Optimal Contests with Incomplete Information and Convex Effort Costs*

Mengxi Zhang†

April 23, 2021

Abstract

I investigate the optimal design of contests when contestants have both private information and convex effort costs. The designer has a fixed prize budget and her objective is to maximize the expected total effort. I first demonstrate that it is always optimal for the designer to employ a grand static contest with as many participants as possible. Further, I identify a sufficient and necessary 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. Lastly, I illustrate how the optimal prize distribution evolves as contest size grows: the prize distribution first becomes more unequal until the optimal level of competition intensity is obtained and then becomes less unequal to maintain the optimal intensity.

*A previous version has been circulated under the title "Contest with Incomplete Information: When to Turn Up the Heat, and How?"

†I owe special thanks to Benny Moldovanu for many insightful discussions and detailed comments. I would also like to thank Cedric Wasser, Zenan Wu, and the participants at 2019 Nanjing International Conference on Theory, 2020 Conference on Mechanism and Institution Design and 2020 European Winter Meeting for helpful conversations and comments. Financial support from the German Research Foundation(Deutsche Forschungsgemeinschaft, 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 compete for a fixed amount of prizes, are widely used in practice to increase participants' performance (e.g., promotion contests, innovation contests, sport contests). It is therefore important to understand how contestants' behavior is shaped by various aspects of the contest's structure and rules. In this paper, I investigate the optimal design of contests for the arguably most relevant case – where contestants are privately informed about their abilities or valuations of the prize and have convex effort costs. Following the tradition in the literature (see Konrad (2009) for a survey), I assume the designer has a fixed prize budget and her objective is to maximize the expected total effort.

The contest design problem is twofold: The designer first needs to choose how to organize the contest; then, she needs to determine how to allocate the prize. My first finding is that, perhaps surprisingly, it is *always* optimal for the designer to conduct a grand static contest with as many participants as possible, which is arguably the most competitive contest format. Previous work has found that the designer may benefit by altering the contest architecture (e.g., conducting multi-round contests or several sub-contests). This is either due to the additional restriction that the designer is not allowed to optimally choose the prize allocation rule (e.g., Moldovanu and Sela, 2006; Fang et al., 2020) or because the environment considered is significantly different from the one studied here (e.g., Fu and Lu, 2012). I further show that, if a *regularity condition* holds, then the designer also finds it optimal to implement the most competitive prize allocation rule: i.e., to always allocate the whole prize to the contestant with the highest effort (also known as *winner-takes-all*). When this condition fails, the designer may prefer to award multiple prizes of descending sizes: the contestant with the highest effort wins the largest prize, the contestant with the second-highest effort wins the second-largest prize, and so on until all the prizes are allocated. I also characterize the optimal prize allocation rule for this case.

I then construct a parameterized setting to further discuss properties of the optimal prize distribution. For instance, I find that high-performing contestants are rewarded more aggressively than low-performing ones: i.e., the expected prize amount allocated to any contestant is increasing and strictly convex in his effort. This result is consistent with the empirical findings of Zenger (1992). I also illustrate how the prize distribution evolves as contest size grows. From the designer's perspective, there exists an optimal level of competition intensity. As contest size increases, the prize distribution first becomes more unequal until the optimal level of competition intensity is obtained and then becomes less unequal to maintain the optimal intensity. The optimal competition level decreases as effort cost becomes more convex, and it increases as the contestants' ability distribution becomes more spread out.

Lastly, I discuss a generalization of the model where the designer wishes to maximize a weighted average of the expected total effort and the expected maximum effort. This generalization is relevant for situations such as innovation contests where the designer cares mostly about the quality of the best product but may also benefit from the design of other products. I again identify a condition under which the winner-takes-all prize structure is optimal in static contests. However, static contests may no longer be the optimal contest format.

Contest design with incomplete information and convex effort costs was first studied by Moldovanu and Sela (2001). The authors investigate the design of a single contest with symmetric contestants, and allow the designer to choose from a single prize or multiple prizes of descending sizes. They find that with concave or linear effort cost, a single prize is always optimal; when the effort cost is convex, multiple prizes may perform better than a single prize. The aim of this paper is not to provide a characterization of the optimal prize structure, which has since been considered as a generally intractable problem.

Olszewski and Siegel (2016, 2020) revisited this problem recently 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 the above mentioned finding to identify the optimal prize structure for contests with "sufficiently many" (either symmetric or asymmetric) contestants. They find that, with convex effort cost, it is always optimal for the designer to award many prizes of descending sizes.

In this paper, I employ a different 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 it is not without loss to only consider large contests.[1] On the other hand, my results offer a method to determine whether a specific contest should be considered "sufficiently large" and provide new insights for 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 static contest) as given and only consider the choice of prize allocation rules. In

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

this paper, I allow the designer to in addition choose the contest architecture and establish the optimality of grand static contests among all feasible contest formats.

Most papers studying contest design with incomplete information assume linear effort cost (Polishchuk and Tonis, 2013; Chawla et al., 2019; Liu et al., 2018; etc.). To illustrate the technical differences, I would like to note that previous studies (e.g., 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. With convex effort costs, I show that the contest problem can still be transformed into an auction design problem, but one with risk-averse bidders. The techniques for solving optimal auction problems with risk-neutral bidders have been well established, thanks to seminal works by Myerson (1981) and Riley and Samuelon (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 quite 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 find a modified first-price auction to be 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 always 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 here do not fell into that category. The optimal mechanisms found by these two papers are also very different from that of the current 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 establish a payoff equivalence result which maps contestants' utilities to the expected share of prize allocated to them (also known as *the reduced form allocation rule*). I then show that the search for a designer-optimal mechanism can be confined to the class of static grand contests, where each contestant's effort only depends on his type and is independent of the contest's outcome. The above results allow me to rewrite the designer's objective as a function of the reduced form prize allocation rule. However, unlike in Myerson (1981), the objective function here is not linear in the reduced form allocation rule. Thus, I cannot rewrite the objective function as a function of the ex-post allocation rule and then maximize it pointwise, as is usually done for contests with linear effort costs. Instead, I have to perform the maximization subject to the n-agent reduced form feasibility constraint. The problem thus becomes mathematically quite different and has "so far proven intractable" (Olszewski and Siegel, 2020). I demonstrate that, with symmetric agents, the problem can be tackled through a variational analysis approach, in particular, by utilizing the infinite-dimensional version of the Kuhn–Tucker Theorem (see Luenberger, 1997). The approach does not apply to the asymmetric case as the reduced form feasibility constraint becomes more complicated for that case. This is consistent with the observation that the analysis of optimal auctions with asymmetric risk-averse bidders is largely absent from the literature. This additional difficulty does not arise in Olszewski and Siegel (2020), as they solved a single agent problem and thus did not need to deal with the reduced form feasibility constraint.

There are also several papers studying 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 Lazear-Rosen Model. Fang, Noe, and Strack (2020) characterize the optimal prize profile for all-pay contests. Letina, Liu and Netzer (2020) take a unifying approach and allow the designer to choose both the prize profile and the contest format. With n contestants, they find it is always optimal for the designer to utilize a nested Tullock contest with $n-1$ identical prizes. The sharp contrast between my results and theirs highlights the role played by private information.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 summarizes the main results. Section 4 discusses some generalizations of the model. Section 5 concludes.

2 The Model

Consider a contest with one designer (female) and n agents (male), $n \geq 2$. 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.[2] Each agent $i \in \{1, 2, ..., n\}$ is

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

characterized by his privately-known type $\theta_i \in \Theta = [0,1]$.[3] A high θ_i means that agent i has a high valuation of the prize. It is common knowledge that agents' types are distributed 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 $[0,1]$.

In the contest, each agent i chooses an effort level $a_i \in [0,\infty)$ at cost $c(a_i)$.[4] The cost function c is strictly increasing, convex, and twice continuously differentiable. $c(0)$ is normalized to be 0. If agent i has type θ_i, chooses effort level a_i and obtains prize x_i, then his utility is given by[5]

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

The agent gets an outside option of 0 if he chooses not to participate in the contest. Note that $\frac{\partial u_i}{\partial \theta_i} > 0$,

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

and

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

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

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

[4]If the designer conducts a multi-stage contest and agent i chooses effort level a_i^t at each stage t, then the agent's total cost of effort is given by $c(\sum_t a_i^t)$.

[5]Moldovanu and Sela (2001) assume the agent's private information is about his ability and 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. I adopt the current setting to ease the comparison with the risk-averse bidder problem.

3 Main Results

3.1 Direct Mechanisms

I restrict attention to direct mechanisms. To allow for both random allocations of prizes and random effort, I introduce a random variable $\omega \in [0,1]$ to capture all randomness in the mechanism.[6] In a direct mechanism $(\boldsymbol{q}, \boldsymbol{a})$, each agent reports his type to the mechanism and a number ω is randomly drawn from $[0,1]$ according to the uniform distribution; 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 reported types and the realization of the random variable ω. Contestants will take the designer's recommendations provided that $(\boldsymbol{q}, \boldsymbol{a})$ forms an equilibrium. Note that, for any given type profile, both the prize allocation and the effort may be random as they also depend on the random number ω. In addition, the effort a_i of agent i may be random even conditional on the allocation of the good, and vice versa.

By the revelation principle, if any outcome pair $(\boldsymbol{q}, \boldsymbol{a})$ can be implemented by a contest, then it can also be implemented by a direct mechanism. Thus, the set of mechanisms being considered here include (but not limited to) all commonly used contest formats, such as:[7]

All-Pay-Contests (Noiseless Contests) : The designer may award a single prize or multiple (fixed) prizes. The allocation of prizes is purely determined by the rank of contestants' effort.

Tullock Contests (see Tullock (1980)): Letting a_i denote the effort exerted by agent i in the contest. Agent i is awarded a prize of size

$$x_i = \begin{cases} \frac{a_i^r}{\sum_{j=1,2\ldots n} a_j^r} & \text{if at least one } a_j \neq 0 \\ \frac{1}{n} & \text{if all } a_j = 0, \end{cases}.$$

where $r > 1$ is a constant.

Lazear-Rosen Model (see Lazear and Rosen (1981)): Agent i receives a score

$$y_i = a_i + \varepsilon_i$$

[6]This is without loss of generality by Halmos and von Neumann (1942).

[7]I slightly adjusted the interpretation of some contest rules to make them consistent with the current framework.

where ε_i is a random variable with mean 0. Agent i obtains the whole prize if y_i is the highest among all y_j and ties are broken randomly.

All variations of the above-described contests (e.g., multi-round contests or sub-contests) are also included.

On the other hand, there may exist implementable direct mechanisms which cannot be implemented by any contests. Optimizing over all implementable direct mechanism is thus a relaxed version of the contest design problem. In what follows, I will first solve for this relaxed problem. I then verify that the optimal direct mechanism can indeed be implemented by a contest, so it is also a solution to the original problem.

The following proposition identifies sufficient and necessary conditions for a direct mechanism to be implementable (i.e. incentive compatible and individual rational). 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 prize allocated to agent i, given that he is of type θ_i. Q_i is also referred to as *the reduced form allocation rule.* I say $\boldsymbol{Q} = (Q_i)_{i=1,2...N}$ is *feasible* if and only if it can be induced by a feasible ex-post allocation rule $\boldsymbol{q}$; I say a feasible $\boldsymbol{Q} = (Q_i)_{i=1,2...N}$ is *implementable* if and only if it can be implemented by an implementable $(\boldsymbol{q}, \boldsymbol{a})$.

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

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

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

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

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

The above proposition states that agents' utilities in any implementable mechanism are essentially uniquely determined by the reduced form allocation rule, as in the risk-neutral bidder problem. However, the other classical result, revenue equivalence, does not hold except when effort cost is linear. The designer's revenue does not only depend on $\boldsymbol{Q}$, but may also depend on the exact details of the mechanism. In the next subsection, I demonstrate that, among all implementable mechanisms, static grand contests are optimal for the designer.

3.2 Independent-Payment is Optimal

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 *independent-payment mechanism* below.

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

The following lemma shows that independent-payment mechanisms are optimal among all implementable mechanisms.

Lemma 1. *For any implementable direct mechanism* $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$*, there exists an independent-payment mechanism which implements the same allocation rule and is at least as profitable for the designer.*

This lemma claims that even if the designer can freely choose the mechanism architecture (e.g., multi-round or single round, one grand contest or several sub-contests, limiting entry) and can make the effort recommendation rule as complicated as she wishes (for instance, the agents put in effort only if they receive more than 70% of the prize), she still prefers to use a simple independent-payment rule. This conclusion stems from the convexity of the cost function and the payoff equivalence result established in the previous subsection: Suppose the designer wants to implement a certain allocation rule $\boldsymbol{q}$. Since the agents have convex effort costs and must obtain the same payoff in any implementable mechanism, the designer finds it optimal to utilize the mechanism which minimizes effort variations while maintaining incentive compatibility. This can be achieved by requiring agents to always exert an effort level which only depends on his type.

In order to search for the optimal mechanism, I can confine my attention to the class of independent-payment mechanisms. For any independent-payment mechanism $(\boldsymbol{q}, \boldsymbol{a})$, let $\boldsymbol{Q}$ denote the the corresponding reduced form allocation rule and write $a_i(\theta_i \mid \boldsymbol{Q}) =$

$a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega)$. I obtain by Proposition 1:

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

and further

$$a_i(\theta_i \mid \boldsymbol{Q}) = c^{-1}\left(\theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt\right).$$

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

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

Note that, although revenue equivalence does not hold in general, it holds within this special class of mechanisms. That is, any independent-payment mechanisms which implement the same expected prize allocation rule generate the same expected revenue for the designer (see Section 3.3.2 for an example).

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

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

The second equality is obtained by using integration by parts and is the same as the revenue function obtained in Myerson (1981).

The mechanism described above can be implemented by a static grand contest. Take any implementable independent-payment mechanism $(\boldsymbol{q}, \boldsymbol{a})$ and let $\boldsymbol{Q}$ be the corresponding reduced form allocation rule. Consider a single static contest which induces the prize allocation rule $\tilde{Q}_i(\tilde{a}_i) = Q_i(\theta_i)$ if $\tilde{a}_i = a_i(\theta_i)$ and $\tilde{Q}_i(\tilde{a}_i) = 0$ otherwise. By construction, it is clear that

$$a_i(\theta_i \mid \boldsymbol{Q}) = c^{-1}\left(\theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt\right)$$

for all θ_i, are positive and constitute a pure strategy Bayesian Nash equilibrium for this game. Assuming all contestants will take the effort recommendation by the designer when such recommendations are Bayesian incentive compatible for them, $(\boldsymbol{q}, \boldsymbol{a})$ can be implemented in the contest setting. As static grand contests are optimal among all implementable direct mechanisms, I conclude that it must also be optimal among all feasible contest formats. Note that, although in this paper I mainly consider symmetric agents, this result also holds when agents are ex-ante asymmetric (see the proof for Lemma 1 for details).[8]

This finding seems to contradict that of Moldovanu and Sela (2006). They consider a similar setting and show that the designer may benefit by splitting the contestants into several sub-contests even when all contestants are symmetric. This is because, in their paper, the designer is restricted to use the winner-takes-all prize structure and thus can only reduce contest competitiveness by altering the contest architecture. My results suggest that the designer can achieve such a goal more efficiently by optimally adjusting the prize allocation rule. Therefore, when she is allowed to choose both the contest structure and the prize allocation rule, she never has an incentive to alter the former (see Section 3.4 for an example where the designer chooses to adjust the allocation rule to reduce contest competitiveness).

3.3 Optimal Allocation of Prize

As the designer's objective is concave in $(Q_i)_{i=1,2\dots n}$, I can restrict my attention to symmetric allocation probabilities and drop below all individual subscripts without causing any loss in generality.[9] In order to employ the variational approach, below I only consider piecewise continuous Q.

For later use, I introduce a new variable $I(\theta) = \int_0^\theta Q(t)dt$. The designer's objective is

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

[8] If the agents are asymmetric, the optimal allocation rule may also be asymmetric.

[9] To see this, consider the case of two agents. Suppose 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). The arguments can be easily generalized to incorporate more agents.

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)

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

(e)

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

where (a) repeats the incentive compatibility constraint identified in Proposition 1; (b) and (c), as demonstrated by Maskin and Riley (1984) and Matthews (1983), are the reduced form feasibility constraints to be considered when Q is non-decreasing; (d) requires that the designer has to allocate the whole prize; (e) defines the new variable $I(\theta)$.

Instead of solving the above problem directly, I consider a relaxed version with the same objective function but only constraints (c),(d) and (e). 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, I first define $x(\theta) = \int_\theta^1 Q(t)dF(t)$. It directly follows that $x'(\theta) = -f(\theta)Q(\theta)$ and $Q(\theta) = -\frac{x'(\theta)}{f(\theta)}$. Moreover, by the definition of I, I obtain $I(\theta) = -\int_0^\theta \frac{x'(t)}{f(t)}dt = \int_0^\theta Q(t)dt$ and $I'(\theta) = -\frac{x'(\theta)}{f(\theta)}$. Note that by construction we have $x(0) = \frac{1}{n}$, $x(1) = 0$, $I(0) = 0$, but $I(1)$ is not fixed. Lastly, let

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

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.

The problem can then by solved by employing the infinite-dimensional generalized Kuhn-Tucker Theorem (see for example Luenberger (1969), P249). The Lagrangian is given by

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

By the Kuhn-Tucker Theorem, any (x^*, I^*) which 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.[10]

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

(3) The complementary slack 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, I need to in addition impose the following "transversality condition" (see for example, Sagan (1992), P73, Theorem 2.8),

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

[10] See Sagan (1992), P124, for Euler-Lagrange conditions with multiple variable.

to ensure $I(1)$ 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.

3.3.1 Winner-Takes-All

I first identify a necessary and sufficient condition under which it is optimal for the designer to *always* award the whole prize to the highest performer (also known as winner-takes-all).

For any Q, let

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

denote the effort exerted by a type-θ agent given the prize allocation rule Q, and let

$$a(\theta) = a(\theta \mid F^{n-1}(\theta)\, \forall s \in [0,1])\,.$$

Further, let[11]

$$\begin{aligned} J(\theta \mid Q) &= \frac{\partial a(\theta \mid Q)}{\partial Q(\theta)} + \frac{\int_\theta^1 \frac{\partial a(s|Q)}{\partial I(s)} dF(s)}{f(\theta)} \\ &= \frac{\theta}{c'\left(a(\theta \mid Q)\right)} - \frac{\int_\theta^1 \frac{1}{c'(a(s|Q))} dF(s)}{f(\theta)} \end{aligned}$$

and

$$\begin{aligned} J(\theta) &= J(\theta \mid F^{n-1}(s)\, \forall s \in [0,1]) \\ &= \frac{\theta}{c'(a(\theta))} - \frac{\int_\theta^1 \frac{1}{c'(a(s))} dF(s)}{f(\theta)} \end{aligned}$$

wherever they are well defined. Note that $J(\theta)$ is always well defined and continuous except possibly at $\theta = 0$. $J(\theta)$ will be referred to as the *generalized virtual value*. I say that the *regularity condition* holds if $J(\theta)$ is non-decreasing on [0,1].

The following theorem states that when the regularity condition holds, the designer finds it optimal to always allocate the whole prize to the agent with the highest valuation, who

[11] In obtaining the second equality, I used $(c^{-1}(t))' = \frac{1}{c'(c^{-1}(t))}$.

also exerts the highest effort in the contest. With this allocation rule, a type-θ agent obtains one unit of prize with probability $F^{n-1}(\theta)$.

Theorem 1. *A sufficient and necessary condition for the winner-takes-all prize structure, i.e., $Q^*(\theta) = F^{n-1}(\theta)$ for all $\theta \in [0,1]$, being optimal is that the regularity condition holds.*

The necessary conditions presented in the last subsection are equivalent to the following two conditions:

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

whenever it is well defined, and the complementary slack condition, 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 for Theorem 1 for detail). Thus $Q^*(\theta)$ is a solution to the relaxed problem. It is then also a solution to the original problem as it satisfies all the ignored constraints.

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

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

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

$$\frac{\int_\theta^1 \frac{\partial a(s|Q)}{\partial I(s)} dF(s)}{f(\theta)} = -\frac{\int_\theta^1 \frac{1}{c'(a(s|Q))} dF(s)}{f(\theta)}$$

captures the indirect marginal cost of increasing $Q(\theta)$. To see this, note that the variable $I(s) = \int_0^s Q(t)dt$ measures the scope of the information rent: A type-s agent has more incentive to misreport when agents with lower types obtain more prizes.[12] Else the same, increasing $Q(\theta)$ will lead to an increase in $I(s)$ for any $s > \theta$, which in turn reduces the effort from any bidder whose type is above θ. Note that both the cost and the benefit depending the current allocation rule Q. Thus, $J(\theta \mid Q)$ describes the total (direct and indirect) marginal

[12] As in many other mechanism design problems, here the agents never have incentives to mis-report to be of a higher type (than they actually are).

returns of assigning more prizes to a type-θ agent, *given the current allocation rule is Q*.

Take any feasible and non-decreasing Q. If $J(\theta \mid Q)$ is increasing, then the designer can improve the mechanism by assigning more prizes to agents with higher types (if such changes are feasible). Further, if $J(\theta \mid Q)$ is increasing for all feasible Q, then the designer will find it optimal to assign the prize to the agent with the highest type as much as the feasibility constraint permits: i.e., she will always give the whole prize 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 $J(\theta)$ is non-decreasing. I would like to point out that when the cost function is linear, $J(\theta)$ is reduced to the standard virtual value found by Myerson (1981), $\theta - \frac{1-F(\theta)}{f(\theta)}$.

Moldovanu and Sela (2001) identify a sufficient and necessary condition under which winner-takes-all performs better than two descending prizes (Proposition 5, p549). This condition is weaker than the regularity condition obtained above, as the latter condition ensures winner-takes-all outperforms any 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 2. *Suppose $c'(0) = 0$. For any F and any c, there exists $\bar{n}$ such that the regularity condition fails for all $n \geq \bar{n}$,.*

Note that the additional assumptions, $c'(0) = 0$, 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.

3.3.2 Multiple Prizes

When the regularity condition fails, the winner-takes-all scheme is no longer optimal. As revenue equivalence holds within the class of independent-payment mechanisms, there may exist multiple schemes to implement the optimal prize allocation rule for this case. For instance, awarding multiple prizes deterministically is equivalent to awarding a single prize with some randomness in the allocation process. For convenience, I will assume the designer uses multiple prizes whenever applicable, and provide an example at the end of this subsection to illustrate the equivalence.

To find the optimal allocation rule for this case, some "ironing" is needed. With linear effort cost, the need for ironing only arises when the monotonicity constraint (i.e., Q is non-decreasing) becomes binding. With convex effort cost, even if the monotonicity constraint never binds, a different type of "ironing" may be required due to the concavity of the objective function. The next theorem describes the optimal (reduced form) allocation rule for this case.

Theorem 2. *Suppose Q^* is increasing and optimal. Then for any $\theta \in [0, 1]$, one of the following statements must be true:*

(a) 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);

(b) 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, (b) holds on a set of positive measure if the regularity condition fails.

Technically, the result directly follows Condition (1) and (2): the optimal Q^* may only consist of two regions, (a) where the feasibility constraint is binding and (b) where it solves

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

To see the intuition, suppose $J(\theta \mid Q^*)$ is strictly decreasing on some non-degenerate interval (α, β). As is already illustrated in the last subsection, this means that the marginal return of assigning more prizes to a higher type is smaller than the marginal return from a lower type, assuming both types fell 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 the lower type. The optimum is achieved when the marginal return becomes constant across types.

By Manelli and Vincent (2010), any Q^* as described in Theorem 2 can be induced by an ex-post allocation rule q^* which is symmetric and non-decreasing in each coordinate. That is, the optimal prize allocation can always be implemented by a modified all-pay-contest: the prizes are always awarded descendingly according to contestants' ranks, but the split of prizes, described by $q^*(\boldsymbol{\theta})$, may vary depending on the realization of the type profile. If the highest type fell into region (a), then the whole prize is allocated to the contestant with the highest type. If the highest type fell into an interval (y, z) which belongs to region (b), then he shares the prizes with any all contestants whose types also fell 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.

As already mentioned in the Introduction, Olszewski and Siegel (2020) identify the optimal prize distribution for the limit case where $n \to \infty$. In particular, their 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 (b) holds almost everywhere on $[0, 1]$.

Again, to gain a better understanding of Theorem 2, consider the example where $c(a) = a^2$ and θ is uniformly distributed on $[0, 1]$. In the last subsection, I have shown that if $n \leq 4$, the regularity condition holds and the designer finds it optimal to always award the whole 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 $[0, 1]$.[13] The solution is uniquely given by $Q(\theta) = \frac{4}{n}\theta^3$, which result in effort level $a(\theta) = \sqrt{\frac{3}{n}}\theta^2$.

As mentioned earlier, there often exist multiple schemes to implement the optimal prize allocation. For instance, the designer can implement the above mentioned allocation rule by conducting an all-pay-contest with multiple descending prizes - the number and sizes of positive prizes depend on the number of contestants. Let x_n^k denote the prize allocated to the $k-th$ highest performer when there are n contestant. For notational convenience, I also 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

[13] Here we just need to solve a standard second order differential equations,

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

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

following recursive formula (see Appendix for 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 it is optimal to award three prizes of $\frac{2}{3}$, $\frac{4}{15}$ and $\frac{1}{15}$ to the first 3 highest performers respectively. Alternatively, the designer can also induce the optimal prize allocation by first asking all contestants to put in effort, then randomly selecting 4 out of n contestants, and awarding the whole prize to the highest performer among the 4 contestants. There may exist other prize schemes that also induce the optimal prize allocation.

3.4 When to (Not) Turn Up the Heat?

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^n denote the optimal allocation rule with n 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 3.3, there must exist a symmetric allocation rule which is at least as profitable to the designer as the one constructed above. Thus, when there are $n+1$ bidders, the designer can earn at least as much as when there are n bidders by using the optimal symmetric mechanism. Moreover, she can earn strictly more when the cost function is strictly convex (and thus c^{-1} is strictly concave). Therefore, the designer's revenue is always increasing in n. However, for different levels of n, the increase in revenue may be caused by different mechanisms.

The optimal prize distribution reacts to the change of n in a more complicated way. As shown in the previous subsections, with uniformly distributed types and quadratic effort cost, the optimal reduced form allocation rule is given by

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

which generates the following effort levels

$$a(\theta) = \begin{cases} \sqrt{\frac{n-1}{n}}\theta^{\frac{n}{2}} & \text{if } n \leq 4 \\ \sqrt{\frac{3}{n}}\theta^2 & \text{if } n > 4. \end{cases}$$

For any value of n, $Q(\theta)$ is increasing and convex in θ. That is, a more able agent will in expectation win more prizes. Moreover, high ability agents are rewarded more aggressively than low ability agents (both the (expected) pay-ability ratio $\frac{Q(\theta)}{\theta}$ and the (expected) pay-output ratio $\frac{Q(\theta)}{a(\theta)}$ are strictly increasing in θ). This result is consistent with the findings of Zenger (1992): for workers within a same job level, the rewards for top workers are often disproportionately high relative to the compensation for non-top workers. In addition, observe that $Q(\theta)$ becomes steeper as n increases to 4, but then becomes flatter as n continues to grow. This observation provides a potentially testable implication of the model: the most unequal prize distribution (within-job pay distribution) should be observed in medium size contests (firms).

Relatedly, in a designer optimal contest with n contestants, the contestants' total effort cost is given by

$$n\int_0^1 \left[\theta Q(\theta) - \int_0^\theta Q(t)dt\right] d\theta = \begin{cases} 1 - \frac{2}{n+1} & \text{if } n \leq 4 \\ \frac{3}{5} & \text{if } n > 4. \end{cases}$$

and the 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. Then an increase in contest size always leads to an increase in contest competitiveness. The contestants (as a whole) work more and get 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 optimal allocation rule becomes less competitive, the contestants' total effort cost and total welfare remain constant, but the designer's revenue still increases due to more efficient "cost sharing" among contestants. Note that, however, the average effort produced by each individual contestant always decreases as n becomes larger.

The above observations lead to a natural conjecture, that is, there exists a finite level of competition intensity which is optimal for the designer. 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 to increase the contest competitiveness. When the contest size is large and the optimal competition intensity is already reached, further increasing competition among contestants only serves to discourage 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 cost. To balance the needs for "cost smoothing" and for maintaining the optimal competition intensity, the optimal prize distribution has to become less competitive as n increases. This observation provides a potential rationale for the findings of Olszewski and Siegel (2019): 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.

For 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 the social planner wants to increase the contestants' welfare by imposing an upper limit on contest sizes. Then she should either keep the limit low ($n < 4$ in this example) or do not impose any such limits at all (decreasing n from 8 to 5 does not affect the contestants' payoff but decreases the designer's revenue).

Lastly, I provide some comparative statics results. When the contestants' abilities are uniformly distributed, the regularity condition is less likely to hold when the cost function is more convex and more likely to hold as the ability distribution becomes more spread out, or equivalently, as the contest task become more skill sensitive. These results suggest that the optimal level of competition intensity increases as the cost function becomes less convex and as ability distribution becomes more spread out.

Proposition 3. *Suppose θ is uniformly distributed on $[0, 1]$. The regularity condition is less likely to hold when the cost function is more convex.*[14]

As effort cost becomes more convex, the need to induce "cost smoothing" increases and the benefit from private information elicitation decreases. Thus the optimal competition intensity decreases. As an illustration, consider the following examples. When θ is uniformly distributed and $c(a) = a$, the regularity condition holds for all n. That is, the optimal competition level is infinity. When $c(a) = a^{\frac{5}{4}}$, the regularity condition holds if and only if

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

$n \leq 8$.[15] The optimal competition intensity is reached at $n = 8$ and $Q(\theta) = \theta^7$. When $c(a) = a^2$, the optimal intensity is reached at $n = 4$ and $Q(\theta) = \theta^3$.

Proposition 4. *Suppose θ is uniformly distributed on $[\frac{1}{2}-k, \frac{1}{2}+k]$, $k \in (0, \frac{1}{2}]$, and $c(a) = a^2$. The regularity condition is more likely to hold as k increases.*

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

4 Discussion

4.1 Withholding the Prize

In the benchmark model, I assume the designer commits to always allocate the whole prize. This is a common assumption in the contest literature. In particular, when effort is observable but non-verifiable and the designer has positive valuation for the prize, there may arise a credibility issue without such a commitment: the designer will always have an ex-post incentive to dispute the performance with agents in order to retain the prize. Committing to always allocating the whole prize can help eliminate such problems.

If effort is verifiable or the designer can creditably commit to any prize allocation rule, the designer may strictly prefer to *not* always allocate the whole prize even if she has no value for it. The following proposition identifies a sufficient condition for this to happen.

Proposition 5. *Suppose $J(\theta)$ is non-decreasing on $[0,1]$ and $\inf_{\theta \in [0,1]} J(\theta) < 0$. If the designer can withhold the prize, then any optimal Q satisfies*

$$\int_0^1 Q(\theta) dF(\theta) < \frac{1}{n}.$$

The proof shows that the designer can improve her payoff by taking away some prizes (and keeping them herself) from contestants with negative generalized virtual values.

[15] For this case,

$$\frac{dJ(\theta \mid Q)}{d\theta} = \frac{8\left[\theta Q(\theta) - I(\theta)\right] - \theta^2 Q'(\theta)}{8\left[\theta Q(\theta) - I(\theta)\right]^{\frac{6}{5}}}$$

and

$$J'(\theta) = \frac{1}{8\left(\frac{n-1}{n}\theta\right)^{\frac{6}{5}}} \frac{(n-1)(8-n)}{n} \theta^n.$$

4.2 Expected Maximum Effort

In this subsection, I consider a generalization of the model where the designer's objective is to maximize a weighted average of the expected total effort and the expected maximum effort. This may be relevant for situations such as innovation contests where the designer values most about the quality of the best product but may also benefit from the design of other products, and promotion contests where the designer cares mostly about the total effort but values the winner's effort a bit more.

Suppose the designer implements a static symmetric contest and chooses the prize allocation rule Q. As the payoff equivalence result (Proposition 1) I obtained earlier holds regardless of the designer's objective, the agent's effort level still equals

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

and the expected total effort is still given by

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

Recall that in any implementable mechanism, Q must be non-decreasing which implies that a must also be non-decreasing. Thus, the maximum effort must be produced by the agent with the highest type. In other words, a type-θ agent's effort only counts if he is of the highest type, which occurs with probability $F^{n-1}(\theta)$, i.e., the probability that all other agents' types are below θ. The expected maximum effort is thus given by

$$\pi_M(Q) = n\int_0^1 a(\theta \mid Q)F^{n-1}(\theta)dF(\theta).$$

Let $\alpha \in [0,1]$ denote the weight the designer puts on the expected total effort and $1-\alpha$ is the weight she puts on the expected maximum effort. The designer's objective is to maximize

$$\pi(Q) = n\int_0^1 \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right] a(\theta \mid Q)dF(\theta)$$

Define

$$h^\alpha(\theta, x'(\theta), I(\theta)) = \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right] c^{-1}\Big(-\theta\frac{x'(\theta)}{f(\theta)} - I(\theta)\Big)f(\theta)$$

I can solve this maximization problem by repeating the same procedure as in Section 3 and replacing h with h^α. To be specific, let

$$\begin{aligned} J^\alpha(\theta \mid Q) &= \frac{\partial\left[\alpha+(1-\alpha)F^{n-1}(\theta)\right]a(\theta \mid Q)}{\partial Q(\theta)} + \frac{\int_\theta^1 \frac{\partial\left[\alpha+(1-\alpha)F^{n-1}(s)\right]a(s|Q)}{\partial I(s)} dF(s)}{f(\theta)} \\ &= \frac{\alpha+(1-\alpha)F^{n-1}(\theta)}{c'\left(a(\theta \mid Q)\right)}\theta - \frac{\int_\theta^1 \frac{\left[\alpha+(1-\alpha)F^{n-1}(s)\right]}{c'(a(s|Q))} dF(s)}{f(\theta)} \end{aligned}$$

and

$$J^\alpha(\theta) = J^\alpha(\theta \mid F^{n-1}(s)\, \forall s \in [0,1]).$$

Clearly, $J^\alpha(\theta) = J(\theta)$ when $\alpha = 1$. I then obtain a generalized version of Theorem 1.[16]

Theorem 3. *Suppose the design uses a symmetric independent-payment mechanism. A sufficient and necessary condition for $Q^*(\theta) = F^{n-1}(\theta)$, for all $\theta \in [0,1]$, being optimal is that $J^\alpha(\theta)$ is non-decreasing on $[0,1]$.*

Before concluding this section, I would like to note that, however, a static grand contest may not be the optimal contest format any more. If the designer only cares about the expected maximum effort, then winner-pay mechanisms always work better than independent-payment mechanisms.

Definition 2. *An implementable direct mechanism $(\boldsymbol{q}, \boldsymbol{b})$ is a winner-pay mechanism with deterministic payments if, for any i and any θ_i, (1) $b_i(\theta_i, \boldsymbol{\theta_{-i}}, r)$ is constant for all $\boldsymbol{\theta_{-i}}$ and all r; (2) $b_i(\theta_i, \boldsymbol{\theta_{-i}}, r) \neq 0$ if and only if $\theta_i \geq \theta_j$ for all $j \neq i$.*

Proposition 6. *Suppose $\alpha = 0$. For any implementable, symmetric independent-payment mechanism $(\boldsymbol{q}, \boldsymbol{a})$, there exists a winner-pay mechanism with deterministic payments $(\boldsymbol{q}, \boldsymbol{b})$ which is more profitable for the designer.*

Characterizing the optimal contest format for any α goes beyond the scope of this paper. Moreover, in many scenarios, more complicated contest formats are more costly to implement, and sometimes not even feasible (see Chawla et al. (2019) for a detailed discussion). Thus, the designer may still choose to conduct a static contest, especially when she puts a relatively large weight on the expected total effort. The above proposition, however, provides

[16] Chawla et al. (2019) provides a solution to the case where c is linear and $\alpha = 0$.

a potential rational for why elimination contests, where contestants pays more when they earns a larger prize, are sometimes used.

5 Conclusion

In this paper, I study the optimal design of contests when contestants have private information and convex effort costs, and the designer wants to maximize the expected total effort.

The solution is twofold: First, it is always optimal for the designer to employ the most competitive contest format: a grand static contest with as many contestants as possible. Second, I provide a sufficient and necessary condition for the most competitive prize allocation rule, winner-takes-all, being optimal. When this condition fails, the designer may find it desirable to award multiple prizes of descending sizes. I also provide a characterization of the optimal prize structure for this case. In addition, I illustrate how the optimal prize distribution evolves as the number of contestants increases.

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

My approach substantially eases the analysis of optimal contest design and I expect it to be useful to other contest/auction design frameworks. As a demonstration, I show that the approach can be easily applied to a generalization of the model where the designer wants to maximize a weighted average of the expected total effort and the expected maximum effort. It is also plausible to further generalize the model. For instance, the designer may want to maximize a weighted average of the total effort, the maximum effort, and the contestants' welfare. The designer may want to design an optimal winner-pay mechanism, etc.

Appendix:

For proofs of Proposition 1 and Lemma 1, 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: (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 $\delta \in (0, 1 - \theta_i']$,

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

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 Lemma 1: I only consider mechanisms for which $V_i(0) = 0$ as the objective is to maximize the designer's revenue. Fix any implementable direct mechanism $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$, I construct an independent-payment mechanism $(\boldsymbol{q}, \boldsymbol{a})$, where $a_i(\theta_i, \boldsymbol{\theta_{-i}}, \omega) = a_i(\theta_i)$ is the solution to

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

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

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

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

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,

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

which further implies

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

That is, I have shown the new mechanism generates at least as much expected total effort as the original one, as desired.

Proof for Theorem 1 & Theorem 3: I write a proof for Theorem 3, which includes Theorem 1 as a special case.

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:

$$\max_Q \int_0^1 \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right] c^{-1}\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right) dF(\theta)$$

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)

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

(e)

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

The above problem can be transformed into a calculus of variations problem. To do this, I first define $x(\theta) = \int_\theta^1 Q(t)dF(t)$. It directly follows that $x'(\theta) = -f(\theta)Q(\theta)$ and $Q(\theta) = -\frac{x'(\theta)}{f(\theta)}$. Moreover, by the definition of I, I obtain $I(\theta) = -\int_0^\theta \frac{x'(t)}{f(t)}dt = \int_0^\theta Q(t)dt$ and

$I'(\theta) = -\frac{x'(\theta)}{f(\theta)}$. Note that by construction $I(0) = 0$, but $I(1)$ is not fixed. Lastly, let

$$h^{\alpha}(\theta, x'(\theta), I(\theta)) = \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right] c^{-1}\left(-\theta\frac{x'(\theta)}{f(\theta)} - I(\theta)\right) f(\theta)$$

Then, the designer's (relaxed) problem becomes

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

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

$$x(\theta) \leq \int_\theta^1 F^{n-1}(t)dF(t), \tag{3}$$

$$I'(\theta) + \frac{x'(\theta)}{f(\theta)} = 0, \tag{4}$$

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

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

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 slack 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:[17]

[17]See Sagan (1992), P124, for Euler-Lagrange conditions with multiple variable.

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

$$\frac{\partial \mathcal{L}}{\partial I} - \frac{d\frac{\partial \mathcal{L}}{\partial I'(\theta)}}{d\theta} = 0$$
$$\Rightarrow -\frac{\partial h^\alpha}{\partial I} - \mu'(\theta) = 0$$
$$\Rightarrow \mu'(\theta) = -\frac{\partial h^\alpha}{\partial I}$$

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):

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

which implies,

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

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

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d\frac{\partial \mathcal{L}}{\partial x'}}{d\theta} = 0$$
$$\Rightarrow \lambda(\theta) - \left[-\frac{d\frac{\partial h^\alpha}{\partial x'}}{d\theta} + \left(\frac{\mu(\theta)}{f(\theta)}\right)' \right] = 0$$
$$\Rightarrow \lambda(\theta) + \frac{d\frac{\partial h^\alpha}{\partial x'}}{d\theta} - \frac{\mu'(\theta)f(\theta) - \mu(\theta)f'(\theta)}{f^2(\theta)} = 0$$
$$\Rightarrow \lambda(\theta) + \frac{d\frac{\partial h^\alpha}{\partial x'}}{d\theta} + \frac{\frac{\partial h^\alpha}{\partial I}}{f(\theta)} - \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 \frac{\partial h^\alpha}{\partial I}(t)dt = 0 \tag{5}$$

Recall that by definition,

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

and

$$J^{\alpha}(\theta \mid Q) = \frac{\partial \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right] a(\theta \mid Q)}{\partial Q(\theta)} + \frac{\int_{\theta}^{1} \frac{\partial \left[\alpha + (1-\alpha)F^{n-1}(s)\right] a(s|Q)}{\partial I(s)} dF(s)}{f(\theta)}.$$

Then Equation (5) implies

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

Thus the assumption that $J^{\alpha}(\theta)$ is non-decreasing almost everywhere on $[0,1]$ directly implies that the left hand side of Equation (5), evaluated at Q^*, must be non-negative almost everywhere 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 convex in (x, x', I, I'): take any (x_1, x_1', I_1, I_1'), (x_2, x_2', I_2, I_2') and $\gamma \in [0,1]$,

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

Thus

$$\begin{aligned}
\mathcal{L}(\theta, x, I, \lambda^*, \mu^*) \geq& \mathcal{L}(x^*, I^*, \lambda^*, \mu^*) + \frac{\partial \mathcal{L}(\theta, x^*, I^*, \lambda^*, \mu^*)}{\partial x}(x - x^*) + \frac{\partial \mathcal{L}(\theta, x^*, I^*, \lambda^*, \mu^*)}{\partial x'}(x' - (x^*)') \\
&+ \frac{\partial \mathcal{L}(\theta, x^*, I^*, \lambda^*, \mu^*)}{\partial I}(I - I^*) + \frac{\partial \mathcal{L}(\theta, x^*, I^*, \lambda^*, \mu^*)}{\partial I'}(I' - (I^*)')
\end{aligned}$$

As

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

for any (x,I) satisfying constraints (3) and (4), and

$$\mathcal{L}(\theta,x^*,I^*,\lambda^*,\mu^*) = -\,h^\alpha(\theta,(x^*)'(\theta),I^*(\theta))),$$

I obtain

$$\begin{aligned}
Y(x^*,I^*) - Y(x,I) &= \int_0^1 h^\alpha(\theta,(x^*)'(\theta),I^*(\theta)))d\theta - \int_0^1 h^\alpha(\theta,x'(\theta),I(\theta))d\theta \\
&\geq \int_0^1 \left[\mathcal{L}(\theta,x,I,\lambda^*,\mu^*) - \mathcal{L}(\theta,x^*,I^*,\lambda^*,\mu^*)\right] d\theta \\
&\geq \int_0^1 \Big[\frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial x}(x-x^*) + \frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial x'}(x'-(x^*)') \\
&\quad + \frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial I}(I-I^*) + \frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial I'}(I'-(I^*)')\Big] d\theta. \quad (6)
\end{aligned}$$

Moreover, by using integration by parts, I obtain

$$\begin{aligned}
&\int_0^1 \frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial x'}(x'-(x^*)')d\theta \\
=&\frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial x'}(x-x^*)\mid_0^1 - \int_0^1 \frac{d\frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial x'}}{dt}(x-x^*)d\theta \\
=&-\int_0^1 \frac{d\frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial x'}}{dt}(x-x^*)d\theta
\end{aligned}$$

The second equality follows $x^*(1) = x(1) = 0$ and $x^*(0) = x(0) = \int_0^1 Q(t)dt$.

Similarly,

$$\int_0^1 \frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial I'}(I'-(I^*)')d\theta = -\int_0^1 \frac{d\frac{\partial\mathcal{L}(x^*,I^*,\lambda^*,\mu^*)}{\partial I'}}{dt}(I-I^*)d\theta$$

Thus Inequality (6) can be rewritten as

$$Y(x^*, I^*) - Y(x, I) \geq \int_0^1 \Big[\left(\frac{\partial \mathcal{L}(x^*, I^*, \lambda^*, \mu^*)}{\partial x} - \frac{d \frac{\partial \mathcal{L}(x^*, I^*, \lambda^*, \mu^*)}{\partial x'}}{d\theta} \right) (x - x^*)$$
$$+ \left(\frac{\partial \mathcal{L}(x^*, I^*, \lambda^*, \mu^*)}{\partial I} - \frac{d \frac{\partial \mathcal{L}(x^*, I^*, \lambda^*, \mu^*)}{\partial I'}}{dt} \right) (I - I^*) \Big] d\theta = 0$$

I thus conclude that Q^* is a global maximizer for the designer's problem, as desired.

Proof of Proposition 2: First note that

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

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

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

is increasing in n. 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 $n \to 0$, 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, the regularity condition fails.

Proof for Theorem 2: As shown in the proof for Theorem 3, if Q^* is optimal, then it must satisfy the following necessary conditions:

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

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

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

It directly follows that, for any θ, one of the following two statements must hold:
(1)

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

and (2) $\lambda(\theta) > 0$ and $x(\theta) - \int_\theta^1 F^{n-1}(t) dF(t) = 0$.

That is, the optimal Q^* only consists of two areas, where $J(\theta \mid Q^*)$ is constant, and where the feasibility constraint becomes binding and $J(\theta \mid Q^*)$ is non-decreasing, as desired.

Proof for Proposition 3: Take any two convex functions c_1, c_2, and define $J_{c_1}(\theta \mid Q)$ and $J_{c_2}(\theta \mid Q)$ accordingly. Suppose there exists a non-decreasing and convex function $K(\cdot)$ such that $c_1 = K(c_2)$. To prove Proposition 3, 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.

Section 3.3 Example: Given the prize allocation rule described in the example, with N agents, the expected prizes earned by a type-θ agent equals

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

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 $\sum_{k=1,2\ldots 5} x_5^k = 1$ 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 $\sum_{k=1,2\ldots N} x_N^k = 1$ and

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

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

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

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 as desired.

Proof for Proposition 4: 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, the regularity condition is more likely to hold as k increases, as desired.

Proof for Proposition 5: By Theorem 1, $Q^*(\theta) = F^{n-1}(\theta)$ for all $\theta \in [0,1]$ is optimal among all feasible and increasing Q which satisfy

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

To prove Proposition 5, it suffices to show there exists a feasible and increasing $\tilde{Q}$ which outperforms Q^*.

As $\inf_{\theta\in[0,1]} J(\theta) < 0$ and $J(\theta)$ is continuous on $(0,1]$, there exists $\delta > 0$ such that $J(\theta) < 0$ on $(0,\delta)$.

Take any $\varepsilon \geq 0$. Consider the allocation rule Q^ε such that $Q^\varepsilon(\theta) = (1-\varepsilon)F^{n-1}(\theta)$ for $\theta \in [0,\delta)$ and $Q^\varepsilon(\theta) = F^{n-1}(\theta)$ otherwise. Any such Q^ε is clearly non-decreasing and feasible. If the designer implements Q^ε, the expected total effort is given by

$$R(Q^\varepsilon) = \int_0^\delta c^{-1}\left((1-\varepsilon)\theta F^{n-1}(\theta) - (1-\varepsilon)\int_0^\theta F^{n-1}(t)dt\right)dF(\theta) + \int_\delta^1 c^{-1}\left(\theta F^{n-1}(\theta) - (1-\varepsilon)\int_0^\delta F^{n-1}(t)dt - \int_\delta^\theta F^{n-1}(t)dt\right)dF(\theta)$$

Taking derivative with respect to ε yields

$$\frac{\partial R(Q^\varepsilon)}{\partial \varepsilon} = -\int_0^\delta \frac{\theta F^{n-1}(\theta) - \int_0^\theta F^{n-1}(t)dt}{c'\left(c^{-1}\left((1-\varepsilon)\theta F^{n-1}(\theta) - (1-\varepsilon)\int_0^\theta F^{n-1}(t)dt\right)\right)}dF(\theta) + \int_\delta^1 \frac{\int_0^\delta F^{n-1}(t)dt}{c'\left(c^{-1}\left(\theta F^{n-1}(\theta) - (1-\varepsilon)\int_0^\delta F^{n-1}(t)dt - \int_\delta^\theta F^{n-1}(t)dt\right)\right)}dF(\theta)$$

Thus

$$\begin{aligned}\frac{\partial R(Q^\varepsilon)}{\partial \varepsilon}\Big|_{\varepsilon=0} &= -\int_0^\delta \frac{\theta F^{n-1}(\theta) - \int_0^\theta F^{n-1}(t)dt}{c'(a(\theta))}dF(\theta) + \int_\delta^1 \frac{\int_0^\delta F^{n-1}(t)dt}{c'(a(\theta))}dF(\theta)\\ &= \int_\delta^1 \frac{1}{c'(a(\theta))}dF(\theta)\int_0^\delta F^{n-1}(t)dt - \int_0^\delta \frac{c(a(\theta))}{c'(a(\theta))}dF(\theta)\\ &= \int_0^\delta \left[F^{n-1}(\theta)\int_\delta^1 \frac{1}{c'(a(t))}dF(t) - \frac{c(a(\theta))}{c'(a(\theta))}f(\theta)\right]d\theta\end{aligned}$$

For any $\gamma \in (0, \delta]$,

$$\begin{aligned} J(\gamma) = \frac{\gamma}{c'(a(\gamma))} - \frac{\int_0^\gamma \frac{1}{c'(a(s))}dF(s)}{f(\gamma)} &< 0 \\ \iff \int_0^\gamma \frac{1}{c'(a(s))}ds &> \frac{\gamma f(\gamma)}{c'(a(r))}.\end{aligned}$$

Taking derivative of

$$\int_0^\gamma \left[F^{n-1}(\theta)\int_\gamma^1 \frac{1}{c'(a(t))}dF(t) - \frac{c(a(\theta))}{c'(a(\theta))}f(\theta)\right]d\theta$$

with respect to γ yields

$$\begin{aligned} &F^{n-1}(\gamma)\int_{\gamma}^{1}\frac{1}{c'(a(t))}dF(t)-\frac{c(a(\gamma))}{c'(a(\gamma))}f(\gamma)-\frac{f(\gamma)}{c'(a(\gamma))}\int_{0}^{\gamma}F^{n-1}(\theta)d\theta \\ &>\frac{f(\gamma)}{c'(a(\gamma))}\left[\gamma F^{n-1}(\gamma)-c(a(\gamma))-\int_{0}^{\gamma}F^{n-1}(\theta)d\theta\right]=0 \end{aligned}$$

Thus,

$$\frac{\partial R(Q^{\varepsilon})}{\partial\varepsilon}\mid_{\varepsilon=0}=\int_{0}^{\delta}\left[F^{n-1}(\theta)\int_{\delta}^{1}\frac{1}{c'(a(t))}dF(t)-\frac{c(a(\theta))}{c'(a(\theta))}f(\theta)\right]d\theta>0.$$

As $R(Q^{\varepsilon})$ is continuous in ε, there exists $\varepsilon'>0$ such that $\frac{\partial R(Q^{\varepsilon})}{\partial\varepsilon}>0$ for any $\varepsilon\in[0,\varepsilon']$. Thus $R(Q^{\varepsilon'})>R(Q^{0})=R(Q^{*})$ as desired. This completes the proof.

Proof for Proposition 6: As shown in Section 4.2, if the designer uses an independent-payment mechanism, then the expected maximum effort is given by

$$\pi_{M}(Q)=n\int_{0}^{1}F^{n-1}(\theta)c^{-1}\left(\theta Q(\theta)-I(\theta)\right)dF(\theta).$$

By Proposition 1, for the winner-pay mechanism to be implementable, the following equality must hold for all θ:

$$\int_{0}^{\theta}Q(t)dt=I(\theta)=Q(\theta)\theta-F^{n-1}(\theta)c(b(\theta))$$

which implies

$$b(\theta)=c^{-1}\left(\frac{\theta Q(\theta)-I(\theta)}{F^{n-1}(\theta)}\right)$$

Then the expected maximum effort, when the designer uses winner-pay mechanisms, is given by

$$\begin{aligned} \pi_{M}^{w}(Q)&=n\int_{0}^{1}F^{n-1}(\theta)b(\theta)dF(\theta) \\ &=n\int_{0}^{1}F^{n-1}(\theta)c^{-1}\left(\frac{\theta Q(\theta)-I(\theta)}{F^{n-1}(\theta)}\right)dF(\theta) \\ &\geq n\int_{0}^{1}F^{n-1}(\theta)c^{-1}\left(Q(\theta)\theta-I(\theta)\right)dF(\theta)=\pi_{M}(Q) \end{aligned}$$

I obtain the inequality by using $F^{n-1}(\theta) \leq 1$ for any $\theta \in [0,1]$.

References

[2017] Baisa, B. (2017). Auction design without quasilinear preference, *Theoretical Economics*, **12**(1), 53-78.

[2019] Chawla, S., Hartline, J. D., & Sivan, B. (2019). Optimal crowdsourcing contests, *Games and Economic Behavior*, **113**, 80-96.

[2020] Drugov, M., & Ryvkin, D. (2020). Tournament rewards and heavy tails. *Journal of Economic Theory*, 190, 105-116.

[2020] Fang, D., Noe, T., & Strack, P. (2020). Turning up the heat: The discouraging effect of competition in contest, *Journal of Political Economy*, **128**(5):1940-75.

[2012] Fu, Q., & Lu, J. (2012). The optimal multi-stage contest, *Economic Theory*, **51**(2), 351-382.

[2020] Gershkov, A., Moldovanu, B., Strack, P. and Zhang, M. (2021). Optimal Auctions: Non-Expected Utility and Constant Risk Aversion. Working Paper.

[1942] Halmos, P. R., & von Neumann, J. (1942). Operator Methods in Classical Mechanics, II. *Annals of Mathematics*, 332-350.

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

[1984] Maskin, E., & Riley, J. (1984). Optimal auctions with risk averse buyers. *Econometrica*, 1473-1518.

[2010] Manelli, A. M., & Vincent, D. R. (2010). Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model. *Econometrica*, 78(6), 1905-1938.

[1987] Matthews, S. (1987). Comparing auctions for risk averse buyers: A buyer's point of view. *Econometrica* 633-646.

[1983] Matthews, S. A. (1983). Selling to risk averse buyers with unobservable tastes, *Journal of Economic Theory*, **30**(2), 370-400.

[2001] Moldovanu, B., & Sela, A. (2001). The optimal allocation of prizes in contests, *American Economic Review*, **91**(3), 542-558.

[2006] Moldovanu, B., & Sela, A. (2006). Contest architecture, *Journal of Economic Theory*, **126**(1), 70-96.

[1981] Myerson, R.B. (1981) Optimal Auction Design, *Mathematics of Operations Research*,**6**(1), 58–73.

[1981] Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, **89**(5), 841-864.

[2020] Letina, I., Liu, S., & Netzer, N. (2020). Optimal Contest Design: A General Approach. Working Paper

[2018] Liu, B., Lu, J., Wang, R., & Zhang, J. (2018). Optimal prize allocation in contests: The role of negative prizes, *Journal of Economic Theory*, **175**, 291-317.

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

[2020] Olszewski, W., & Siegel, R. (2020). Performance-maximizing large contests, *Theoretical Economics*, **15**(1), pp.57-88.

[2016] Olszewski, W., & Siegel, R. (2016). Large contests. *Econometrica*, **84**(2), 835-854.

[2013] Polishchuk, L., & Tonis, A. (2013). Endogenous contest success functions: a mechanism design approach, *Economic Theory*, **52**(1), 271-297.

[1981] Riley, J. and Samuelson, W. (1981), Optimal Auctions, *American Economic Review*, **71**(3), 381-392.

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

[2012] chweinzer, P., & Segev, E. (2012). The optimal prize structure of symmetric Tullock contests. *Public Choice*, 153(1-2), 69-82.

[1980] Tullock, G. (1980). Efficient rent seeking. In James M. Buchanan, Robert D. Tollison, Gordon Tullock, (Eds.), *Toward a theory of the rent-seeking society*. College Station, TX: Texas A& M University Press, 97-112.

[2017] Xiao, J. (2017). Ability Grouping in All-Pay Contests. Working Paper.

[1987] Yaari, M. E. (1987). The Dual Theory of Choice under Risk. *Econometrica*: 95-115.

[1992] Zenger, T. R. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Administrative Science Quarterly*, 198-219.