \geq \ldots \geq s_n$, the first term is weakly decreasing in s. Taking the derivative of the second term with respect to s yields

$$y_1'(s) - 1 = \frac{2y_1(s) - s}{\left[1 + s^2 \lambda_1''\left(y_1(s)\right)\right] s} - 1 \leq \frac{2y_1(s) - s}{s} - 1 = \frac{2}{s}\left[y_1(s) - s\right] < 0,$$

where the first equality follows from Equation (14); the first inequality follows from $\lambda_i'' \geq 0$ and $y_1(s) \geq \frac{s}{2}$, and the second inequality follows from $y_i(s) < s$ [see Equation (13)]. Therefore, the second term is strictly decreasing in s, which in turn implies that $Y(s)$ is strictly decreasing for $s \in [s_2, \infty]$.

It is straightforward to see that for all four cases, we have that $y_i(s) = \delta_i$ for $s \geq s_i^{\dagger\dagger}$. Let

$s^{\dagger\dagger} := s_2 + \sum_{i=1}^n s_i^{\dagger\dagger} + \sum_{i=1}^n \delta_i$. It is clear that $s^{\dagger\dagger} \geq s_2$. Moreover, we have that

$$Y(s^{\dagger\dagger}) = \sum_{i=1}^n y_i(s^{\dagger\dagger}) - s^{\dagger\dagger} = \sum_{i=1}^n \delta_i - \left(s_2 + \sum_{i=1}^n s_i^{\dagger\dagger} + \sum_{i=1}^n \delta_i \right) = -s_2 - \sum_{i=1}^n s_i^{\dagger\dagger} \leq 0.$$

Therefore, there exists a unique positive solution to $Y(s) = 0$ for $s \in [s_2, s^{\dagger\dagger}]$. This completes the proof. ■

Proof of Theorem 2

Proof. The analysis for the case $x_t^* > 0$ is provided in the main text, and it suffices to prove the theorem for the case $x_t^* = 0$. Because $\beta_t^* > 0$, we must have $p_t^* > 0$. If $p_t^* = 1$, then we must have $\boldsymbol{x}^* = \mathbf{0}$. Clearly, the equilibrium outcome (i.e., $\boldsymbol{x}^*$ and $\boldsymbol{p}^*$) can be replicated by the following contest rule with zero headstarts:

$$(\alpha_i, \beta_i) := \begin{cases} (1, 0) & \text{for} \quad i = t, \\ (0, 0) & \text{for} \quad i \neq t. \end{cases}$$

Therefore, it remains to focus on the case in which $p_t^* \in (0, 1)$. Denote by $x^{\dagger\dagger}$ the unique solution to the following equation:

$$p_t^*(1 - p_t^*)v_t = c'(x^{\dagger\dagger}) \cdot \frac{h(x^{\dagger\dagger})}{h'(x^{\dagger\dagger})}.$$

Note that the left-hand side of the above equation is strictly positive. Therefore, $x^{\dagger\dagger} > 0 = x_t^*$. Consider the following contest rule with weights $\widehat{\boldsymbol{\alpha}} \equiv (\widehat{\alpha}_1, \ldots, \widehat{\alpha}_n)$ and headstarts $\widehat{\boldsymbol{\beta}} \equiv (\widehat{\beta}_1, \ldots, \widehat{\beta}_n)$ such that

$$\left(\widehat{\alpha}_i, \widehat{\beta}_i\right) := \begin{cases} \left(\frac{\alpha_t^* h(x_t^*) + \beta_t^*}{h(x^{\dagger\dagger})}, 0\right) & \text{for} \quad i = t, \\ (\alpha_i^*, \beta_i^*) & \text{for} \quad i \neq t. \end{cases}$$

Denote the equilibrium effort profile and winning probabilities under the alternative contest rule $(\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$ by $\widehat{\boldsymbol{x}}^* \equiv (\widehat{x}_1^*, \ldots, \widehat{x}_n^*)$ and $\widehat{\boldsymbol{p}}^* \equiv (\widehat{p}_1^*, \ldots, \widehat{p}_n^*)$, respectively. It can be verified that

$$\widehat{x}_i^* = \begin{cases} x^{\dagger\dagger} & \text{for} \quad i = t, \\ x_i^* & \text{for} \quad i \neq t. \end{cases}$$

Moreover, we have that $\widehat{p}_i^* = p_i^*$ for all $i \in \mathcal{N}$ because $\widehat{\alpha}_t \cdot h(x^{\dagger\dagger}) + \widehat{\beta}_t = \alpha_t^* \cdot h(x_t^*) + \beta_t^*$ by construction. Therefore, the contest designer's payoff under $(\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$ is weakly higher than that under $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ by Assumption 2. This completes the proof. ■

Proof of Theorem 3

Proof. Part (i) of the theorem is trivial, and it remains to show part (ii). It is clear that $x_i = 0$ is a strictly dominant strategy if $\alpha_i = 0$. For $(p_i, p_j) > (0, 0)$, we must have $(x_i, x_j) > (0, 0)$. Therefore, the following first-order conditions must be satisfied by Equation (9):

$$\begin{aligned} x_i &= g\left(\log(p_i\left(1-p_i\right)) + \log(v_i)\right), \\ x_j &= g\left(\log(p_j\left(1-p_j\right)) + \log(v_j)\right). \end{aligned}$$

Note that Equation (1) implies that

$$\frac{p_i}{p_j} = \frac{\frac{\alpha_i \cdot h(x_i)}{\sum_{k=1}^n \alpha_k \cdot h(x_k)}}{\frac{\alpha_j \cdot h(x_j)}{\sum_{k=1}^n \alpha_k \cdot h(x_k)}} = \frac{\alpha_i \cdot h(x_i)}{\alpha_j \cdot h(x_j)}.$$

Combining the above conditions, we can obtain that

$$\frac{\alpha_i}{\alpha_j} = \frac{p_i/h(x_i)}{p_j/h(x_j)} = \frac{\frac{p_i}{h\left(g\left(\log(p_i\left(1-p_i\right))+\log(v_i)\right)\right)}}{\frac{p_j}{h\left(g\left(\log(p_j\left(1-p_j\right))+\log(v_j)\right)\right)}}.$$

The last equation clearly holds for the set of weights specified in Equation (11). This completes the proof. ■

Proof of Theorem 4

Proof. With slight abuse of notation, let us define $x(p_k, v_k) := g\left(\log(p_k\left(1-p_k\right)) + \log(v_k)\right)$. Then the equilibrium effort x_k in Equation (9) can be written as $x(p_k, v_k)$ for all $k \in \mathcal{N}$. Define $\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{v}) := \big(x(p_1, v_1), \ldots, x(p_n, v_n)\big)$. It follows immediately that $\tau\left(\boldsymbol{x}(\boldsymbol{p}, \boldsymbol{v})\right) = \boldsymbol{x}\left(\tau(\boldsymbol{p}), \tau(\boldsymbol{v})\right)$. Moreover, Equation (9) implies that $x(0, v) = 0$ for all $v > 0$.

Suppose, to the contrary, that there exists some contestant $j \in \mathcal{N}$ with $v_j < v_i$ such that $p_i^* = 0 < p_j^*$. Then we can obtain

$$\begin{aligned} \Lambda\left(\boldsymbol{x}(\boldsymbol{p}^*, \boldsymbol{v}), \boldsymbol{p}^*, \boldsymbol{v}\right) &\leq \Lambda\left(\boldsymbol{x}(\boldsymbol{p}^*, \boldsymbol{v}), \boldsymbol{p}^*, \tau_{ij}(\boldsymbol{v})\right) \\ &= \Lambda\left(\tau_{ij}\left(\boldsymbol{x}(\boldsymbol{p}^*, \boldsymbol{v})\right), \tau_{ij}(\boldsymbol{p}^*), \boldsymbol{v}\right) \\ &= \Lambda\left(\boldsymbol{x}\left(\tau_{ij}(\boldsymbol{p}^*), \tau_{ij}(\boldsymbol{v})\right), \tau_{ij}(\boldsymbol{p}^*), \boldsymbol{v}\right) \\ &< \Lambda\left(\boldsymbol{x}\left(\tau_{ij}(\boldsymbol{p}^*), \boldsymbol{v}\right), \tau_{ij}(\boldsymbol{p}^*), \boldsymbol{v}\right). \end{aligned}$$

The first inequality follows from $x(p_i^*, v_i) = 0$ and part (ii) of Assumption 3; the first equality follows from part (i) of Assumption 3 and the fact that $\tau_{ij}\left(\tau_{ij}(\boldsymbol{v})\right) = \boldsymbol{v}$; the second equality follows from $\tau_{ij}\left(\boldsymbol{x}(\boldsymbol{p}^*, \boldsymbol{v})\right) = \boldsymbol{x}\left(\tau_{ij}(\boldsymbol{p}^*), \tau_{ij}(\boldsymbol{v})\right)$; and the last strict inequality follows from $x(p_i^*, v_i) = x(p_i^*, v_j) = 0$, $x(p_j^*, v_j) < x(p_j^*, v_i)$, the postulated $p_j^* > 0$, and part (iii) of Assumption 3. Therefore, the contest designer's payoff under the optimal vector of winning probabilities $\boldsymbol{p}^*$ is strictly lower than that under $\tau_{ij}(\boldsymbol{p}^*)$, which is a contradiction. This completes the proof. ■

Proof of Proposition 1

Proof. It is obvious that $p_1^* = p_2^* = \frac{1}{2}$ from Equation (9) when $n = 2$, and it remains to prove the result for the case $n \geq 3$. We first prove part (i) of the proposition. Suppose, to the contrary, that only two players remain active in the optimal contest. It is clear that $p_1^* = p_2^* = \frac{1}{2}$ in the optimum. Consider the following profile of equilibrium winning probabilities $\boldsymbol{p} = (\frac{1}{2}, \frac{1}{2} - \epsilon, \epsilon, 0, \ldots, 0)$. It can be verified that the total effort under $\boldsymbol{p}$ is equal to

$$\Lambda(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}) = g\left(\log(\frac{1}{4}) + \log(v_1)\right) + g\left(\log(\frac{1}{4} - \epsilon^2) + \log(v_2)\right) + g\left(\left(\log(\epsilon(1-\epsilon)\right) + \log(v_3)\right).$$

Simple algebra shows that $\partial\Lambda/\partial\epsilon > 0$ when ϵ is sufficiently small. Therefore, at least three players will remain active in the optimum.

Next, we prove part (ii). Suppose, to the contrary, that $p_i^* \geq \frac{1}{2}$ for some $i \in \mathcal{N}$. If $p_i^* > \frac{1}{2}$, then the contest designer can assign probability $1 - p_i^*$ to contestant i and probability $p_j^* + (2p_i^* - 1)$ to an arbitrary contestant $j \neq i$. Because at least three players remain active in the optimum, we must have $p_i^* + p_j^* < 1$. This in turn implies that $|p_j^* + (2p_i^* - 1) - \frac{1}{2}| < |p_j^* - \frac{1}{2}|$, and thus contestant j's effort strictly increases. Furthermore, it follows from Equation (9) that contestant i's effort remains the same. Therefore, the total effort strictly increases after the adjustment. If $p_i^* = \frac{1}{2}$, then there exists an active player $j \in \mathcal{N}$ such that $p_j \in (0, \frac{1}{2})$, because at least three players remain active in the optimum. In such a scenario, the designer can increase the total effort by reducing p_i^* by a sufficiently small amount and increasing p_j^* by the same amount. This completes the proof. ■

Proof of Proposition 2

Proof. It is useful to first prove the following intermediate result.

Lemma 1 *Consider a contest with three players who are indexed by i, j, and k. Suppose that the contest designer aims to maximize the expected winner's effort. Then setting $p_i = p_j = p_k = \frac{1}{3}$ is suboptimal.*

Proof. Without loss of generality, we assume that $v_i \geq v_j \geq v_k$. The difference between the expected winner's effort under $(p_i, p_j, p_k) = (\frac{1}{2}, \frac{1}{2}, 0)$ and that under $(p_i, p_j, p_k) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ can be derived as

$$\begin{aligned}
&\left[\frac{1}{2}g\left(\log\left(\frac{1}{4}\right)+\log(v_i)\right)+\frac{1}{2}g\left(\log\left(\frac{1}{4}\right)+\log(v_j)\right)\right] \\
&\quad-\left[\frac{1}{3}g\left(\log\left(\frac{2}{9}\right)+\log(v_i)\right)+\frac{1}{3}g\left(\log\left(\frac{2}{9}\right)+\log(v_j)\right)+\frac{1}{3}g\left(\log\left(\frac{2}{9}\right)+\log(v_k)\right)\right] \\
&>\frac{1}{6}\left[g\left(\log\left(\frac{2}{9}\right)+\log(v_i)\right)-g\left(\log\left(\frac{2}{9}\right)+\log(v_j)\right)\right] \\
&\geq 0,
\end{aligned}$$

where the strict inequality follows from $\frac{1}{4} > \frac{2}{9}$, $v_j \geq v_k$, and the monotonicity of $g(\cdot)$. Therefore, setting $p_i = p_j = p_k = \frac{1}{3}$ is suboptimal. This completes the proof. ■

Now we can prove the proposition. Suppose, to the contrary, that three or more players remain active in the optimal contest. Then there exist $i, j, k \in \mathcal{N}$ such that $p_i^{**} \geq p_j^{**} > 0$ and $p_i^{**} \geq p_k^{**} > 0$. Lemma 1 implies that $\min\{2p_j^{**} + p_k^{**}, p_j^{**} + 2p_k^{**}\} < 1$. Without loss of generality, we assume that $v_j \geq v_k$.

Suppose that the contest designer assigns probability $p_{jk}^{**} := p_j^{**} + p_k^{**}$ to player j and 0 to player k, and does not change the equilibrium winning probability of all other players. Then the difference between the expected winner's effort under the new profile of winning probabilities and that under $\boldsymbol{p}^{**} \equiv (p_1^{**}, \ldots, p_n^{**})$ can be derived as

$$\begin{aligned}
&(p_j^{**}+p_k^{**})g\left(\log\left(p_{jk}^{**}(1-p_{jk}^{**})\right)+\log(v_j)\right) \\
&\quad-\left[p_j^{**}g\left(\log\left(p_j^{**}(1-p_j^{**})\right)+\log(v_j)\right)+p_k^{**}g\left(\log\left(p_k^{**}(1-p_k^{**})\right)+\log(v_k)\right)\right] \\
&=p_j^{**}\left[g\left(\log\left(p_{jk}^{**}(1-p_{jk}^{**})\right)+\log(v_j)\right)-g\left(\log\left(p_j^{**}(1-p_j^{**})\right)+\log(v_j)\right)\right] \\
&\quad+p_k^{**}\left[g\left(\log\left(p_{jk}^{**}(1-p_{jk}^{**})\right)+\log(v_j)\right)-g\left(\log\left(p_k^{**}(1-p_k^{**})\right)+\log(v_k)\right)\right] \\
&>0,
\end{aligned}$$

where the inequality follows from $\min\{2p_j^{**} + p_k^{**}, p_j^{**} + 2p_k^{**}\} < 1$, $v_j \geq v_k$, and the monotonicity of $g(\cdot)$. A contradiction. Therefore, only two contestants would remain active in the optimal contest. Moreover, they must be the two ex ante strongest players by Theorem 4.

It remains to show that the ex ante stronger player always wins with a strictly higher probability than the underdog. Suppose, to the contrary, that $v_1 > v_2$ and $0 < p_1^{**} \le p_2^{**}$, with $p_1^{**} + p_2^{**} = 1$. We consider the following two cases:

Case I: $p_1^{**} < p_2^{**}$. Then the designer can increase the expected winner's effort by assigning probability p_1^{**} to player 2 and p_2^{**} to player 1. This would lead to a change in the expected winner's effort that amounts to

$$\begin{aligned}&\Big[p_1^{**} g\left(\log\left(p_1^{**}p_2^{**}\right) + \log(v_2)\right) + p_2^{**} g\left(\log\left(p_1^{**}p_2^{**}\right) + \log(v_1)\right)\Big]\\ &\quad - \Big[p_1^{**} g\left(\log\left(p_1^{**}p_2^{**}\right) + \log(v_1)\right) + p_2^{**} g\left(\log\left(p_1^{**}p_2^{**}\right) + \log(v_2)\right)\Big]\\ &= \left(p_2^{**} - p_1^{**}\right)\Big[g\left(\log\left(p_1^{**}p_2^{**}\right) + \log(v_1)\right) - g\left(\log\left(p_1^{**}p_2^{**}\right) + \log(v_2)\right)\Big] > 0,\end{aligned}$$

which is a contradiction.

Case II: $p_1^{**} = p_2^{**} = \frac{1}{2}$. Let the designer assign winning probability $\frac{1}{2} + \epsilon$ to player 1 and $\frac{1}{2} - \epsilon$ to player 2. The adjustment leads to a change in the expected winner's effort that amounts to

$$\begin{aligned}\Xi(\epsilon) :=& \left[\left(\frac{1}{2} + \epsilon\right) g\left(\log\left(\frac{1}{4} - \epsilon^2\right) + \log(v_1)\right) + \left(\frac{1}{2} - \epsilon\right) g\left(\log\left(\frac{1}{4} - \epsilon^2\right) + \log(v_2)\right)\right]\\ &- \frac{1}{2}\left[g\left(\log\left(\frac{1}{4}\right) + \log(v_1)\right) + g\left(\log\left(\frac{1}{4}\right) + \log(v_2)\right)\right].\end{aligned}$$

It is straightforward to verify that $\Xi(0) = 0$ and $\Xi'(0) = g\left(\log\left(\frac{v_1}{4}\right)\right) - g\left(\log\left(\frac{v_2}{4}\right)\right) > 0$. Therefore, $\Xi(\epsilon) > 0$ for sufficiently small $\epsilon > 0$, which is again a contradiction. This completes the proof. ■

Proof of Proposition 3

Proof. Recall that Proposition 1 states that $p_i^*, p_j^* < \frac{1}{2}, \forall\, i, j \in \mathcal{N}$. Suppose, to the contrary, that $v_i > v_j$ and $p_i^* \le p_j^*$. We consider the following two cases:

Case I: $p_i^* < p_j^*$. Let the contest designer assign probability p_j^* to player i and p_i^* to player j, and not change the equilibrium winning probability of all other players. Define $\Omega_{k_1k_2} := \log(p_{k_1}^*\left(1 - p_{k_1}^*\right)) + \log(v_{k_2})$ for $k_1, k_2 \in \{i, j\}$. It can be verified that $\Omega_{ii}, \Omega_{jj} \in (\Omega_{ij}, \Omega_{ji})$

and $\Omega_{ii} + \Omega_{jj} = \Omega_{ij} + \Omega_{ji}$. Furthermore, the difference between the total effort under the alternative profile of winning probabilities and that under $\boldsymbol{p}^* \equiv (p_1^*, \dots, p_n^*)$ is equal to

$$\Big[g\left(\Omega_{ij}\right) + g\left(\Omega_{ji}\right)\Big] - \Big[g\left(\Omega_{ii}\right) + g\left(\Omega_{jj}\right)\Big] > 0,$$

where the strict inequality follows from $\Omega_{ii}, \Omega_{jj} \in (\Omega_{ij}, \Omega_{ji})$, $\Omega_{ii} + \Omega_{jj} = \Omega_{ij} + \Omega_{ji}$, and the strict convexity of $g(\cdot)$. A contradiction.

Case II: $p_i^* = p_j^*$. Let the contest designer assign probability $p_i^* + \epsilon$ to player i and $p_j^* - \epsilon$ to player j, and not change the equilibrium winning probability of all other players. It can be verified that such adjustment generates strictly more total effort to the designer for a sufficiently small $\epsilon > 0$. This completes the proof. ∎

Proof of Proposition 4

Proof. The proof follows from Theorems 3 and 4, and the fact that the total effort $r \sum_{i=1}^n p_i(1-p_i)v_i$ is quadratic in p_i for all $i \in \mathcal{N}$. It is omitted for brevity. ∎

Proof of Proposition 5

Proof. Part (ii) of the proposition follows directly from part (i), and it suffices to prove part (i). With slight abuse of notation, we add r into α_i and α_j to emphasize the fact that the optimal weights $\boldsymbol{\alpha}^* \equiv (\alpha_1^*, \dots, \alpha_n^*)$ depend on the bidding efficiency r. Note that $\boldsymbol{p}^* \equiv (p_1^*, \dots, p_n^*)$ and κ are independent of r by Proposition 4. Moreover, we have that

$$\mathcal{T}(r) := \log\left(\frac{\alpha_i^*(r)}{\alpha_j^*(r)}\right) = (1-r)\log\left(\frac{p_i^*}{p_j^*}\right) - r\log\left(\frac{1-p_i^*}{1-p_j^*}\right) - r\log\left(\frac{v_i}{v_j}\right).$$

Clearly, $\mathcal{T}(r)$ is linear in r, and $\mathcal{T}(r) \gtreqless 0$ is equivalent to $\alpha_i^*(r) \gtreqless \alpha_j^*(r)$. Note that

$$\lim_{r \searrow 0} \mathcal{T}(r) = \log\left(\frac{p_i^*}{p_j^*}\right) > 0,$$

and

$$\mathcal{T}(1) = -\log\left(\frac{1-p_i^*}{1-p_j^*} \times \frac{v_i}{v_j}\right) = -\log\left(\frac{v_i + \frac{\kappa-2}{\sum_{s=1}^{\kappa}\frac{1}{v_s}}}{v_j + \frac{\kappa-2}{\sum_{s=1}^{\kappa}\frac{1}{v_s}}}\right) < 0,$$

where the second equality follows from Equation (12). Therefore, there exists a unique cutoff $\overline{r}_{ij} \in (0,1)$ such that $\alpha_i^*(r) \gtreqless \alpha_j^*(r)$ if $r \lesseqgtr \overline{r}_{ij}$. This completes the proof. ∎