Optimal Contests with Incomplete Information and Convex Effort Costs

Mengxi Zhang*

March 21, 2023

Abstract

I investigate the design of effort-maximizing mechanisms when agents have both private information and convex effort costs, and the designer has a fixed prize budget. I first demonstrate that it is always optimal for the designer to utilize a contest with as many participants as possible. Further, I identify a necessary and sufficient condition for the winner-takes-all prize structure to be optimal. When this condition fails, the designer may prefer to award multiple prizes of descending sizes. I also provide a characterization of the optimal prize allocation rule for this case. Finally, I illustrate how the optimal prize distribution evolves as the contest size grows.

*I thank the editor and three anonymous reviewers for their constructive comments, which helped me to improve the manuscript. I am indebted to Benny Moldovanu for many insightful discussions. I am also grateful to Shuo Liu, Thomas Tröger, Cédric Wasser, Zenan Wu, Juuso Välimäki, and seminar/conference participants at UC3M, Aalto University, Durham University, University of Mannheim, Peking University, Hunan University, Global Seminar on Contests & Conflict, 2022 Stony Brook Conference, 7th Annual Conference on Contests: Theory and Evidence, 2020 European Winter Meeting, and 2019 Nanjing International Conference on Theory for helpful conversations and comments. Financial support from the German Research Foundation(DFG) via Germany's Excellence Strategy - EXC 2047/1 - 390685813 and CRC TR-224 (project B01) are gratefully acknowledged. Contact: Department of Economics, University of Bonn, mzhang@uni-bonn.de.

1 Introduction

Contests, in which participants/agents exert costly and irreversible effort to compete for a set of prizes, are widely used in practice to improve participants' performance (e.g., sales contests, innovation contests, and sport contests). The ubiquity of contests naturally raises two questions: First, how can the contest's structure and rules be optimally designed to best motivate participants? Second, is it possible to further improve participants' performance by adopting mechanisms other than contests? In this paper, I address these two questions for the arguably most relevant case, in which participants are privately informed about their abilities or valuations of the prize and have convex effort costs. Following the tradition in the contest literature (see Konrad (2009) for a survey), I assume that the designer has a fixed prize budget and that her objective is to maximize the expected total effort.

To answer the second question, I first characterize the set of all implementable, i.e., incentive compatible and individually rational, mechanisms by establishing a payoff equivalence result that maps agents' utilities to the expected share of prize allocated to them (this is also known as the *interim prize allocation rule*). I then show that the set of effort-maximizing mechanisms always includes a contest, i.e. a mechanism in which the designer defines a set of prizes and asks the agents to invest non-refundable effort before the allocation of the prizes are determined. In other words, there does not exist any implementable mechanism that strictly outperforms contests from the designer's perspective. This finding provides a potential rationale for the prevalence of contests.

Turning to the first question, the contest design problem is twofold: The designer first needs to choose how to structure the contest. That is, she needs to decide if she wants to limit entry or to admit all potential participants, and if she wants to host a single contest with all contestants or to organize several simultaneous/sequential subcontests. Then the designer needs to determine how to allocate the prize budget. I find that it is always optimal for the designer to host a single contest with as many contestants as possible. Previous work has found that the designer may benefit by altering the contest structure. This is mainly due to an additional restriction that the designer cannot optimally choose the prize allocation rule (e.g., Moldovanu and Sela, 2006; Fu and Lu, 2012; Fang et al., 2020). Regarding the prize allocation, I show that if a regularity condition holds, then the designer finds it optimal to always allocate the entire prize budget to the contestant who exerts the highest effort (this is often referred to as *winner-takes-all*). When this condition fails, the designer may prefer to award multiple prizes of descending sizes: The contestant who exerts the highest effort wins the largest prize, the contestant who exerts the second-highest effort wins the second-largest prize, and so on until all of the prizes are allocated. I also characterize the optimal prize allocation rule for this case.

I then construct a parameterized setting to illustrate how the optimal prize distribution evolves as the contest size grows. The findings suggest that, with convex effort costs, the optimal level of competition intensity, measured by the total effort costs borne by agents, is often less than the maximum possible level. As the contest size increases, the prize distribution first becomes more unequal until the optimal level of competition intensity is obtained, and then it becomes less unequal to maintain the optimal intensity. The optimal competition level decreases as the effort cost becomes more convex, and it increases as the contestants' ability distribution becomes more spread out.

The existing contest design literature mostly focuses on investigating how to improve the design of contests and largely ignores the comparison of contests with other non-contest mechanisms. To the best of my knowledge, this is the first paper that establishes the optimality of contests among all implementable mechanisms with respect to motivating agents to exert effort.[1]

The first paper that takes a comprehensive approach to studying contest design with incomplete information and convex effort costs is Moldovanu and Sela (2001).[2] The authors investigate the design of a single contest with symmetric contestants, and they allow the designer to choose either a single prize or multiple prizes of descending sizes. They find that when the effort cost is concave or linear, a single prize is always optimal, whereas when the effort cost is convex, multiple prizes may perform better than a single prize. They do not aim to provide a characterization of the optimal prize structure, which has since been considered to be a generally intractable problem.

Olszewski and Siegel (2016, 2020) have recently revisited this problem and show that it can be solved in the context of large contests. Olszewski and Siegel (2016) find that when the number of contestants goes to infinity, the n-agent contest design problem converges to a single agent problem. In Olszewski and Siegel (2020), the authors utilize this finding to identify the optimal prize structure for contests with "sufficiently many" (either symmetric or asymmetric) contestants. They find that, with convex effort costs, it is always optimal for the designer to award many prizes of descending sizes. In this paper, I employ a differ-

[1]Letina et al. (2020) find that contests are optimal within a large set of mechanisms. They consider a very different model in which agents do not have private information and contests are used to ensure that a lenient reviewer will have an incentive to punish shirking agents.

[2]Glazer and Hassin (1988) analyze a similar problem but they only provide a solution for the special case where the agents' types are uniformly distributed and have linear effort costs.

ent approach and demonstrate that when the environment is symmetric, the optimal prize allocation problem can be solved for any arbitrary number of contestants. The existence of a regularity condition under which the winner-takes-all prize structure is optimal suggests that restricting attention to large contests is not without loss.[3] On the other hand, my results offer a method to determine whether a specific contest should be considered "sufficiently large" and provides new insights about why awarding many prizes is optimal in large contests.

Both Moldovanu and Sela (2001) and Olszewski and Siegel (2020) take the contest format—a single contest—as given and only consider the choice of prize allocation rules. In this paper, I allow the designer to additionally choose the contest structure and establish the optimality of a single contest among all feasible contest structures.

Most papers that study contest design with incomplete information assume linear effort costs (e.g., Polishchuk and Tonis, 2013; Chawla et al., 2019; Liu et al., 2018). Regarding the technical differences, previous studies such as Polishchuk and Tonis (2013) have demonstrated that, with linear effort costs, the contest design problem is mathematically equivalent to an auction design problem with risk-neutral bidders. In contrast, assuming convex effort costs, I show that the contest problem can still be transformed into an auction design problem, but one with risk-averse bidders. Techniques for solving optimal auction problems with risk-neutral bidders are well established, thanks to the seminal works by Myerson (1981) and Riley and Samuelson (1983), and thus can be readily applied to the corresponding contest problems. However, with risk-averse bidders, the problem becomes much more complex, and explicit solutions are generally unknown. Optimal auction design with risk-averse bidders was first studied by Maskin and Riley (1984) and Matthews (1983). Maskin and Riley (1984) consider very general risk preferences and identify several properties of the revenue-maximizing mechanism, without attempting to obtain the explicit solution for any specific case. Matthews (1983) investigates the special case with constant absolute risk aversion and finds that a modified first-price auction is optimal. More recently, Gershkov et al. (2021) characterize the optimal mechanism for a risk-neutral seller who faces bidders equipped with non-expected utility preferences that exhibit constant risk aversion (i.e., both constant absolute risk aversion and constant relative risk aversion). They find that it is optimal for the seller to utilize a full insurance mechanism. Matthews (1983) and Gershkov et al. (2021) both consider preferences with a constant attitude towards risk, while the preferences studied

[3] When this condition fails, the optimal prize allocation rule in general still depends on the number of contestants.

in this paper do not fall into that category. The optimal mechanisms found by these two papers are also very different from that of the present paper. Other papers in the risk-averse bidder literature typically do not attempt to characterize the optimal mechanism and instead focus on comparing the performance of specific selling schemes (e.g., Matthews, 1987; Baisa, 2017).

In this paper, I first show that the search for an effort-maximizing mechanism can be confined to a special class of direct mechanisms, referred to as *all-pay mechanisms*, in which each agent's effort depends only on his type and is independent of either other agents' types or the prize allocation. Within this special class of mechanisms, the designer's objective function can be expressed as a function of the interim prize allocation rule. However, unlike in Myerson (1981), the objective function investigated here is not linear in the interim prize allocation. Thus, the objective function cannot be rewritten as a function of the ex-post prize allocation rule and then maximized pointwise, as is usually done for contests with linear effort costs. Instead, the maximization must be performed subject to the n-agent interim feasibility constraint. The problem is thus mathematically quite different. I demonstrate that, with symmetric agents, the problem can be tackled through a variational analysis approach, specifically, by utilizing the infinite-dimensional version of the Kuhn–Tucker Theorem (see Luenberger, 1997). This approach does not apply to the asymmetric case, as the interim feasibility constraint becomes more complicated for that case. This is consistent with the observation that analysis of optimal auctions with asymmetric risk-averse bidders is largely absent from the literature. This additional difficulty does not arise for Olszewski and Siegel (2020), as their approach replaces the n-agent feasibility constraint with a single agent feasibility constraint. Finally, I demonstrate that, for any all-pay mechanism, a contest that implements the same interim prize allocation and induces all types to put in the same effort as in this all-pay mechanism can be constructed. This finding proves the optimality of contests and implies that the optimal contest shares the same prize allocation rule as the optimal all-pay mechanism.

Several papers also study contest design with complete information and convex effort costs. For example, Schweinzer and Segev (2012) investigate the optimal prize structure of Tullock contests. Drugov and Ryvkin (2020) study the optimal allocation of prizes in the Lazear-Rosen Model. Similar to the present paper, Drugov and Ryvkin (2020) also find that when the designer can set prizes optimally, she always prefers to have as many contestants as possible. Fang, Noe, and Strack (2020) characterize the optimal prize profile for noiseless contests. Letina, Liu, and Netzer (2022) take a unifying approach and allow the designer to choose both the prize profile and the contest format. With n contestants, they find that it is always optimal for the designer to utilize a nested Tullock contest with $n-1$ identical prizes. By contrast, I find that when contestants have different and privately known types, the designer may find it optimal to utilize winner-takes-all or several descending prizes. The sharp contrast between my results and those of Letina, Liu, and Netzer (2022) highlights the role played by private information.

The rest of this paper is organized as follows. Section 2 presents the model. Sections 3 and 4 summarize the main results. Finally, Section 5 concludes the paper.

2 The Model

There is one designer (female) and $n \geq 2$ agents (male). All parties are risk neutral. The designer wants to maximize the expected total effort and has a fixed prize budget, which is normalized to 1, to allocate to the agents.[4] Each agent $i \in N = \{1, 2, ..., n\}$ is characterized by his privately-known type $\theta_i \in \Theta = [0, 1]$.[5] A high θ_i means that agent i has a high valuation of the prize. It is common knowledge that agents' types are I.I.D. according to a distribution $F : \Theta \to [0, 1]$. F is twice continuously differentiable, and the corresponding density function f is strictly positive on Θ. Each agent i can choose an effort level $a_i \in \mathbb{R}_+$ at cost $c(a_i)$. The cost function $c : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, convex, and twice continuously differentiable, and satisfies $c(0) = 0$.

If agent i of type θ_i chooses effort level a_i and obtains prize x_i, then his utility is given by

$$u_i(\theta_i, a_i, x_i) = \theta_i x_i - c(a_i).$$

The agent gets an outside option of 0 if he chooses not to participate in the game.[6] Note that

$$\frac{\partial u_i}{\partial \theta_i} = x_i > 0,$$

[4]As contestants' payoffs are linear in the prizes, this is equivalent to the alternative setting where the prize is non-divisible and the designer chooses who wins the prize.

[5]The type space can be any bounded interval. The choice of $[0, 1]$ is simply a normalization.

[6]Moldovanu and Sela (2001) assume that the agent's private information is about his ability and that the agent's preference is described by

$$v_i(\theta_i, a_i, x_i) = x_i - \frac{c(a_i)}{\theta_i}.$$

Their setting is mathematically equivalent to the one used in this paper.

$$\frac{\partial u_i}{\partial(-a_i)} = -\frac{\partial u_i}{\partial a_i} = c'(a_i) > 0,$$

and

$$\frac{\partial^2 u_i}{\partial^2(-a_i)} = \frac{\partial^2 u_i}{\partial^2 a_i} = -c''(a_i) < 0.$$

If we re-interpret $-a_i$ as the agent's wealth, u_i can also be thought of as the preference of a bidder whose utility is increasing in both his type and his wealth and who has a diminishing return to wealth. Such a bidder's preference satisfies Assumption A in Maskin and Riley (1984, p. 1476) and can thus be viewed as a risk-averse bidder.

A (single) contest $(\boldsymbol{s}, \boldsymbol{\mu})$ is characterized by a *prize splitting rule* $\boldsymbol{s}$ and a *contest success function* $\boldsymbol{\mu}$. The prize splitting rule $\boldsymbol{s} = (s_1, ...s_n) : \Pi_{j\in N}\mathbb{R}_+ \to [0,1]^n$ maps any effort profile $\boldsymbol{e} = (e_1, ..., e_n)$ into a vector of n non-increasing prizes ($s_1 \geq s_2 \geq ... \geq s_n \geq 0$) such that $\sum_{j\in N} s_j = 1$. Since I allow for zero prizes, it is without loss of generality to assume that there are always n prizes. The contest success function $\boldsymbol{\mu} = (\mu_i^k)_{i,k\in N} : \Pi_{i\in N}\mathbb{R}_+ \to [0,1]^{n^2}$ determines the possibly random allocation of the prizes to the agents as a function of the effort profile. Specifically, for any given effort profile $\boldsymbol{e}$, $\mu_i^k(\boldsymbol{e})$ describes the probability that agent i receives the prize $s_k(\boldsymbol{e})$. In addition, the contest success function satisfies: for any given $\boldsymbol{e}$, $\sum_{k\in N} \mu_i^k(\boldsymbol{e}) = 1$ for all $i \in N$ (every agent receives a prize with probability 1) and $\sum_{i\in N} \mu_i^k(\boldsymbol{e}) = 1$ for all $k \in N$ (every prize is allocated to an agent with probability 1). This formulation of a contest includes all of the common contest formats, such as noiseless contests, Tullock contests, and Lazear-Rosen tournaments. In this formulation, I allow both the size of the prizes and the contest success function to be endogenously determined by the effort profile. This is equivalent to the more customary setting where the prize splitting rule is fixed but the contest success function can be endogenously determined by the effort profile (e.g., Tullock contests). I choose the current formulation to ease the comparison with the literature, and will further discuss the above mentioned equivalence in Section 3.3.

The timeline of the game is as follows. First, the designer announces $(\boldsymbol{s}, \boldsymbol{\mu})$. Second, after observing their private types and contest rules, all agents simultaneously choose their effort. Third, the designer allocates prizes according to the realized effort profile and contest rules. Let $\sigma_i : \Theta \to \triangle\mathbb{R}+$ be agent i's strategy, and let $\boldsymbol{\sigma} = (\sigma_1 ... \sigma_n)$ denote a strategy profile. Similarly, let $b_i : \Theta \to \mathbb{R}+$ denote agent i's pure strategy, and let $\boldsymbol{b} = (b_1 ... b_n)$ denote a pure strategy profile. I say that a contest $(\boldsymbol{s}, \boldsymbol{\mu})$ implements a strategy profile $\boldsymbol{\sigma}$ (or a pure strategy profile $\boldsymbol{b}$) if $\boldsymbol{\sigma}$ (or $\boldsymbol{b}$) forms a Bayesian Nash equilibrium in the game induced by $(\boldsymbol{s}, \boldsymbol{\mu})$.

3 A Single Contest is Optimal

In this section, I first show that, assuming the designer commits to allocate the entire prize with probability 1, a single contest is optimal among all implementable (i.e., incentive compatible and individually rational) mechanisms.[7] It then directly follows that hosting a single contest is optimal among all feasible contest structures.[8]

3.1 Direct Mechanisms

By the revelation principle, it can be assumed without loss of generality that the designer uses a direct mechanism. To allow for both random allocations of the prize and random efforts, I introduce a random variable $\omega \in [0,1]$ to capture all randomness in the mechanism. In a direct mechanism $(\boldsymbol{q}, \boldsymbol{a})$, each agent reports his type to the mechanism, and a single number ω is randomly drawn from $[0,1]$ according to the uniform distribution. ω is independent of $\boldsymbol{\theta}$ and is the same for all agents. The designer specifies for each agent i a prize allocation rule $q_i : \Pi_{i\in N}\Theta \times [0,1] \to [0,1]$ and a recommended effort plan $a_i : \Pi_{i\in N}\Theta \times [0,1] \to [0,\infty)$, as functions of both the reported types and the realization of the random variable ω. The allocation rule must also satisfy $\sum_{i\in N} q_i(\boldsymbol{\theta}, \omega) = 1$ for all $(\boldsymbol{\theta}, \omega)$, as the designer never withholds any prize. For any given type profile, both the prize allocation and the effort may be random, as they also depend on the random number ω.[9] In addition, the allocation of the prize may be random conditional on the effort profile, and vice versa. As is customary in the mechanism design literature, I assume that the agents will take the designer's recommendation provided that it is incentive compatible for them to do so.

The following proposition identifies necessary and sufficient conditions for a direct mechanism to be implementable. For any allocation rule $\boldsymbol{q}$, let

$$Q_i(\theta_i) = \mathbb{E}[q_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega) \mid \theta_i]$$

denote the expected share of the prize to be allocated to agent i, given that he is of type θ_i. Q_i is also referred to as the *interim allocation rule.* An interim allocation rule $\boldsymbol{Q} = (Q_i)_{i=1,2\ldots N}$

[7] If the designer can retain (some of) the prize, then a single contest with reserve effort is optimal.

[8] If the designer conducts a multi-stage mechanism and agent i chooses effort level a_i^t at each stage t, then the agent's total cost of effort is given by $c(\sum_t a_i^t)$. In other words, I assume the designer cannot influence the cost of producing effort by simply adding more periods.

[9] This approach of modeling randomness in the mechanism is without loss of generality by Halmos and von Neumann (1942).

is *feasible* if and only if it can be induced by an allocation rule $\boldsymbol{q}$. For any θ_i and θ_i', let

$$V_i(\theta_i, \theta_i') = \mathbb{E}\left[\theta_i q_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega) - c(a_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i, \theta_i'\right]$$

denote agent i's expected utility if he is of type θ_i but reports that he is of type θ_i', assuming that all other agents report truthfully. I slightly abuse notation by using $V_i(\theta_i) = V_i(\theta_i, \theta_i)$ to denote the agent's expected utility when he reports truthfully.

Proposition 1. *(Payoff Equivalence) Let* $(\boldsymbol{q}, \boldsymbol{a})$ *be any direct mechanism, and let* $\boldsymbol{Q}$ *denote the corresponding interim allocation rule. This mechanism is implementable if and only if for all* i *and* θ_i*,*

(a) $\boldsymbol{Q}$ *is feasible;*

(b) $Q_i(\theta_i)$ *is non-decreasing in* θ_i*;*

(c) $V_i(\theta_i) = V_i(0) + \int_0^{\theta_i} Q_i(t)dt$*; and*

(d) $V_i(0) \geq 0$.

The above proposition states that agents' utilities in any implementable mechanism are essentially uniquely determined by the interim prize allocation, as in the risk-neutral bidder problem. It directly follows that, in any implementable mechanism, the expected total effort cost borne by agents,

$$\begin{aligned} \mathbb{E}\left[c(a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i\right] =& \mathbb{E}\left[\theta_i q_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega) \mid \theta_i\right] - V_i(\theta_i) \\ =& \theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt - V_i(0), \end{aligned}$$

is also essentially uniquely determined by $\boldsymbol{Q}$. However, the designer's objective is to maximize the expected total effort, not the effort cost. So the other classical result, revenue equivalence, does not hold unless the effort cost is linear. The designer's "revenue" (i.e., the expected total effort) not only depends on $\boldsymbol{Q}$, but may also depend on the exact details of the mechanism.

3.2 All-Pay Mechanisms

Since my objective is to maximize the designer's revenue, in the following analysis I only consider mechanisms for which $V_i(0) = 0$. For later use, I define the term *all-pay mechanism* here:[10]

[10] I thank Juuso Välimäki for suggesting this name to me.

Definition 1. *An implementable direct mechanism* $(\boldsymbol{q}, \boldsymbol{a})$ *is an all-pay mechanism if and only if for any* i *and* θ_i, $a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)$ *is constant for all* $\boldsymbol{\theta_{-i}}$ *and* ω.

In an all-pay mechanism, any agent's effort depends only on his type and is independent of any other agents' types and the realization of the random variable ω. The definition also implies that the effort a type-θ agent is required to exert cannot be made contingent on the realized allocation of prizes. As will be explained in Section 3.3, this "all-pay" property of effort can also be induced by contests. But it cannot be induced, for instance, by first-price auctions: in a first-price auction, an agent is only required to exert effort when he obtains the entire prize.

The following proposition shows that all-pay mechanisms are optimal among all implementable mechanisms.

Proposition 2. *For any implementable direct mechanism* $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$, *there exists an all-pay mechanism that implements the same allocation rule and that is at least as profitable for the designer.*

This proposition claims that even if the designer can make the effort recommendation rule as complicated as she wishes (e.g., the agents are requested to exert effort only when they receive the entire prize or when they receive more than 70% of the prize), she still prefers to use a simple all-pay rule. This conclusion stems from the convexity of the cost function and the payoff equivalence result established in the previous subsection, as follows. Suppose that the designer wants to implement a certain allocation rule $\boldsymbol{q}$. Since the agents have convex effort costs and must obtain the same interim payoff in any implementable mechanism, the designer finds it optimal to utilize the mechanism that minimizes effort variations while maintaining incentive compatibility. This can be achieved by requiring agents to always exert a level of effort that depends only on their type.

Proposition 2 implies that, in searching for an optimal mechanism, we can confine our attention to the class of all-pay mechanisms. For any all-pay mechanism $(\boldsymbol{q}, \boldsymbol{a})$, let $\boldsymbol{Q}$ denote the corresponding interim allocation rule, and let $a_i(\theta_i \mid \boldsymbol{Q}) = a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)$ for all $\boldsymbol{\theta_{-i}}$ and all ω. By Proposition 1, we obtain the following:

$$\begin{aligned}\int_0^{\theta_i} Q_i(t)dt &= V_i(\theta_i)\\ &= \mathbb{E}\left[\theta_i q_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega) - c(a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i\right]\\ &= \theta_i Q_i(\theta_i) - c(a_i(\theta_i \mid \boldsymbol{Q}))\\ \Rightarrow a_i(\theta_i \mid \boldsymbol{Q}) &= c^{-1}\left(\theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt\right).\end{aligned}$$

As c is strictly increasing and convex, its inverse function c^{-1} is well defined, strictly increasing, and concave. The expected total effort is given by

$$\begin{aligned}R(\boldsymbol{Q}) &= \sum_{i=1,2\dots n} \int_0^1 a_i(\theta_i \mid \boldsymbol{Q}) dF(\theta_i)\\ &= \sum_{i=1,2\dots n} \int_0^1 c^{-1}\left(\theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt\right) dF(\theta_i).\end{aligned}$$

Using the linear effort function $c(a_i) = a_i$, the above formula becomes

$$\begin{aligned}R(\boldsymbol{Q}) &= \sum_{i=1,2\dots n} \int_0^1 \left[\theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt\right] dF(\theta_i)\\ &= \sum_{i=1,2\dots n} \int_0^1 Q_i(\theta_i)\left[\theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}\right] dF(\theta_i).\end{aligned}$$

The second equality is obtained through integration by parts and is the same as the revenue function obtained in Myerson (1981). Further, as c^{-1} is concave, we can without loss of generality restrict our attention to symmetric interim allocation rules, and thus I drop all individual subscripts for Q_i.[11] So the designer's problem is reduced to choosing a feasible and non-decreasing interim allocation rule Q to maximize

$$n \int_0^1 a(\theta \mid Q) dF(\theta).$$

[11]To see this, consider the case of two agents. Suppose that there exists an optimal pair (Q_1^*, Q_2^*) such that $Q_1^* \neq Q_2^*$. As the two agents are symmetric, (Q_2^*, Q_1^*) must also be optimal. But then the symmetric allocation rule $(\frac{Q_2^*+Q_1^*}{2}, \frac{Q_2^*+Q_1^*}{2})$ is also feasible and increasing, and at least as profitable for the seller (since c^{-1} is concave). Moreover, when c^{-1} is strictly concave, $(\frac{Q_2^*+Q_1^*}{2}, \frac{Q_2^*+Q_1^*}{2})$ is strictly more profitable than (Q_1^*, Q_2^*). The argument can be easily generalized to incorporate more agents.

3.3 The Optimality of Contests

To state the results more formally, I introduce some new terminology: I say that a contest $(\boldsymbol{s}, \boldsymbol{\mu})$ is a *noiseless contest* if for every effort profile $\boldsymbol{e} \in \Pi_{i \in N}\mathbb{R}$, all $i, k \in N$, $\mu_i^k(\boldsymbol{e}) \in \{0, 1\}$.[12] In addition, taking any interim allocation rule Q, I say that a contest $(\boldsymbol{s}, \boldsymbol{\mu})$ *implements* Q (in pure strategies) if this contest implements a pure strategy profile $\boldsymbol{b}$ such that, for all i and all θ_i,

$$Q(\theta_i) = \mathbb{E}\left[\sum_{k \in N} \mu_i^k((b(\theta_i, \boldsymbol{\theta_{-i}}))s_k(\boldsymbol{b}(\theta_i, \boldsymbol{\theta_{-i}}))\middle|\theta_i\right].$$

In order to show that the set of optimal mechanisms also includes a contest, it suffices to show that Proposition 3 holds.

Proposition 3. *For any feasible and non-decreasing Q, there exists a noiseless contest $(\boldsymbol{s}, \boldsymbol{\mu})$ such that:*

(I) $(\boldsymbol{s}, \boldsymbol{\mu})$ implements Q and the symmetric pure strategy profile

$$b(\theta) = a(\theta \mid Q) = c^{-1}\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)$$

for all $\theta \in [0, 1]$.

(II) For all $\boldsymbol{e} \in \Pi_{i \in N}\mathbb{R}$, all $i, k \in N$, $\mu_i^k(\boldsymbol{e}) = 1$ only if

$$\left|\{j \mid e_j > e_i\}\right| < k \leq \left|\{j \mid e_j \geq e_i\}\right|.$$

The above proposition states that, taking any feasible and non-decreasing interim allocation rule Q, a noiseless contest that implements Q and the strategy profile $b(\theta) = a(\theta \mid Q)$ for all θ can always be constructed. In this noiseless contest, the prize profile may vary depending on the realization of the effort profile but prizes are always awarded in descending order according to contestants' ranks. This is equivalent to stating that, for any all-pay mechanism, there exists a noiseless contest that implements the same interim prize allocation and induces all types to put in the same effort as in this all-pay mechanism. Therefore, since the set of

[12] The definition does not allow for random tie-breaking. This is without loss as I allow for prizes of equal sizes.

optimal mechanisms always includes an all-pay mechanism, it must also include a noiseless contest.

Note that there often exist multiple contests that can induce the same effort distribution. For instance, awarding multiple (positive) prizes deterministically is equivalent to awarding a single (positive) prize with some randomness in the allocation process. To see how this works, take any noiseless contest $(\boldsymbol{s}, \boldsymbol{\mu})$ such that $\boldsymbol{s}(\boldsymbol{e}) \neq (1, 0...0)$ for some $\boldsymbol{e}$, and let $\boldsymbol{p} = (p_1, ...p_n) : \Pi_{i \in N}\mathbb{R}_+ \to [0,1]^n$ denote the corresponding mapping from effort profiles to prizes. We can construct another contest $(\hat{\boldsymbol{s}}, \hat{\boldsymbol{\mu}})$ as follows: for every effort profile $\boldsymbol{e}$, let $\hat{\boldsymbol{s}}(\boldsymbol{e}) = (1, 0...0)$ and $\hat{\boldsymbol{\mu}}(\boldsymbol{e}) = \boldsymbol{p}(\boldsymbol{e})$. By construction, it is clear that in these two contests, the same effort profile always leads to the same interim prize allocation. By following very similar arguments as in the proof for Proposition 3, one can verify that the newly constructed contest implements the same interim allocation rule and strategy profile as the original contest. For convenience, I will henceforth consider only noiseless contests with single or multiple (positive) prizes, and I will provide an example in Section 4.2 to further illustrate the equivalence.

Since a single contest is optimal among all implementable mechanisms, it directly follows that there is no added benefit to altering the contest structure.

Before concluding this section, I note that although I mainly consider symmetric agents in this paper, the optimality of all-pay mechanisms also extends to the case where agents have asymmetric type distributions and/or asymmetric cost functions (see the proof of Proposition 2 for details). With asymmetric agents, the optimal effort can still be implemented by a noiseless contest, but the optimal prize allocation rule may be asymmetric. Therefore, the prizes allocated to agents are not necessarily non-decreasing in their effort.

4 Optimal Prize Allocation

Having established the optimality of a noiseless contest with descending prizes, I next turn to characterizing the optimal prize allocation. In order to employ the variational approach, I only consider piecewise differentiable rules Q.

As illustrated at the end of Section 3.2, the designer's objective is to choose a feasible and non-decreasing interim allocation rule Q to maximize

$$R(Q) = \int_0^1 c^{-1}\Big(\theta Q(\theta) - \int_0^\theta Q(t)dt\Big)dF(\theta).$$

Maskin and Riley (1984) and Matthews (1983) find that to ensure that a non-decreasing Q is also feasible, it suffices to verify that Q satisfies the following two constraints: for any $\theta \in [0,1]$, $Q(\theta) \in [0,1]$ and

$$\int_\theta^1 Q(t)dF(t) \leq \int_\theta^1 F^{n-1}(t)dF(t).$$

The above inequality will be referred to as the *interim feasibility constraint*. It requires that, for any θ, the expected share of the prize to be allocated to types in $[\theta,1]$ is never higher than the probability that such a type exists. So when the designer selects the optimal Q for any subset A of types, she must also take into consideration the effect on the feasibility of Q on $[\theta,1] \setminus A$ for all θ.

As the aforementioned two papers do not require the designer to always exhaust the prize budget, I need to impose the additional constraint

$$n \int_0^1 Q(t)dF(t) = 1$$

to capture this extra restriction. This equation states that the total prizes assigned to the n agents always add up to 1, i.e., the entire prize budget.

The designer's problem can thus be summarized as

$$\max_Q \int_0^1 c^{-1}\Big(\theta Q(\theta) - \int_0^\theta Q(t)dt\Big)dF(\theta),$$

subject to the constraints that, for any $\theta \in [0,1]$,

(a) Q is non-decreasing;

(b) $Q(\theta) \in [0,1]$;

(c)

$$\int_\theta^1 Q(t)dF(t) \leq \int_\theta^1 F^{n-1}(t)dF(t);$$

(d)

$$\int_0^1 Q(t)dF(t) = \frac{1}{n}.$$

As mentioned in the Introduction, Olszewski and Siegel (2020) identify the optimal prize distribution for the limit case where $n \to \infty$. In their general cost case, they solved a problem that is similar to the one above, but within the large contest framework.[13] Their approach utilizes the observation that, with a large number of contestants, in equilibrium prizes are almost assortatively allocated to agents according to the rank-order quantile of their types in the type distribution. Equipped with this insight, the authors then show that the interim allocation Q in the designer's problem can be (approximately) replaced by the assortative allocation of prizes. The n-agent prize allocation problem can thus be transformed into a single agent problem in which the designer's task is to choose the optimal distribution of prizes. In this transformed problem, the interim feasibility constraint (c) is replaced by a total budget constraint that is similar to constraint (d). The designer can therefore select the optimal prize for each type independently, provided that the total prize budget is not exceeded. In this paper, as I allow for any arbitrary number of contestants, the aforementioned approximation does not work. The optimization must be performed subject to this additional constraint, and thus a different technique is needed. I will further discuss the role played by the interim feasibility constraint after presenting Theorem 2.

Instead of solving the above problem directly, I consider a relaxed version with the same objective function but only constraints (c) and (d). If the solution to the relaxed problem satisfies the ignored constraints, then it is also the solution to the original problem. The relaxed problem can be transformed into a multi-variable calculation of variation problem. To do this, first define two variables

$$x(\theta) = \int_\theta^1 Q(t)dF(t)$$

and

$$I(\theta) = \int_0^\theta Q(t)dt.$$

It follows that

$$Q(\theta) = -\frac{x'(\theta)}{f(\theta)},$$

$$I(\theta) = -\int_0^\theta \frac{x'(t)}{f(t)}dt,$$

and

$$I'(\theta) = -\frac{x'(\theta)}{f(\theta)}.$$

Note that the definitions of x and I directly imply that $x(1) = 0$ and $I(0) = 0$, constraint (d)

[13] See Olszewski and Siegel (2020) Appendix A2 for detail.

is equivalent to requiring that $x(0) = \frac{1}{n}$, and $I(1)$ is not fixed. Finally, define the function

$$h(\theta, x'(\theta), I(\theta)) = c^{-1}\Big(- \theta \frac{x'(\theta)}{f(\theta)} - I(\theta)\Big) f(\theta).$$

Then the designer's (relaxed) problem becomes

$$\max_{x,I} Y(x, I) = \int_0^1 h(\theta, x'(\theta), I(\theta)) d\theta$$

s.t. for any $\theta \in [0, 1]$,

$$x(\theta) \leq \int_\theta^1 F^{n-1}(t) dF(t),$$

$$I'(\theta) + \frac{x'(\theta)}{f(\theta)} = 0,$$

and the boundary conditions $x(0) = \frac{1}{n}$, $x(1) = 0$, and $I(0) = 0$ must be satisfied. In this formulation, the first constraint corresponds to constraint (c) in the original problem, the second constraint relates I to x, and constraint (d) in the original problem is captured by requiring that $x(0) = \frac{1}{n}$.

The problem can then be solved using the infinite-dimensional generalized Kuhn-Tucker Theorem (see, e.g., Luenberger (1969), p. 249). The Lagrangian is given by

$$\begin{aligned}\mathcal{L}(x, I) =& h(\theta, x'(\theta), I(\theta)) - \lambda(\theta) \left(x(\theta) - \int_\theta^1 F^{n-1}(t) dF(t) \right) \\ & - \mu(\theta) \left(I'(\theta) + \frac{x'(\theta)}{f(\theta)} \right).\end{aligned}$$

By the Kuhn-Tucker Theorem, any (x^*, I^*) that maximizes Y must satisfy the following necessary conditions:

(1) The Euler-Lagrange conditions

$$\frac{\partial \mathcal{L}}{\partial I} - \frac{d \frac{\partial \mathcal{L}}{\partial I'(\theta)}}{d\theta} = 0$$

and

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d \frac{\partial \mathcal{L}}{\partial x'}}{d\theta} = 0$$

must hold wherever they are well-defined;

(2) The Lagrangian multiplier $\lambda(\theta) \geq 0$;

(3) The complementary slackness condition

$$\lambda(\theta)\left(x(\theta) - \int_{\theta}^{1} F^{n-1}(t)dF(t)\right) = 0;$$

(4) The boundary conditions $x_0(0) = \frac{1}{n}$, $x(1) = 0$, and $I(0) = 0$.

As $I(1)$ is not given, we also need to impose the following *transversality condition* (see Sagan (1992), p. 73, Theorem 2.8):

$$\frac{\partial \mathcal{L}}{\partial I'} \mid_{\theta=1} = 0,$$

to ensure that $I(1)$ is chosen optimally.

Note that there is a clear analogy between the infinite dimensional Kuhn-Tucker Theorem and the finite dimensional one, with Euler-Lagrange conditions replacing the role of the commonly used "first order conditions." I will further discuss the implications of these necessary conditions in the following analysis.

4.1 Winner-Takes-All

I first identify a necessary and sufficient condition under which the designer finds it optimal to allocate the entire prize to the agent who exerts the highest effort in the contest. This prize allocation scheme is commonly referred to as *winner-takes-all.* Proposition 1 implies that the agent who exerts the highest effort must be the one with the highest type. Thus, with winner-takes-all, an agent of type θ obtains one unit of the prize with probability $F^{n-1}(\theta)$, or equivalently, expects to receive $Q(\theta) = F^{n-1}(\theta)$ units of the prize.

Recall that in Section 3.2,

$$a(\theta \mid Q) = c^{-1}\left(\theta Q(\theta) - \int_{0}^{\theta} Q(t)dt\right)$$

was defined to be the effort exerted by an agent of type θ given the interim allocation rule Q. In addition, let

$$a(\theta) = a(\theta \mid F^{n-1})$$

denote the effort exerted by an agent of type θ when $Q(\theta) = F^{n-1}(\theta)$ for all θ, and let

$$J(\theta \mid Q) = \frac{\theta}{c'\left(a(\theta \mid Q)\right)} - \frac{\int_\theta^1 \frac{1}{c'(a(s|Q))} dF(s)}{f(\theta)}$$

and

$$\begin{aligned} J(\theta) &= J(\theta \mid F^{n-1}) \\ &= \frac{\theta}{c'(a(\theta))} - \frac{\int_\theta^1 \frac{1}{c'(a(s))} dF(s)}{f(\theta)}, \end{aligned}$$

wherever they are well defined. Note that $J(\theta)$ is always well defined and continuous, except possibly at $\theta = 0$. I refer to $J(\theta)$ as the *generalized virtual value*, and I will sometimes refer to the following assumption as the *regularity condition*:

Assumption R The function $J(\theta)$ is non-decreasing on [0,1].

The following theorem shows that Assumption R is a necessary and sufficient condition for the winner-takes-all prize allocation rule being optimal.

Theorem 1. *Allocating the entire prize to the agent who exerts the highest effort, i.e., $Q^*(\theta) = F^{n-1}(\theta)$ for all $\theta \in [0,1]$, is optimal for the designer if and only if Assumption R holds.*

The necessary conditions presented at the beginning of this section are equivalent to the following two conditions:

$$\frac{dJ(\theta \mid Q)}{d\theta} = \lambda(\theta) \geq 0, \tag{1}$$

whenever this condition is well defined, and the complementary slackness condition that for any θ,

$$\lambda(\theta) \left(\int_\theta^1 Q(t) dF(t) - \int_\theta^1 F^{n-1}(t) dF(t) \right) = 0. \tag{2}$$

It directly follows that for $Q^*(\theta) = F^{n-1}(\theta)$ to be optimal, $J(\theta)$ must be non-decreasing on $[0,1]$. As the objective function is concave, all constraints are linear, and $Q^*(\theta)$ is continuously differentiable, this necessary condition is also sufficient for Q^* being optimal (see proof of Theorem 1 for details). Thus, $Q^*(\theta)$ is a solution to the relaxed problem, and it is therefore also a solution to the original problem because it satisfies all of the ignored constraints.

The intuition underlying Theorem 1 is as follows: Fix any feasible and non-decreasing Q. The first term of $J(\theta \mid Q)$,[14]

$$\frac{\theta}{c'\left(a(\theta \mid Q)\right)} = \frac{\partial a(\theta \mid Q)}{\partial Q(\theta)},$$

measures the direct marginal benefit of assigning more of the prize to a bidder of type θ. The second term,

$$-\frac{\int_\theta^1 \frac{1}{c'(a(s|Q))} dF(s)}{f(\theta)} = \frac{1}{f(\theta)} \int_\theta^1 \frac{\partial a(s \mid Q)}{\partial \left(\int_0^\theta Q(t)dt\right)} dF(s),$$

captures the indirect marginal cost of increasing $Q(\theta)$. To see this, note that $\int_0^s Q(t)dt$ measures the scope of information rent: An agent of type s has more incentive to misreport when agents of lower types obtain larger prizes.[15] All else being the same, an increase in $Q(\theta)$ will lead to an increase in $\int_0^s Q(t)dt$ for any $s > \theta$, which in turn reduces the effort from any bidder whose type is higher than θ. Thus, for any given allocation rule Q, $J(\theta \mid Q)$ describes the total (direct and indirect) marginal return of assigning more of the prize to an agent of type θ.

Again, take any feasible and non-decreasing Q. If $J(\theta \mid Q)$ is increasing, then the designer can improve the mechanism by assigning more of the prize to agents of higher types (if such changes are feasible). Further, if $J(\theta \mid Q)$ is increasing for all feasible Q, the designer will find it optimal to assign as much of the prize as the feasibility constraint permits to the agent with the highest type—i.e., she will allocate the entire prize with probability 1 to such an agent. As the designer's problem is concave, to check that $J(\theta \mid Q)$ is non-decreasing for any feasible Q, it suffices to check that $J(\theta)$ is non-decreasing. Note that when the cost function is linear, $J(\theta)$ is reduced to the standard virtual value found by Myerson (1981), namely, $\theta - \frac{1-F(\theta)}{f(\theta)}$.

Moldovanu and Sela (2001) identify a necessary and sufficient condition under which winner-takes-all performs better than two prizes in descending order (Proposition 5, p. 549). This condition is weaker than the regularity condition obtained above, as the latter ensures

[14] In obtaining the following two equations, I use $(c^{-1}(t))' = \frac{1}{c'(c^{-1}(t))}$.

[15] As in many other mechanism design problems, here the agents never have incentives to misreport that they are of a higher type than they actually are.

that winner-takes-all outperforms all other prize allocation rules.

Consistent with the findings of Olszewski and Siegel (2020), I show that when the contest becomes sufficiently large, the regularity condition always fails, and thus the winner-takes-all prize structure cannot be optimal:

Proposition 4. *Suppose that $c'(0) = 0$. For any F and c, there exists $\bar{n}$ such that Assumption R fails for all $n \geq \bar{n}$.*

Note that the additional assumption, $c'(0) = 0$, is also imposed by Olszewski and Siegel (2020) to obtain their corresponding results.

To gain a better understanding of the above results, consider the example where $c(a) = a^2$ and θ is uniformly distributed on $[0, 1]$. Computation yields

$$\frac{dJ(\theta \mid Q)}{d\theta} = \frac{4\left[\theta Q(\theta) - I(\theta)\right] - \theta^2 Q'(\theta)}{4\left[\theta Q(\theta) - I(\theta)\right]^{\frac{3}{2}}}$$

wherever it is well defined, and

$$J'(\theta) = \frac{1}{4\left(\frac{n-1}{n}\theta\right)^{\frac{3}{2}}} \frac{(n-1)(4-n)}{n}\theta^n.$$

If $2 \leq n \leq 4$, then $J'(\theta) \geq 0$ for all θ in $[0, 1]$, and thus the regularity condition holds; otherwise, the condition fails.

4.2 Multiple Prizes

When Assumption R fails, the winner-takes-all scheme is no longer optimal and the designer prefers to utilize multiple (positive) prizes of descending sizes. For this case, I characterize the optimal interim allocation rule under the following assumption, which is essentially the same as Assumption 2 in Olszewski and Siegel (2020):

Assumption M For any $\theta \in [0, 1]$,

$$\frac{2}{c'(c^{-1}(1))} + \frac{f'(\theta)[1 - F(\theta)]}{c'(0) f^2(\theta)} > 0,$$

where, if $c'(0) = 0$, the second term is set to equal ∞, $-\infty$, or 0 when its numerator is positive, negative, or 0, respectively.

This assumption holds, for example, when the distribution F is convex (regardless of n and c).

The next theorem states that the optimal interim allocation rule may only consist of two regions: where the entire prize is awarded to the agent with the highest effort, and where the marginal return of assigning more of the prize is constant across types.

Theorem 2. *Suppose that Assumption M holds and that c is strictly convex. Let Q^* denote the optimal interim allocation rule. For any $\theta \in [0, 1]$, one of the following statements must be true:*

(I) There exists an interval (x, y) such that $\theta \in (x, y)$, $Q^*(\theta) = F^{n-1}(\theta)$, and $J(\theta \mid Q^*)$ is non-decreasing on (x, y);

(II) There exists an interval (y, z) such that $\theta \in (y, z)$, $Q^*(\theta) \neq F^{n-1}(\theta)$, and $J(\theta \mid Q^*)$ is constant on (y, z).

Moreover, (II) holds on a set of positive measure if Assumption R fails.

Technically, the above result follows directly from Conditions (1) and (2) given in Section 4.1: The optimal Q^* may only consist of the two parts: (I) where the feasibility constraint is binding and (II) where Q^* solves

$$\frac{dJ(\theta \mid Q)}{d\theta} = \lambda(\theta) = 0.$$

To see the intuition, suppose that $J(\theta \mid Q^*)$ is decreasing on some non-degenerate interval (α, β). As was already illustrated in the last subsection, this means that the marginal return of assigning more of the prize to a higher type is smaller than the marginal return from assigning more to a lower type, assuming that both types fall in (α, β). Thus, the designer can improve the mechanism by reducing the prize allocated to the higher type and increasing the prize allocated to a lower type. The optimum is achieved when the marginal return becomes constant across types.

In the context of large contests, Olszewski and Siegel (2020)'s Lemma 3 can be used to obtain the optimal prize distribution when contestants have convex effort costs and (possibly) non-linear valuations of the prize. With symmetric contestants and linear valuations, their Lemma 3 corresponds to the special case where Theorem 2 (II) holds almost everywhere on $[0, 1]$. This is because, as explained in the beginning of Section 4, the interim feasibility constraint becomes irrelevant in their environment.

As explained in Section 3.3, for any Q^* as described in Theorem 2, there exists a noiseless contest to implement it: The prizes are always awarded in descending order according to contestants' ranks, but the prize profile may vary depending on the realization of the type profile. If the highest type falls into region (I), then the entire prize is allocated to the contestant with the highest type. If the highest type falls into an interval (y, z) that belongs to region (II), then he shares the prize with any other contestants whose types also fall into the same interval. The exact sharing rule is determined such that the rate of return for a marginal increase in Q becomes constant across types. When Assumption R fails, the designer finds it optimal to at least sometimes allocate multiple prizes.

Theorem 2 offers a formula to compute the optimal interim allocation rule, not a closed-form solution. It is therefore not possible to directly identify a closed-form solution for the optimal prize splitting rule. However, once all of the model's parameters are given, one can utilize Theorem 2 to compute the optimal prize structure. To gain a better understanding of how this can be done, consider the case where $c(a) = a^2$ and θ is uniformly distributed on $[0, 1]$. In the last subsection, it was shown that if $n \leq 4$, the regularity condition holds and the designer finds it optimal to always award the entire prize to the highest bidder. When $n > 4$, $J(\theta)$ is decreasing on $[0, 1]$, and the optimal Q^* must solve

$$\frac{dJ(\theta \mid Q)}{d\theta} = \frac{4\left[\theta Q(\theta) - I(\theta)\right] - \theta^2 Q'(\theta)}{4\left[\theta Q(\theta) - I(\theta)\right]^{\frac{3}{2}}} = 0$$

(i.e. $J(\theta \mid Q^*)$ is constant) almost everywhere on $[0, 1]$.[16] The solution is uniquely given by $Q(\theta) = \frac{4}{n}\theta^3$, which results in effort level $a(\theta) = \sqrt{\frac{3}{n}}\theta^2$.

The designer can implement the aforementioned interim allocation rule by conducting a noiseless contest with multiple positive prizes in descending order, with the number and sizes of the prizes depending on the number of contestants. Let x_n^k denote the prize allocated to the *k-th* highest performer when there are n contestants. For notational convenience, define $x_n^k = 0$ for any $k > n$ or $k = 0$. When $n = 5$, the highest performer is awarded $x_5^1 = \frac{4}{5}$, the second highest performer is awarded $x_5^2 = \frac{1}{5}$, and all other contestants receive nothing (i.e., $x_5^3 = x_5^4 = x_5^5 = 0$). For any $n > 5$, the optimal $k - th$ prize can be computed using the

[16] Here we just need to solve a standard second order differential equation

$$4\left[I'(\theta) - I(\theta)\right] - \theta^2 I''(\theta) = 0$$

with the boundary conditions $I(0) = 0$ and $I(1) = 1$.

following recursive formula (see the Appendix for the computational details):

$$x_n^k = \frac{n-1}{n} x_{n-1}^k + \frac{k-1}{n} \left(x_{n-1}^{k-1} - x_{n-1}^k \right).$$

The number of positive prizes equals $n-3$. When $n=6$, for instance, the above formula implies that it is optimal to award three prizes of $\frac{2}{3}$, $\frac{4}{15}$, and $\frac{1}{15}$ to the 3 highest performers.

As mentioned earlier, there often exist multiple schemes to implement the optimal prize allocation. For instance, when $n=5$, the designer can also induce the optimal prize allocation by awarding the entire prize to the highest performer with probability $\frac{4}{5}$ and to the second highest performer with probability $\frac{1}{5}$.

4.3 Small Contests vs. Large Contests

To maximize the expected total effort, the designer always prefers to have more contestants. This observation directly follows from the concavity of the objective function and the optimality of symmetric mechanisms: Let Q^k denote the optimal allocation rule with k bidders. When there are $n+1$ bidders, it is clear that the designer can earn the same expected revenue by allocating prizes to the first n bidders according to Q^n and never allocating any prize to the last bidder. As argued in Section 3.2, when c^{-1} is (strictly) concave, the optimal symmetric allocation generates (strictly) more expected total effort than any asymmetric rules. Thus, when there are $n+1$ agents, the designer can induce (strictly) more effort by using Q^{n+1} instead of the asymmetric allocation rule constructed above. Therefore, the designer's revenue is always non-decreasing in n, strictly so when c is strictly convex. However, for different levels of n, the increase in revenue may be driven by different forces.

The optimal prize distribution responds to the change of n in a more complicated way. To understand this, first note that, for any interim allocation rule Q, the contestants' total effort cost is given by

$$\Gamma(n, F, Q) = n \int_0^1 Q(\theta) \left[\theta - \frac{1-F(\theta)}{f(\theta)} \right] dF(\theta).$$

I will refer to $\Gamma(n, F, Q)$ as the *competition intensity*, as it measures how aggressively the contestants as a whole engage in costly competition. For any distribution F with a monotone hazard rate, the maximal competition intensity can be achieved by awarding a single prize to the highest performer and is strictly increasing in n. Let

$$\Gamma_{\max} = \lim_{n \to \infty} \Gamma(n, F, F^{n-1})$$

represent the maximal competition intensity the designer can induce with one unit of the prize for a given distribution F. As

$$\Gamma(n, F, F^{n-1}) < n \int_0^1 F^{n-1}(\theta) dF(\theta) = 1$$

and it is increasing in n, $\Gamma_{\max}$ is always well defined and finite. The designer's objective, however, is to maximize the expected total effort, not the total effort costs/competition intensity. The two objectives coincide when the effort cost functions are linear, but they may diverge with convex effort costs: Fixing the total effort cost, the designer ideally wants to equally divide the cost among as many participants as possible. However, a flatter effort distribution can only be induced by a flatter prize distribution, which in turn leads to a decrease in the total effort cost. To select the optimal prize structure, the designer needs to balance the need to induce more aggressive competition among the contestants and the need to smooth the effort costs.

As shown in the previous subsections, for uniformly distributed types and quadratic effort costs, the optimal interim prize allocation rule is given by

$$Q^*(\theta) = \begin{cases} \theta^{n-1} & \text{if } n \leq 4 \\ \frac{4}{n}\theta^3 & \text{if } n > 4. \end{cases}$$

Observe that $Q^*(\theta)$ becomes steeper as n increases to 4, but then becomes flatter as n continues to grow. The contestants' total effort cost is given by

$$\Gamma(n, F, Q^*) = \begin{cases} 1 - \frac{2}{n+1} & \text{if } n \leq 4 \\ \frac{3}{5} & \text{if } n > 4. \end{cases}$$

and their expected total payoff equals

$$n \int_0^1 \left[\int_0^\theta Q^*(t) dt \right] d\theta = \begin{cases} \frac{1}{n+1} & \text{if } n \leq 4 \\ \frac{1}{5} & \text{if } n > 4. \end{cases}$$

When n is small, the winner-takes-all prize structure is optimal. In this case, an increase in the contest size always leads to an increase in the competition intensity. The contestants (as a whole) work more and become worse off. On the other hand, when n is large, the designer no longer finds it optimal to utilize the most competitive prize allocation rule. In this case, as n further increases, the contestants' total effort cost and total welfare remain constant, but the designer's revenue still increases due to more efficient "cost sharing" among more contestants. Also note that the average effort produced by each individual contestant always decreases as n becomes larger.

These observations lead to a natural conjecture: With convex effort costs, there often exists a designer-optimal competition intensity that is strictly smaller than $\Gamma_{\max}$. When the contest size is small, even with the winner-takes-all prize structure, the overall competition intensity is still lower than the optimal level. The designer thus wants to recruit more contestants while keeping the winner-takes-all prize structure to increase the contest's competitiveness. When the contest size is large and the optimal competition intensity has already been reached, further increasing the prize inequality among contestants will only serve to discourage the total effort. However, the designer still prefers to have more contestants, as then she can induce more efficient "cost sharing" among ex ante symmetric contestants with convex effort costs. To balance the needs for "cost smoothing" and for maintaining the optimal competition intensity, the optimal prize distribution has to become flatter as n increases. This observation provides a potential rationale for the findings of Olszewski and Siegel (2020): If there are many contestants and the designer still uses winner-takes-all, the induced competition level will be higher than optimal. Therefore, the designer always prefers to award multiple prizes in large contests.

Regarding welfare implications, note that increasing the contest size always benefits the designer but may or may not hurt the contestants, depending on the initial contest size. Suppose that a social planner wants to increase the contestants' total welfare by imposing an upper limit on contest sizes. Then she should either keep the limit low ($n < 4$ in this example) or not impose any such limits at all (as decreasing n from 8 to 5 does not affect the contestants' payoff but decreases the designer's payoff).

Finally, I provide some comparative statics results. When the contestants' abilities are uniformly distributed, winner-takes-all is less likely to be optimal when the cost function is more convex and more likely to be optimal as the ability distribution becomes more spread out, or equivalently, as the contest task becomes more skill sensitive.

Proposition 5. *Suppose that θ is uniformly distributed on $[0, 1]$. Consider any two functions*

c_1 and c_2 such that c_1 is more convex than c_2.[17] *If Assumption R does not hold when the agents' cost function is given by c_2, it also does not hold when the agents' cost function is given by c_1.*

As the effort cost function becomes more convex, the need to induce "cost smoothing" increases and the marginal benefit of increasing the effort cost decreases. Thus, the optimal competition intensity may be reached earlier. As an illustration, consider the following examples. When θ is uniformly distributed and $c(a) = a$, the regularity condition holds for all n. The optimal competition level equals $\Gamma_{\max} = 1$ and is never attained. When $c(a) = a^{\frac{5}{4}}$, the regularity condition holds if and only if $n \leq 8$. The optimal competition intensity equals $\frac{7}{9}$ and is reached at $n = 8$ and $Q(\theta) = \theta^7$.[18] When $c(a) = a^2$, the optimal intensity equals $\frac{3}{5}$ is reached at $n = 4$ and $Q(\theta) = \theta^3$.

Proposition 6. *Suppose that θ is uniformly distributed on $[\frac{1}{2} - k, \frac{1}{2} + k]$, $k \in (0, \frac{1}{2}]$, and that $c(a) = a^2$. If Assumption R when $k = k'$, then it also holds for any $k > k'$.*

When there is a higher level of uncertainty in the contestants' abilities, competition leads to a greater efficiency gain and is thus more desirable.

5 Conclusion

In this paper, I first demonstrate that contests are the optimal mechanisms to use when agents have private information and convex effort costs, and the designer wants to maximize the expected total effort. I then study the optimal design for contests. The solution is twofold: First, it is always optimal for the designer to employ a single contest with as many contestants as possible. Second, I provide a necessary and sufficient condition for the winner-takes-all prize allocation rule being optimal. When this condition fails, the designer finds it optimal to utilize multiple prizes of descending sizes. I also provide a characterization of the optimal prize distribution for this case. In addition, I illustrate how the optimal prize distribution evolves as the number of contestants increases. Finally, I provide some comparative statics results, which may serve as testable implications of the model. Roughly speaking, the designer has more incentives to encourage competition among the contestants

[17] c_1 is more convex than c_2 if and only if there exists a non-decreasing convex function K such that $c_1(\cdot) = K(c_2(\cdot))$.

[18] For $n \geq 8$, the optimal interim allocation is given by $Q^*(\theta) = \frac{8}{n}\theta^7$, which yields a constant competition intensity of $\frac{7}{9}$.

when the size of the contest is small, when there exists a greater dispersion in the contestant pool, when the task is more skill sensitive, and when the cost function is less convex.

My approach substantially eases the analysis of the optimal contest design, and I expect it to be useful for other contest/auction design frameworks. For instance, the approach can be used to analyze a generalization of the model in which the designer wants to maximize a weighted average of the total effort, the maximum effort, and the contestants' welfare.

Appendix:

For proofs of Proposition 1 and 2, I allow both F_i and c_i to be asymmetric. For all other proofs, the contestants' are assumed to be symmetric.

Proof for Proposition 1: The set of implementable mechanisms is defined by the following conditions. First, it must be incentive compatible for all agents to report their types truthfully. That is, for any θ_i and any θ_i',

$$V_i(\theta_i) \geq V_i(\theta_i, \theta_i').$$

Second, it is assumed that the seller cannot force any bidder to participate. Thus the individual rationality condition,

$$V_i(\theta_i) \geq 0$$

must be satisfied for any θ_i.

(1) (Necessity) The necessity of (d) directly follows the individual rationality constraint. The incentive compatibility constraint requires that for any $\theta_i > \theta_i'$,

$$\begin{aligned} V_i(\theta_i) &\geq V_i(\theta_i, \theta_i') \\ &= \mathbb{E}\left[\theta_i q_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega) - c(a_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i, \theta_i'\right] \\ &= V_i(\theta_i') + (\theta_i - \theta_i')Q_i(\theta') \end{aligned}$$

and similarly,

$$V_i(\theta_i') \geq V_i(\theta_i) + (\theta_i' - \theta_i)Q_i(\theta_i).$$

The above two inequalities together imply,

$$Q_i(\theta) \geq \frac{V_i(\theta_i) - V_i(\theta_i')}{\theta_i - \theta_i'} \geq Q_i(\theta_i').$$

It immediately follows that Q_i must be non-decreasing. In addition, the above inequalities can be rewritten as, for all θ_i' and all $\rho \in (0, 1 - \theta_i']$,

$$\rho Q_i(\theta_i' + \rho) \geq V_i(\theta_i' + \rho) - V_i(\theta_i') \geq \rho Q_i(\theta_i').$$

Since Q_i is non-decreasing and bounded, it is Riemann integrable. This yields, for any

$\theta_i \in [0,1]$,

$$V_i(\theta_i) = V_i(0) + \int_0^{\theta_i} Q_i(t)dt.$$

(2) (Sufficiency) The individual rationality constraint directly follows (c) and (d). Below I show that the incentive compatibility constraint also holds. Consider any $\theta_i > \theta_i'$. (b) and (c) implies

$$\begin{aligned} V_i(\theta_i) - V_i(\theta_i') = & \int_{\theta_i'}^{\theta_i} Q_i(t)dt \\ \geq & Q_i(\theta_i')(\theta_i - \theta_i'). \end{aligned}$$

Hence,

$$\begin{aligned} V_i(\theta_i) \geq & \mathbb{E}\left[\theta_i q_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega) - c(a_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i'\right] \\ & + Q_i(\theta')(\theta_i - \theta_i') \\ = & \mathbb{E}\left[\theta_i q_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega) - c(a_i(\theta_i', \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i, \theta_i'\right] \\ = & V(\theta_i, \theta_i') \end{aligned}$$

Thus, any agent i of type θ_i has no incentive to misreport to be of type θ_i'. Similarly, I can obtain

$$V_i(\theta_i') \quad \geq \quad V_i(\theta_i', \theta_i).$$

That is, agent i of type θ_i' also has no incentive to misreport to be of type θ_i. Since the choice of θ_i and θ_i' are arbitrary, the truthtelling constraint holds for all $\theta_i \in [0,1]$.

Proof for Proposition 2: Fix any implementable direct mechanism $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$, I construct an all-pay mechanism $(\boldsymbol{q}, \boldsymbol{a})$, where $a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega) = b_i(\theta_i)$ is the solution to

$$c_i(b_i(\theta_i)) = \mathbb{E}\left[c_i(\tilde{a}_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i\right].$$

As c is strictly increasing, such $b_i(\theta_i)$ always exists and is unique.

Suppose the designer uses $(\boldsymbol{q}, \boldsymbol{a})$. By construction, type θ_i agent's utility from truthtelling equals

$$\begin{aligned}
&\mathbb{E}\left[q_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)\theta_i - c_i(b_i(\theta_i)) \mid \theta_i\right] \\
=&\mathbb{E}\left[q_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)\theta_i - c_i(\tilde{a}_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)) \mid \theta_i\right] \\
=&\int_0^{\theta_i} Q_i(t)dt.
\end{aligned}$$

I obtain the second equality by applying Proposition 1 and the assumption that $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$ is implementable. Again, by Proposition 1, the constructed mechanism is implementable as it gives any agent i of type θ_i an utility of $\int_0^{\theta_i} Q_i(t)dt$.

As c_i is convex, I obtain for all i and all θ_i,

$$b_i(\theta_i) \geq c_i^{-1}\left(\mathbb{E}\left[c_i\left(\tilde{a}_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)\right) \mid \theta_i\right]\right) \geq \mathbb{E}\left[\tilde{a}_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega) \mid \theta_i\right]$$

using Jensen's inequality. It follows that

$$\sum_{i=1,2...n} E\left[a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)\right] = \sum_{i=1,2...n} E\left[b_i(\theta_i)\right] \geq \sum_{i=1,2...n} E\left[\tilde{a}_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)\right].$$

That is, $(\boldsymbol{q}, \boldsymbol{a})$ generates at least as much expected total effort as $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$, as desired.

Proof for Proposition 3: Consider any feasible and non-decreasing Q. By Manelli and Vincent (2010), any such Q can be induced by a symmetric allocation rule $\boldsymbol{\delta}(\boldsymbol{\theta}) = (\delta_1(\boldsymbol{\theta}), ..., \delta_n(\boldsymbol{\theta}))$ in which each δ_i is non-decreasing in θ_i.

To construct a contest which implements

$$b(\theta) = a(\theta \mid Q) = c^{-1}\left(\theta Q(\theta) - \int_0^{\theta} Q(t)dt\right),$$

we first construct a function that maps effort profiles to type profiles: For any effort profile $\boldsymbol{e} = (e_1, ...e_n)$ and any $i \in N$, define

$$\tilde{\theta}_i(e_i) = \sup\{t \mid e_i \geq a_i(t \mid Q)\}.$$

We then construct a noiseless contest $(\boldsymbol{s}, \boldsymbol{\mu})$ as follows: for any effort profile $\boldsymbol{e}$, let $s(\boldsymbol{e})$ be a descending reordering of the vector $\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}(\boldsymbol{e}))$. The success function $\boldsymbol{\mu}$ satisfies $\mu_i^k(\boldsymbol{e}) \in \{0, 1\}$ for all $\boldsymbol{e}$, i, k and maps $\boldsymbol{s}(\cdot)$ onto $\boldsymbol{\delta}(\tilde{\boldsymbol{\theta}}(\cdot))$. Such a contest success function clearly exists.

By the construction, it is clear that for all θ,

$$b(\theta) = a(\theta \mid Q)$$

constitutes a pure strategy Bayesian Nash equilibrium for the game induced by $(\boldsymbol{s}, \boldsymbol{\mu})$: An agent of type θ has no incentive to deviate to any other type's strategy, as $a(\theta \mid Q)$ is incentive compatible for him, and such an agent also has no incentive to deviate to exert any off-path effort level e', as it is strictly dominated by $a(\tilde{\theta}_i(e') \mid Q)$. Moreover, in this equilibrium, agent i of type θ_i expects to receive

$$\mathbb{E}\left[\delta_i(\theta_i, \boldsymbol{\theta_{-i}})\middle|\theta_i\right] = Q(\theta_i)$$

units of the prize. So the above described contest $(\boldsymbol{s}, \boldsymbol{\mu})$ indeed implements Q as desired.

Proof for Theorem 1: Note that $Q^*(\theta)$ satisfies all constraints (a)-(d). Therefore, to show it is the solution to the original designer's problem, it suffices to show that it is the solution to the following relaxed problem:

$$\max_{Q} \int_0^1 c^{-1}\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right) dF(\theta)$$

subject to the constraint that, for any $\theta \in [0, 1]$,

(c)

$$\int_\theta^1 Q(t)dF(t) \leq \int_\theta^1 F^{n-1}(t)dF(t)$$

(d)

$$\int_0^1 Q(t)dF(t) = \frac{1}{n}.$$

The above problem can be transformed into a calculus of variations problem. To do this, I first define two variables

$$x(\theta) = \int_\theta^1 Q(t)dF(t)$$

and

$$I(\theta) = \int_0^\theta Q(t)dt$$

It directly follows that

$$Q(\theta) = -\frac{x'(\theta)}{f(\theta)},$$

$$I(\theta) = -\int_0^\theta \frac{x'(t)}{f(t)} dt$$

and

$$I'(\theta) = -\frac{x'(\theta)}{f(\theta)}.$$

Note that the definitions of x and I directly imply that $x(1) = 0$ and $I(0) = 0$; constraint (d) is equivalent to requiring $x(0) = \frac{1}{n}$; $I(1)$ is not fixed. Lastly, let

$$h(\theta, x'(\theta), I(\theta)) = c^{-1} \left(-\theta \frac{x'(\theta)}{f(\theta)} - I(\theta) \right) f(\theta)$$

Then, the designer's (relaxed) problem becomes

$$(P) \quad \max_{x,I} Y(x, I) = \int_0^1 h(\theta, x'(\theta), I(\theta)) d\theta$$

s.t., for any $\theta \in [0, 1]$,

$$x(\theta) \leq \int_\theta^1 F^{n-1}(t) dF(t),$$

$$I'(\theta) + \frac{x'(\theta)}{f(\theta)} = 0,$$

$x(0) = \frac{1}{n}$, $x(1) = 0$ and $I(0) = 0$.

I utilize the infinite-dimensional generalized Kuhn-Tucker Theorem (see for example Luenberger (1969), P249) to solve the problem. The Lagrangian is given by

$$\mathcal{L}(\theta, x, I, \lambda, \mu) = h(\theta, x'(\theta), I(\theta)) - \lambda(\theta) \left(x(\theta) - \int_\theta^1 F^{n-1}(t) dF(t) \right) - \mu(\theta) \left(I'(\theta) + \frac{x'(\theta)}{f(\theta)} \right)$$

Suppose (x^*, I^*) maximizes Y. Then it has to satisfy the following four necessary conditions:

(1) The Euler-Lagrange conditions wherever they are well-defined. I will discuss these conditions in details below.

(2) The Lagrangian multiplier $\lambda(\theta) \geq 0$.

(3) The complementary slackness condition

$$\lambda(\theta) \left(x(\theta) - \int_\theta^1 F^{n-1}(t) dF(t) \right) = 0.$$

(4) The boundary conditions $x_0(0) = \int_0^1 F^{n-1}(t)dF(t)$, $x(1) = 0$ and $I(0) = 0$.

Now let us return to the previously mentioned Euler-Lagrange conditions. Here I have two such conditions:

(a) The Euler-Lagrange conditions with respect to I:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial I} - \frac{d\frac{\partial \mathcal{L}}{\partial I'(\theta)}}{d\theta} &= 0 \\ \Rightarrow \frac{\partial h}{\partial I} + \mu'(\theta) &= 0 \\ \Rightarrow \mu'(\theta) &= -\frac{\partial h}{\partial I} \end{aligned}$$

The above equation uniquely defines $\mu'(\theta)$ but not $\mu(\theta)$. However, note that, for any non-uniform distribution, the value of $I(1)$ is not given and thus must be chosen optimally. This can be done by imposing the following *transversality condition* (see for example, Sagan (1992), P73, Theorem 2.8):

$$0 = \frac{\partial \mathcal{L}}{\partial I'}\mid_{\theta=1} = -\mu(1),$$

which implies,

$$\begin{aligned} \mu(\theta) &= \mu(1) - \int_\theta^1 \mu'(t)dt \\ &= \int_\theta^1 \frac{\partial h}{\partial I}(t)dt. \end{aligned}$$

(b) The Euler-Lagrange conditions with respect to x:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} - \frac{d\frac{\partial \mathcal{L}}{\partial x'}}{d\theta} &= 0 \\ \Rightarrow -\lambda(\theta) - \left[\frac{d\frac{\partial h}{\partial x'}}{d\theta} - \left(\frac{\mu(\theta)}{f(\theta)}\right)'\right] &= 0 \\ \Rightarrow \lambda(\theta) + \frac{d\frac{\partial h}{\partial x'}}{d\theta} - \frac{\mu'(\theta)f(\theta) - \mu(\theta)f'(\theta)}{f^2(\theta)} &= 0 \\ \Rightarrow \lambda(\theta) + \frac{d\frac{\partial h}{\partial x'}}{d\theta} + \frac{\frac{\partial h}{\partial I}}{f(\theta)} - \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 \frac{\partial h}{\partial I}(t)dt &= 0 \end{aligned} \tag{3}$$

Recall that by definition,

$$a(\theta \mid Q) = c^{-1}\left(-\theta\frac{x'(\theta)}{f(\theta)} - I(\theta)\right)$$

and

$$J(\theta \mid Q) = \frac{\partial a(\theta \mid Q)}{\partial Q(\theta)} + \frac{\int_\theta^1 \frac{\partial a(s|Q)}{\partial I(s)} dF(s)}{f(\theta)}.$$

Then Equation (3) implies

$$\begin{aligned}\lambda(\theta) =& \frac{d\theta a'(\theta \mid Q)}{d\theta} + a'(\theta \mid Q) + \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 a'(s \mid Q) dF(s) \\ =& \frac{dJ(\theta \mid Q)}{d\theta}\end{aligned}$$

In obtaining the last equality, I use $(c^{-1}(t))' = \frac{1}{c'(c^{-1}(t))}$.

Thus the assumption that $J(\theta)$ is non-decreasing on $[0,1]$ directly implies that the left hand side of Equation (3), evaluated at Q^*, must be non-negative on $[0,1]$. It is then easy to verify that Q^* satisfies all the necessary conditions listed above.

To prove sufficiency, note that the concavity of c^{-1} and the linearly of both constraints imply $\mathcal{L}$ is concave in (x, x', I, I'): take any (x_1, x_1', I_1, I_1'), (x_2, x_2', I_2, I_2') and $\gamma \in [0,1]$,

$$\begin{aligned}&\mathcal{L}(\theta, \gamma x_1 + (1-\gamma)x_2, \gamma I_1 + (1-\gamma)I_2, \lambda, \mu) \\ =& - c^{-1}\left(-\theta\frac{\gamma x_1' + (1-\gamma)x_2'}{f(\theta)} - \gamma I_1 - (1-\gamma)I_2\right) f(\theta) \\ &+ \lambda(\theta)\left[\gamma x_1 + (1-\gamma)x_2 - \int_\theta^1 F^{n-1}(t)dF(t)\right] + \mu(\theta)\left[\gamma I_1' + (1-\gamma)I_2' + \frac{\gamma x_1' + (1-\gamma)x_2'}{f(\theta)}\right] \\ \leq& \gamma\mathcal{L}(\theta, x_1, I_1, \lambda, \mu) + (1-\gamma)\mathcal{L}(\theta, x_2, I_2, \lambda, \mu).\end{aligned}$$

Then by Proposition 1 in Luenberger (1969), Section 7.8, as Q^* is also continuously differentiable, it is indeed a solution for the relaxed problem.

Proof of Proposition 4: First note that, for any given F and c

$$a(\theta) = c^{-1}\left(\theta F^{n-1}(\theta) - \int_0^\theta F^{n-1}(t)dt\right)$$

is increasing in θ and goes to 0 as $n \to \infty$; and

$$a(1) = c^{-1}\left(1 - \int_0^1 F^{n-1}(t)dt\right)$$

is increasing in n and weakly smaller than $1 - \int_0^1 F^{n-1}(t)dt$. Thus I obtain

$$J(1) = \frac{1}{c'(a(1))} \leq \frac{1}{c'\left(1 - \int_0^1 F(t)dt\right)}$$

The right hand of the above inequality is independent of n. Take any $\theta \in (0, 1)$,

$$\begin{aligned} J(\theta) &= \frac{1}{c'(a(\theta))}\left[\theta - \frac{\int_\theta^1 \frac{c'(a(\theta))}{c'(a(s))}dF(s)}{f(\theta)}\right] \\ &\geq \frac{1}{c'(a(\theta))}\left[\theta - \frac{1 - F(\theta)}{f(\theta)}\right] \end{aligned}$$

In obtaining the inequality I used $a(\theta) \leq a(s)$ for any $s > \theta$ and c is convex. Since $a(\theta) \to 0$ as $n \to \infty$, c' is continuous and $c'(0) = 0$, there exists $\bar{n}$ such that for any $n > \bar{n}$,

$$J(\theta) \geq \frac{1}{c'(a(\theta))}\left[\theta - \frac{1 - F(\theta)}{f(\theta)}\right] \geq \frac{1}{c'\left(1 - \int_0^1 F(t)dt\right)} \geq J(1)$$

That is, Assumption R fails.

Proof for Theorem 2: The proof consists of 3 steps.

Step 1: Letting Q^* denote the solution to the designer's relaxed problem (with only constraints (c) and (d)). As shown in the proof for Theorem 1, Q^* must satisfy the following necessary conditions:

$$\frac{dJ(\theta \mid Q^*)}{d\theta} = \lambda(\theta),$$

$\lambda(\theta) \geq 0$, and the complementary slackness condition

$$\lambda(\theta)\left(x(\theta) - \int_\theta^1 F^{n-1}(t)dF(t)\right) = 0.$$

whenever these conditions are defined.

The above conditions are equivalent to requiring, for any θ at which Q^* is differentiable (and thus $J(\theta \mid Q)$ is differentiable), one of the following two statements holds:

(1)

$$\frac{dJ(\theta \mid Q^*)}{d\theta} \geq 0$$

and

$$Q^*(\theta) = F^{n-1}(\theta).$$

or (2)

$$\frac{dJ(\theta \mid Q^*)}{d\theta} = \lambda(\theta) = 0,$$

and

$$Q^*(\theta) \neq F^{n-1}(\theta).$$

Step 2: I further show that Q^* does not have any jump discontinuous point. Let (x^*, I^*) denote the solution to Problem (P) (defined in the proof for Theorem 1). To show Q^* does not have any jump discontinuous point is equivalent to showing $(x^*)'$ has no such points.

Suppose there exists $l \in [0,1]$ at which $(x^*)'(l^-) \neq (x^*)'(l^+)$. Then for (x^*, I^*) to be optimal, the Weierstrass-Erdmann corner conditions have to be satisfied at this point l:

$$\frac{\partial h}{\partial x'} \mid_{(x^*)'(l^-)} = \frac{\partial h}{\partial x} \mid_{(x^*)'(l^-)}$$

and

$$[x' \frac{\partial h}{\partial x'} - h] \mid_{(x^*)'(l-)} = [x' \frac{\partial h}{\partial x'} - h] \mid_{(x^*)'(l+)}$$

The first condition is equivalent to

$$c'\left(c^{-1}\left(-l\frac{(x^*)'(l^-)}{f(l)} - I(l)\right)\right) = c'\left(c^{-1}\left(-l\frac{(x^*)'(l^+)}{f(l)} - I(l)\right)\right)$$

As both c^{-1} and c' are strictly increasing, the above inequality implies $(x^*)'(l^-) = (x^*)'(l^+)$

which contradicts the assumption $(x^*)'(l^-) \neq (x^*)'(l^+)$. So such l cannot exist. This proves $(x^*)'$, and thus also Q^*, does not contain any jump discontinuous point.

Step 1 and Step 2 together imply that Q^* only contains the two regions described in Theorem 2.

Step 3: To show that Q^* is also a solution to the original problem, one needs to in addition show that (a) Q^*it is non-decreasing and that (b) $Q^*(\theta) \in [0,1]$ for all θ.
(a) As Q^* has no jump discontinuous point and $F^{n-1}(\theta)$ is non-decreasing, it suffices to show the solution to $\frac{dJ(\theta|Q)}{d\theta}$ is also non-decreasing. Note that

$$\begin{aligned}
\frac{dJ(\theta \mid Q)}{d\theta} =& \frac{d\theta a'(\theta \mid Q)}{d\theta} + a'(\theta \mid Q) + \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 a'(s \mid Q) dF(s) \\
=& 2a'(\theta \mid Q) + \frac{f'(\theta)}{f^2(\theta)} \int_\theta^1 \frac{1}{c'(a(s \mid Q))} dF(s) + \frac{da'(\theta \mid Q)}{d\theta} \\
=& \left(c^{-1}\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)\right)'' \theta Q'(\theta) + \frac{2}{c'(a(\theta \mid Q))} + \frac{f'(\theta)}{f^2(\theta)} \int_\theta^1 \frac{1}{c'(a(s \mid Q))} dF(s)
\end{aligned}$$

Suppose Q^* is *not* monotone non-decreasing. As Q^* has no jump discontinuous point and is piecewise differentiable, there must exists θ such that $(Q^*(\theta))'$ is negative. Also, c^{-1} is strictly concave. So the first term is positive at θ.

The second term is always positive. If $f'(\theta) > 0$, then last term is also positive. I obtain

$$\frac{dJ(\theta \mid Q)}{d\theta} > 0$$

and $Q^*(\theta) \neq F^{n-1}(\theta)$, which contradicts the necessary conditions given in Step 1. If $f'(\theta) < 0$, by Assumption M, I also obtain

$$\begin{aligned}
\frac{dJ(\theta \mid Q)}{d\theta} \geq& \frac{2}{c'(c^{-1}(1))} + \frac{f'(\theta)[1 - F(\theta)]}{c'(0)f^2(\theta)} + \left(c^{-1}\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)\right)'' \theta Q'(\theta) \\
>& \left(c^{-1}\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)\right)'' \theta Q'(\theta) > 0
\end{aligned}$$

By the same arguments as above, I conclude this also contradicts the assumption that Q^* is the solution to the relaxed problem. Thus, Q^* must be non-decreasing.

(b) I show that any Q^* which satisfies constraints (a), (c), (d) must also satisfy constraint

(b) $Q^*(\theta) \in [0,1]$ for all θ.

Suppose $Q^*(\theta) < 0$ on a set of positive measure. As Q^* is non-decreasing, there must exist ρ such that $Q^*(\theta) < 0$ on $[0,\rho]$. Constraint (d)

$$\int_0^1 Q(t)dF(t) = \frac{1}{n}.$$

then requires

$$\int_\rho^1 Q(t)dF(t) > \frac{1}{n} = \int_0^1 F^{n-1}(t)dF(t)$$

which contradicts the interim feasibility constraint (c).

Similarly, suppose there exists ρ' such that $Q^*(\theta) > 1$ on $[\rho', 1]$. This implies

$$\int_{\rho'}^1 Q(t)dF(t) > \int_{\rho'}^1 F^{n-1}(t)dF(t)$$

which contradicts constraint (c).

As Q^* satisfies constraints (a)-(d), I conclude it is also a solution to the original problem.

Proof for Proposition 5: Take any two convex functions c_1 and c_2 for which there exists a non-decreasing and convex function $K(\cdot)$ such that $c_1 = K(c_2)$. Let $J_{c_1}(\theta)$ and $J_{c_2}(\theta)$ denote the corrsponding generalized virtual value functions. To prove Proposition 5, it suffices to show $J'_{c_2}(\theta) \geq 0$ implies $J'_{c_1}(\theta) \geq 0$.

Recall that by definition, I have

$$J_{c_1}(\theta) = \frac{\theta}{c'_1(a(\theta))} - \int_\theta^1 \frac{1}{c'_1(a(s))}ds$$

Thus

$$
\begin{aligned}
J'_{c_1}(\theta) &= \frac{2}{c_1'(a(\theta))} - \frac{\theta c_1''(a(\theta))a'(\theta)}{[c_1'(a(\theta))]^2} \\
&= \frac{2c_1'(a(\theta)) - \theta a'(\theta)c_1''(a(\theta))}{[c_1'(a(\theta))]^2} \\
&= \frac{2K'(c_2(a(\theta))c_2'(a(\theta)) - \theta a'(\theta)[K'(c_2(a(\theta))c_2''(a(\theta)) + K''(c_2(a(\theta)))(c_2'(a(\theta)))^2]}{[c_1'(a(\theta))]^2} \\
&\geq \frac{2K'(c_2(a(\theta))c_2'(a(\theta)) - \theta a'(\theta)K'(c_2(a(\theta))c_2''(a(\theta))}{[c_1'(a(\theta))]^2} \\
&\geq \left[\frac{c_2'(a(\theta))}{c_2'(a(\theta))}\right] J'_{c_2}(\theta)
\end{aligned}
$$

It is then clear that, $J'_{c_2}(\theta) \geq 0$ implies $J'_{c_1}(\theta) \geq 0$, as desired.

Proof for Proposition 6: For this case, we have

$$
\frac{dJ(\theta \mid Q)}{d\theta} = \frac{4\left[\theta Q(\theta) - I(\theta)\right] - \theta^2 Q'(\theta)}{4\left[\theta Q(\theta) - I(\theta)\right]^{\frac{3}{2}}}.
$$

As $4\left[\theta Q(\theta) - I(\theta)\right]^{\frac{3}{2}}$ is always positive, the regularity condition holds if and only if

$$
\begin{aligned}
&4\left[\theta F^{n-1}(\theta) - \int_{\frac{1}{2}-k}^{\theta} F^{n-1}(t)dt\right] \geq (n-1)\theta^2 F^{n-2}(\theta)f(\theta) \\
\iff\ & \theta\left[\frac{\theta - (\frac{1}{2}-k)}{2k}\right]^{n-1} - \int_{\frac{1}{2}-k}^{\theta}\left[\frac{t-(\frac{1}{2}-k)}{2k}\right]^{n-1} dt \geq \frac{(n-1)\theta^2}{8k}\left[\frac{\theta-(\frac{1}{2}-k)}{2k}\right]^{n-2} \\
\iff\ & \theta\left[\theta - (\frac{1}{2}-k)\right] - \frac{1}{n}\left[\theta - (\frac{1}{2}-k)\right]^2 \geq \frac{(n-1)\theta^2}{4} \\
\iff\ & \frac{1}{n}\left[\theta - (\frac{1}{2}-k)\right]\left[(n-1)\theta + (\frac{1}{2}-k)\right] \geq \frac{(n-1)\theta^2}{4}
\end{aligned}
$$

The right hand side is independent of k. As $n \geq 2$, $(n-1)\theta \geq \theta$. It follows that the left hand decreases in $\frac{1}{2} - k$, and thus increases in k. Thus, the above inequality is more likely to hold when k increases.

That is, Assumption R is more likely to hold as k increases, as desired.

Section 4.2 Example: Given the prize allocation rule described in the example, with N

agents, the expected prizes earned by a type-θ agent equals

$$Q_N(\theta) = \sum_{k=1,2...N} \frac{(N-1)!}{(N-k)!(k-1)!}\theta^{N-k+1}(1-\theta)^{k-1}x_N^k$$

Recall that, by definition, $x_n^k = 0$ for any $k > n$ or $k = 0$.

When $n = 5$, with $x_5^1 = \frac{4}{5}$, $x_5^2 = \frac{1}{5}$, and $x_5^3 = x_5^4 = x_5^5 = 0$, it is easy to verify that $\sum_{k=1,2...5} x_5^k = 1$ and the expected prizes earned by a type-θ agent equals

$$\frac{4}{5}\theta^4 + \frac{1}{5} \times \frac{4!}{3!1!}\theta^3(1-\theta) = \frac{4}{5}\theta^3$$

as desired.

I next show that if $\sum_{k=1,2...N} x_N^k = 1$ and

$$Q_N(\theta) = \sum_{k=1,2...N} \frac{(N-1)!}{(N-k)!(k-1)!}\theta^{N-k+1}(1-\theta)^{k-1}x_N^k = \frac{4}{n}\theta^3$$

holds for $N = n-1$, then it also holds for $N = n$. For $N = n$,

$$\begin{aligned}
Q_n(\theta) = & \sum_{k=1,2...n} \frac{(n-1)!}{(n-k)!(k-1)!}\theta^{n-k+1}(1-\theta)^{k-1}x_n^k \\
= & \sum_{k=1,2...n} \frac{(n-1)!}{(n-k)!(k-1)!}\theta^{n-k+1}(1-\theta)^{k-1}\left[\frac{n-1}{n}x_{n-1}^k + \frac{k-1}{n}\left(x_{n-1}^{k-1} - x_{n-1}^k\right)\right] \\
= & \sum_{k=1,2...n} \frac{(n-1)!}{(n-k)!(k-1)!}\theta^{n-k+1}(1-\theta)^{k-1}\left[\frac{n-k}{n}x_{n-1}^k + \frac{k-1}{n}x_{n-1}^{k-1}\right] \\
= & \frac{n-1}{n}\theta \sum_{k=1,2...n-1} \frac{(n-2)!}{(n-k-1)!(k-1)!}\theta^{(n-1)-k+1}(1-\theta)^{k-1}x_{n-1}^k \\
& + \frac{n-1}{n}(1-\theta) \sum_{k=2,3...n} \frac{(n-2)!}{(n-k)!(k-2)!}\theta^{n-k+1}(1-\theta)^{k-2}x_{n-1}^{k-1} \\
= & \frac{n-1}{n}\theta \sum_{k=1,2...n-1} \frac{(n-2)!}{(n-k-1)!(k-1)!}\theta^{(n-1)-k+1}(1-\theta)^{k-1}x_{n-1}^k \\
& + \frac{n-1}{n}(1-\theta) \sum_{s=1,2...n-1} \frac{(n-2)!}{(n-s-1)!(s-1)!}\theta^{(n-1)-s+1}(1-\theta)^{s-1}x_{n-1}^s \\
= & \frac{n-1}{n}\frac{4}{n-1}\theta^3 = \frac{4}{n}\theta^3
\end{aligned}$$

as desired. Note that I used $n-k=0$ when $k=n$ and $k-1=0$ when $k=1$ to obtain the fourth equality; I used change of variable $k=s+1$, $x_0^{n-1}=0$ and $x_{n-1}^k=0$ for $k>n-3$ to obtain the fifth equality.

The last step is to check

$$
\begin{aligned}
\sum_{k=1,2\ldots n} x_n^k &= \sum_{k=1,2\ldots n} \frac{n-1}{n}x_{n-1}^k + \frac{k-1}{n}\left(x_{n-1}^{k-1} - x_{n-1}^k\right) \\
&= \sum_{k=1,2\ldots n-1} \frac{n-k}{n}x_{n-1}^k + \sum_{s=1,2\ldots n-1} \frac{s}{n}x_{n-1}^s \\
&= \sum_{k=1,2\ldots n-1} x_{n-1}^k = 1
\end{aligned}
$$

also holds.

References

[1] Baisa, Brian. "Auction design without quasilinear preferences." Theoretical Economics 12, no. 1 (2017): 53-78.

[2] Chawla, Shuchi, Jason D. Hartline, and Balasubramanian Sivan. "Optimal crowdsourcing contests." Games and Economic Behavior 113 (2019): 80-96.

[3] Drugov, Mikhail, and Dmitry Ryvkin. "Tournament rewards and heavy tails." Journal of Economic Theory 190 (2020): 105116.

[4] Fang, Dawei, Thomas Noe, and Philipp Strack. "Turning up the heat: The discouraging effect of competition in contests." Journal of Political Economy 128, no. 5 (2020): 1940-1975.

[5] Fu, Qiang, and Jingfeng Lu. "The optimal multi-stage contest." Economic Theory 51, no. 2 (2012).

[6] Gershkov, Alex, Benny Moldovanu, Philipp Strack, and Mengxi Zhang. "Optimal auctions: Non-expected utility and constant risk aversion." The review of economic studies 89, no. 5 (2022): 2630-2662.

[7] Glazer, Amihai, and Refael Hassin. "Optimal contests." Economic Inquiry 26, no. 1 (1988): 133-143.

[8] Halmos, Paul and John von Neumann (1942). Operator Methods in Classical Mechanics, II. *Annals of Mathematics*, 332-350.

[9] Konrad, Kai A. (2009). *Strategy and dynamics in contests*. Oxford University Press.

[10] Lazear, Edward P., and Sherwin Rosen. "Rank-order tournaments as optimum labor contracts." Journal of political Economy 89, no. 5 (1981): 841-864.

[11] Letina, Igor, Shuo Liu, and Nick Netzer. "Delegating performance evaluation." Theoretical Economics 15, no. 2 (2020): 477-509.

[12] Letina, Igor, Shuo Liu, and Nick Netzer. "Optimal contest design: Tuning the heat." Journal of Economic Theory (2023): 105616.

[13] Liu, Bin, Jingfeng Lu, Ruqu Wang, and Jun Zhang. "Optimal prize allocation in contests: The role of negative prizes." Journal of Economic Theory 175 (2018): 291-317.

[14] Luenberger, David G. (1997). *Optimization by vector space methods*. John Wiley & Sons.

[15] Manelli, Alejandro M., and Daniel R. Vincent. "Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model." Econometrica 78, no. 6 (2010): 1905-1938.

[16] Matthews, Steven. "Comparing auctions for risk averse buyers: A buyer's point of view." Econometrica: Journal of the Econometric Society (1987): 633-646.

[17] Matthews, Steven. "Selling to risk averse buyers with unobservable tastes." Journal of Economic Theory 30, no. 2 (1983): 370-400.

[18] Moldovanu, Benny, and Aner Sela. "The optimal allocation of prizes in contests." American Economic Review 91, no. 3 (2001): 542-558.

[19] Moldovanu, Benny, and Aner Sela. "Contest architecture." Journal of Economic Theory 126, no. 1 (2006): 70-96.

[20] Maskin, Eric, and John Riley. "Optimal auctions with risk averse buyers." Econometrica (1984): 1473-1518..

[21] Myerson, Roger B. "Optimal auction design." Mathematics of operations research 6, no. 1 (1981): 58-73.

[22] Olszewski, Wojciech, and Ron Siegel. "Performance-maximizing large contests." Theoretical Economics 15, no. 1 (2020): 57-88.

[23] Olszewski, Wojciech, and Ron Siegel. "Large contests." Econometrica 84, no. 2 (2016): 835-854.

[24] Polishchuk, Leonid, and Alexander Tonis. "Endogenous contest success functions: a mechanism design approach." Economic Theory 52 (2013): 271-297.

[25] Riley, John G., and William F. Samuelson. "Optimal auctions." The American Economic Review 71, no. 3 (1981): 381-392.

[26] Sagan, Hans. (1992): *Introduction to the Calculus of Variations*, New York: McGraw-Hill.

[27] Schweinzer, Paul, and Ella Segev. "The optimal prize structure of symmetric Tullock contests." Public Choice 153 (2012): 69-82.