Discussion Paper Series – CRC TR 224

Discussion Paper No. 156
Project B 01

Contest with Incomplete Information: When to Turn Up the Heat, and How?

Mengxi Zhang*

February 2020

*Institute for Microeconomics, University of Bonn, E-mail: mzhang@uni-bonn.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Contest with Incomplete Information: When to Turn Up the Heat, and How?

Mengxi Zhang*

Institute for Microeconomics, University of Bonn

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 static, grand all-pay-contest with as many as possible participants. In addition, 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 structure for this case. Lastly, I illustrate how the optimal prize structure evolves as contest size grows: the prize structure first becomes more unequal until the optimal level of competition intensity is obtained, and then becomes less unequal to maintain the optimal intensity.

1 Introduction

Contests, in which contestants compete for a fixed amount of prizes, are widely used in practice to increase participants' performance (e.g. promotion contest, innovation contest, sport contest). It is therefore important to understand how contestants' behavior in the contest are shaped by various aspects of the contest's structure and rules. In this paper,

*My special thanks goes to Benny Moldovanu for his constructive suggestions and detailed comments. Financial support from the German Science Foundation via the Hausdorff Center for Mathematics and CRC TR-224 (Project B01) are gratefully acknowledged. Contact: mzhang@uni-bonn.de.

I investigate the optimal design of contests for the arguably most relevant case, where the contestants are privately informed about their abilities and have convex effort costs. As is commonly assumed in the literature (see Konrad (2009) for a survey), the designer has a fixed prize budget to allocate to the contestants 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, and 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 implement a static, grand all-pay-contest with as many as possible participants, 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. (2019)), or because the environment considered is significantly different from the one studied here (e.g., Xiao (2017)).[1] I further show that, if a *regularity condition* holds, then the designer also finds it optimal to implement the most competitive prize allocation rule, that is, 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 provide a characterization of the optimal prize allocation rule for this case.

I then provide some examples to illustrate how the optimal prize structure evolves as contest size grows. From the designer's perspective, there exists an optimal level of competition intensity. As contest size increases, the prize structure first becomes more unequal until the optimal level of competition intensity is obtained, and then becomes less unequal to maintain the optimal intensity. I also provide some comparative statics results, which suggest that the optimal competition level decreases as effort cost becomes more convex, and 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

[1] More detailed discussion can be found in Section 4.1.

optimal in static all-pay-contests. I also find that, this condition is more likely to hold when the designer allocates a higher weight to the expected maximum effort and the ability distribution is convex. However, static all-pay-contests may no longer be the optimal contest format. For instance, if the designer only cares about the expected maximum effort, then she prefers to utilize a winner-pay contest instead.

Contest design with incomplete information and convex effort cost have been first studied by Moldovanu and Sela (2001). They examine the design of static all-pay-contests and allow the designer to choose from a single prize or multiple prizes. They show 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. They do not aim to provide a characterization of the optimal prize structure.

More recently, Olszewski and Siegel revisited this problem in the context of large contests. Olszewski and Siegel (2016) demonstrate 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 (2019), they characterize the optimal prize structure for static all-pay-contests with "sufficiently many" 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 allow for any arbitrary number of contestants and identify a condition under which the winner-takes-all prize structure is optimal, suggesting that it is not without loss to only consider large contests.[2] On the other hand, my results offer a method to determine whether a specific contest should be considered as "sufficiently large", and provide new insights for why awarding many prizes is optimal in large contests.

Another recent paper, by Fang, Noe, and Strack (2019), investigates the optimal prize allocation rule for contests with homogeneous contestants and convex effort cost. They find that it is effort-maximizing to reward identical prizes to all but the worst-performing players, which is almost the least competitive approach to allocate prizes. In addition, they argue that increasing contest competitiveness in general discourages effort. By contrast, I show that, when contestants are ex post heterogeneous and have private information, the designer may choose to utilize both the most competitive contest format and the most competitive prize allocation rule. The sharp contrast between our results highlights the role played by contestant heterogeneity and incomplete information.

Moreover, none of the above mentioned papers establish the optimality of the static,

[2]When this condition fails, the optimal prize allocation rule is not winner-takes-all but in general still depends on the number of contestants.

grand all-pay-contest among all feasible contest formats.

To illustrate the technical challenges, I would like to note that previous studies (Polishchuk and Tonis (2013), Chawla et al. (2019), Liu et al. (2018), etc...) have demonstrated that, with linear effort cost, the optimal contest design problem is mathematically very similar to the optimal auction design problem with risk neutral bidders. With convex effort cost, 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, due to the seminal work 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 have been first studied by Maskin and Riley (1984) and Matthews (1983). Maskin and Riley (1984) consider quite general risk preference and identify several properties of the revenue maximizing mechanism, without attempting to characterize the explicit solution for any specific case. Matthews (1983) provides an explicit solution for the special case with constant absolute risk aversion. More recently, Gershkov et al. (2020) characterize the optimal mechanism for a risk neutral seller who faces Yaari's [1987] dual risk-averse bidders. Most of the results found in Matthews (1983) and Gershkov et al. (2020) do not hold for other risk preferences, including the one studied here. Other papers in this literature typically does not attempt to provide a characterization of 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 and then show the optimal contests must be all-pay. These results allow me to rewrite the designer's objective as a function of the interim allocation rule (also known as the *reduced form allocation rule*). However, unlike in Myerson (1981), the objective function here is not linear in the interim allocation rule. Thus I cannot rewrite the objective function as a function of the ex post allocation rule and then maximize it pointwisely, as usually done for contests with linear cost. 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 (2019)). In this paper, I am able to find a solution by employing a variational analysis approach and utilizing the infinite-dimensional version of the Kuhn-Tucker theorem (see Luenberger (1997)).

The rest of paper is organized as follows. Section 2 presents the model. Section 3 summarizes the main results. I discuss some extensions and applications in Section 4. Section

5 concludes.

2 The Model

Consider a contest with one designer (she) and n agents (he), $n \geq 2$. 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. Each agent $i \in \{1, 2, ..., n\}$ is characterized by his privately-known type $\theta_i \in \Theta = [0, 1]$.[3] A high θ_i means that agent i either has high ability/low cost in performing the contest task or 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 $[0, 1]$. For illustrative convenience, here I assume the prize is divisible and designer chooses how to divide the prize among contestants. All my results also hold for the alternative setting where the prize is non-divisible and the designer chooses who wins the prize. I will discuss the equivalence of the two settings in more details in Section 4.2.

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 any strictly increasing, convex, and twice differentiable function with $c(0) = 0$.[5] If agent i has type θ_i, chooses effort level a_i and obtains prize x_i, then his utility is given by[6]

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

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

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

[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] $c(0) = 0$ is a normalization.

[6] Moldovanu and Sela (2001) assumed 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 auction design problem.

and

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

So 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 income, and has diminishing return to income. 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.[7]

Note that, however, the contest design problem I investigated here differs from a standard risk averse bidder problem in two important ways: First, "bids" (effort) in contests are typically not contractible while bids in auctions are usually considered contractible. Second, contestants can only make positive "bids" while in auctions negative bids are allowed. In what follows, I solve the contest design problem in two steps: I first provide a characterization of the solution to the *dual problem with risk averse bidders* (with contractible bids and allowing for both positive and negative bids). Next, I verify that the solution to the dual problem is also a solution to the original contest problem.

3 Main Results

3.1 Implementable Mechanism

I invoke the revelation principle and restrict attention to direct mechanisms. To allow for both random allocations and random effort, I introduce a random variable $r \in [0,1]$ to capture all randomness in the mechanism.[8] In a direct mechanism $(\boldsymbol{q}, \boldsymbol{a})$, each agent reports his type to the mechanism and a number r 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 demands effort $a_i : \Pi_{i\in N}\Theta \times [0,1] \to [0,\infty)$, as functions of the reported types and the realization of the random variable r. 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 r. In addition, the effort a_i of agent i may be random even conditional on the allocation of the good.

[7] Maskin and Riley (1984) assume the object being sold is non-divisible. But as will be illustrated in Section 4.2, the two settings are equivalent here.

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

For any θ_i and any θ_i' , let

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

denote agent i's expected utility if he is of type θ_i but reports to be type θ_i' (assuming 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.

Given any prior distribution F and any cost function c, the designer's problem is to select an *implementable* direct mechanism $(\boldsymbol{q}, \boldsymbol{a})$ to maximize her expected payoff. As I want to first solve the dual problem with risk averse bidders, in the following part of this section, I assume the effort level is contractible and can be either positive or negative. If effort level must be positive and is not contractible (but observable to all parties), as often assumed in the contest literature, the set of implementable mechanism may be smaller than the one characterized here.[9] However, static all-pay-contests can be implemented regardless of the verifiability of effort.[10] So if such contests are optimal when effort is verifiable, they must also be optimal when effort is not verifiable.

The set of implementable mechanisms is defined by the following conditions. First, each agent must be given the incentive to report his private information 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 either bidder to participate. Thus the individual rationality condition,

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

must be satisfied for any θ_i.

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

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

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

[9]If $(\boldsymbol{q}, \boldsymbol{a})$ can be implemented when effort is not contractible, it can also be implemented when effort is contractible. The reverse may not be true.

[10]The implementability of all-pay-contests with non-verifiable effort will be discussed at the end of Section 3.2.

if there exists a feasible ex post allocation rule $\boldsymbol{q}$ to induce it.

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, regardless of contest rules, the agents' utility in any implementable mechanism are essentially uniquely determined by the reduced form allocation rule, as in the classic auction problem. However, the other classical result, revenue equivalence for the designer, no longer holds here. In other words, the designer's revenue does not only depend on $\boldsymbol{Q}$, but may also depend on the exact details of the contest design. In the next subsection, I demonstrate that, among all feasible contest formats, all-pay-contests with deterministic effort are optimal for the designer.

3.2 All-Pay is Optimal

For later use, I first define the term *all-pay-mechanism with deterministic payments* below.

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

The following lemma shows that all-pay-mechanisms with deterministic payments are optimal among all implementable mechanisms.

Lemma 1. *For any implementable direct mechanism* $(\boldsymbol{q}, \tilde{\boldsymbol{a}})$*, there exists an all-pay-mechanism with deterministic payments 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 contest structure (e.g. multi-round or single round, one grand contest or several sub-contests...) and can make contest rules as complicated as she wishes (for example, the contestants pay only if they receive more than 70% of the prize), she still prefers to use a simple static all-pay-contest. 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' are "risk averse" and must obtain the same payoff in any implementable mechanism, the designer finds it optimal to utilize the mechanism which minimizes the degree of risk exposure. In the environment studied here, this can be achieved by requiring agents to always exert the same effort regardless of the outcome of the contest.

In order to search for the optimal mechanism, I can confine my attention to the class of all-pay-mechanisms with deterministic payments. 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). That is, the designer always prefers to use a static, grand all-pay contests, which is arguably the most competitive contest formats, even if the agents are ex ante asymmetric.[11]

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 any all-pay-mechanism $(\boldsymbol{q}, \boldsymbol{a})$, let $\boldsymbol{Q}$ denote the the corresponding reduced form allocation rule and write $a_i(\theta_i) = a_i(\theta_i, \boldsymbol{\theta_{-i}}, r)$ for short. I obtain by Proposition 1:

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

and further

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

Letting $G = c^{-1}$. As c is increasing and convex, G is increasing and concave. The expected

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

total effort is given by

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

Note that 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 objective function obtained in Myerson (1981).

Before concluding this section, I would like to note that the all-pay mechanism mentioned above can still be implemented in the contest setting. Take any implementable all-pay mechanism $(\boldsymbol{q}, \boldsymbol{a})$ and let $\boldsymbol{Q}$ be the corresponding reduced form allocation rule. Consider an all-pay-contest with 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

$$
a_i(\theta_i) = G\left(\theta_i Q_i(\theta_i) - \int_0^{\theta_i} Q_i(t)dt\right)
$$

for all θ_i, are all positive and constitute a pure strategy Bayesian Nash equilibrium for this game. So if I assume the contestants will take the effort level recommended by the designer when the recommended effort levels are Bayesian incentive compatible for them, $(\boldsymbol{q}, \boldsymbol{a})$ can still be implemented even if the effort is not contractible and must be positive. Thus, I conclude that the solution to the risk averse bidder problem is also the solution to the original contest problem, as desired.

3.3 Optimal Allocation of Prize

As the designer's objective is concave in $(Q_i)_{i=1,2\ldots n}$, I can restrict my attention to symmetric mechanisms and drop below individual subscripts of interim allocation probabilities without causing any loss in generality.[12] In order to employ the variational approach, below I only consider piecewise continuous Q.

The designer's task is to choose some feasible and increasing allocation rule Q to maximize her revenue. For convenience, let $I(\theta) = \int_0^\theta Q(t)dt$ and define

$$H(\theta, Q(\theta), I(\theta)) = G\left(\theta Q(\theta) - I(\theta)\right).$$

So the designer's objective is to choose optimal Q to maximize

$$\int_0^1 H(\theta, Q(\theta), I(\theta))dF(\theta)$$

subject to the constraint 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)

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

(e)

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

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

The above problem can be transformed into a multi-variable calculation of variation problem and then solved by utilizing the infinite-dimensional version of the Kuhn-Tucker theorem (see Luenberger (1997) and Sagan (1969)). To keep the text more focused and concise, I delegate all technical details to the Appendix and only present the results below.

[12] To see this, consider, for example, the case of two bidders. Suppose there exists an optimal pair (Q_1^*, Q_2^*) such that $Q_1^* \neq Q_2^*$. By symmetry, (Q_2^*, Q_1^*) is also optimal. But then the symmetric allocation rule $(\frac{Q_2^*+Q_1^*}{2}, \frac{Q_2^*+Q_1^*}{2})$ is also feasible and it is at least as profitable for the seller (since G is concave). The generalization to more bidders is straightforward.

3.3.1 The Regular Case

I first identify a necessary and sufficient condition under which it is optimal for the designer to utilize a winner-takes-all prize structure. For any Q, let

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

and define

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

to be the *generalized virtual value.* Note that the value of $J(\theta)$ is uniquely determined by the distribution function F and cost function c. I define the term "regularity condition" below.

Definition 2. *The regularity condition holds if and only if $J(\theta)$ is non-decreasing almost everywhere 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 also exerts the highest effort in the contest. Thus a type-θ agent obtains 1 unit of prize with probability $F^{n-1}(\theta)$, i.e. $Q^*(\theta) = F^{n-1}(\theta)$.

Theorem 1. *Suppose the designer uses an all-pay mechanism with deterministic payments. A sufficient and necessary condition for the winner-takes-all prize structure, i.e. $Q^*(\theta) = F^{n-1}(\theta)$ for all θ, being optimal is that the regularity condition holds.*

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

$$\frac{\partial H(\theta, Q(\theta), I(\theta))}{\partial Q(\theta)}$$

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

$$\frac{\int_\theta^1 \frac{\partial H(s,Q(s),I(s))}{\partial I(s)} dF(s)}{f(\theta)},$$

captures the *indirect* marginal cost of increasing $Q(\theta)$. A type-s agent has more incentive to misreport when agents with lower types obtain more prizes, which is captured by $I(s) = \int_0^s Q(t)dt$.[13] That is, the variable $I(s)$ measures the scope of the information rent. Else the

[13] 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).

same, increasing $Q(\theta)$ will lead to an increase in $I(s)$ for any $s > \theta$, which in turn reduces the expected payment from any bidder whose type is above θ. So $J(\theta \mid Q)$ describes the total (direct and indirect) marginal 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 a larger share of prize 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 highest type as much as the feasibility constraint permits - i.e. she will always give the whole prize to the agent with the highest type. 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 (see the relevant proof in the Appendix for details). Lastly, 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), $\theta - \frac{1-F(\theta)}{f(\theta)}$. The case where the regularity condition fails will be discussed in the next subsection.

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

$$H(\theta, Q(\theta), I(\theta)) = \sqrt{\theta Q(\theta) - I(\theta)},$$

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

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 The Non-Regular Case: Multiple Prizes

When the regularity condition fails, the designer may prefer to award multiple smaller prizes rather than one single large prize to the contestants. To find the optimal allocation rule in this case, some "ironing" is needed. For problem with risk neutral bidders, the need for ironing only arises when the monotonicity constraints (i.e. Q is non-decreasing) becomes binding. If this occurs, the designer will choose to award multiple prizes of equal sizes. For risk averse bidder problem, even if the monotonicity constraint never binds, a different type of "ironing" may be required due to the concavity of the objective function. The designer may find it optimal to award multiple prizes of descending sizes. The next theorem describes the optimal allocation rule for this case.

Theorem 2. *Suppose the designer uses an all-pay mechanism with deterministic payments. Suppose Q^* is increasing and optimal. Then for any $\theta \in [0,1]$, one of the following statement must be true:*

(1) 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);

(2) There exists an interval (y,z), such that $\theta \in (y,z)$ and $J(\theta \mid Q^*)$ is constant on (y,z).

Moreover, (2) holds on a set of positive measure if the regularity condition fails.

This theorem states that the designer either assigns the whole prize to the contestant with the highest type, or "irons" the allocation rule to the extent that the rate of return for a marginal increase in Q becomes constant across types. 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 allocating less prizes to the higher type and more prizes to the lower type. The optimum is achieved when marginal return becomes constant across types. The above observation is consistent with the findings of Fang et al. (2019), where they show when contestants are homogeneous, a more unequal prize structure always results in less total effort.

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 strictly 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 $[0,1]$.[14] The solution is uniquely given by $Q(\theta) = \frac{4}{n}\theta^3$, which result in effort level $a(\theta) = \sqrt{\frac{3}{n}}\theta^2$.

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

The optimal mechanism can be implemented, for instance, by conducting an all-pay-contest with the following prize allocation rule: Let a_i be the effort of agent i and x_i be prize allocated to agent i. If at least one contestant exerts positive effort, then

$$x_i = \frac{4}{n} \frac{a_i^{\frac{3}{2}}}{\sum_{j=1,2\ldots n} a_j^{\frac{3}{2}}}.$$

Otherwise, the prize is equally divided among all contestants.

3.4 When to Turn Up the Heat, and How?

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 assigning the prize to the first n bidders according to Q^n and never allocating the 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. So 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 G is strictly concave). So 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 allocation rule 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 allocation rule is given by

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

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

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

which generates the following effort level

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

For any value of n, $Q(\theta)$ is increasing and convex in θ. That is, a more able agent always expects to win more prizes. Moreover, high ability agents are rewarded more aggressively than low ability workers (both the pay-ability ratio $\frac{Q(\theta)}{\theta}$ and the 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 when $n \leq 4$, $Q(\theta)$ becomes steeper as n increases to 4, but then becomes less steeper as n continues to increase. This observation provides a potentially testable implication of the model: the most unequal prize structure (within-job pay structure) 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 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 structure. In this case, as n further increases, the optimal allocation rule becomes less competitive, the contestants' total effort and total welfare remain constant, but the designer's revenue still increases.[15]

The above comparison suggests that, there is a finite level of competition intensity which is optimal for the designer. When the contest size is small, even with the winner-takes-all

[15]Each individual contestant always work less and get worse off when n becomes larger.

prize structure, the overall competition intensity is still lower than the optimal level. So the designer 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 will only discourage effort. However, the designer still prefers to have more contestants, as then she can induce more efficient "risk sharing" among "risk averse" contestants. To balance the needs for "risk sharing" and for "optimal competition intensity", the optimal prize structure has to become less competitive as n increases. This observation provides a potential rational 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 the optimal level.

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. These results suggest the optimal level of competition intensity decreases as the cost function becomes more convex, and increases as the contestants' ability distribution becomes more spread out, or equivalently, as the contest task become more skilled sensitive.

Proposition 2. *Suppose θ is uniformly distributed on $[0,1]$. The regularity condition is less likely to hold when the cost function is more convex.*

As effort cost becomes more convex, the need to induce "risk sharing" 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 $n \leq 8$.[16] The optimal competition intensity is reached at $n = 8$ and $Q(\theta) = \theta^7$. When

[16]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

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

$c(a) = a^2$, the optimal intensity is reached at $n = 4$ and $Q(\theta) = \theta^3$.

Proposition 3. *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.*

The intuition underlying the above result is that competition leads to more efficiency gain when there is a higher level of heterogeneity in the contestant pool.

4 Discussion

4.1 Grand Contest versus Sub-Contests

The revelation principle states that for any incentive compatible indirect mechanism, there always exists a direct mechanism which delivers exactly the same outcome to all parties in every state. So any incentive compatible sub-group contest can also be implemented by a direct mechanism, and thus has been proven to be (weakly) dominated by the grand contest. To be more specific, any sub-group contest corresponds to some group specific allocation rule. For illustration, consider the example where there there are 4 contestants who are divided into two groups, with contestants 1 and 2 in the first group and the others in the second group. Further, the contest rules are such that the prize is equally shared by the winners of each group. Such a contest corresponds to the following allocation rule: for any $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)$, $q_1 = \frac{1}{2}$ if and only if $\theta_1 \geq \theta_2$ and $q_1 = 0$ otherwise ; $q_2 = \frac{1}{2}$ if and only if $\theta_2 > \theta_1$ and $q_2 = 0$ otherwise; $q_3 = \frac{1}{2}$ if and only if $\theta_3 \geq \theta_4$ and $q_3 = 0$ otherwise;$q_4 = \frac{1}{2}$ if and only if $\theta_4 > \theta_3$ and $q_4 = 0$ otherwise.

As argued in Section 3.3, due to the concavity of the designer's objective function, any asymmetric allocation rule is dominated by a symmetric one (strictly so if G is strictly concave). So any sub-group contests which correspond to asymmetric allocation rules are not optimal, and any sub-contests which correspond to symmetric allocation rule can be replicated by a grand contest with the same allocation rule. Thus I can conclude that the designer always (weakly) prefers to run a grand contest where every contestant is invited to participate.

As noted in the Introduction, previous work has found the designer may benefit by conducting sub-group contests. For instance, Moldovanu and Sela (2006) show that if the effort cost function is convex and the designer is restricted to use a single prize, then the designer may benefit by splitting the contestants into several sub-contests. In my setting, if the designer is only allowed to use the winner-takes-all prize structure, then she may find it beneficial to alter the contest architecture to reduce contest competitiveness. This is consistent with the findings of Moldovanu and Sela (2006). However, the designer can achieve such a goal more efficiently by optimally adjusting the prize allocation rule. Fang et al. (2019) find that consolidating two identical contests into a larger contest reduces expected total effort. They also do not allow the designer to optimally adjust the prize allocation rule after the consolidation. Xiao (2017) examines a situation where ex ante asymmetric contestants can be grouped into different sub-contests. He assumes complete information and shows that the designer would prefer to assign agents with similar ability into the same sub-contest. In Section 3.2, I demonstrate that even if contestants are ex ante symmetric, the designer would still prefer to use a grand contest with all contestants. However, the result may change if I impose the additional restriction that the designer can only assign the prize to the highest performer, as assumed in Xiao (2017).

4.2 Non-Divisible Prize

In the earlier parts of this paper, I assumed the prize to be divisible and designer chooses how to divide the prize among contestants. In this subsection, I demonstrate that all my results also apply to the case where the prize is non-divisible and the designer chooses who wins the prize. In this case, agent i's preference is given by

$$u_i(\theta_i, a_i, \mathbb{I}_i) = \theta_i \mathbb{I}_i - c(a_i)$$

where $\mathbb{I}_i = 1$ if agent i obtains the prize and $\mathbb{I}_i = 0$ otherwise. Again, I can invoke the revelation principle and restrict attention to direct mechanisms. In a direct mechanism $(\boldsymbol{p}, \boldsymbol{a})$, each agent reports his type to the mechanism and a number r is randomly drawn from $[0,1]$ according to the uniform distribution; the designer decides who wins the prize according to the allocation rule $p_i : \Pi_{i\in N}\Theta \times 0,1 \to \{0,1\}$ and demands effort $a_i : \Pi_{i\in N}\Theta \times [0,1] \to [0,\infty)$, as functions of the reported types and the realization of the random variable r. Comparing with the direct mechanisms characterized in Section 3, the only difference is that the allocation rule p has to be either 0 or 1. However, for any given type profile, the allocation of the prize can still be random as r is random.

Fix any $(\boldsymbol{p}, \tilde{\boldsymbol{a}}_{\boldsymbol{i}})$, if agent i of type θ_i reports to be type θ_i' (assuming all other agents report truthfully), his utility is given by

$$U_i(\theta_i, \theta_i') = \mathbb{E}\left[\theta_i p_i(\theta_i', \boldsymbol{\theta}_{-\boldsymbol{i}}, r) - c(a_i(\theta_i', \boldsymbol{\theta}_{-\boldsymbol{i}}, r)) \mid \theta_i, \theta_i'\right].$$

By using essentially the same arguments as those used in the proof for Proposition 1 and Lemma 1, I can show the payoff equivalence result still holds and all-pay-mechanisms with deterministic payments are still optimal in this alternative setting. Thus I obtain the same maximization problem for the designer and the only difference is that Q_i is replaced by

$$P_i(\theta_i) = \mathbb{E}[p_i(\theta_i, \boldsymbol{\theta_{-i}}, r) \mid \theta_i],$$

i.e. agent i's expected winning probability. As there is a one-to-one mapping between Q_i and P_i, I can obtain exactly the same solutions as the ones obtained in Section 3, and only need to re-interpret the share of the prize as the allocation probability. To be precise, when the regularity condition holds, the designer always assigns the prize to the contestants with the highest effort. When the regularity condition fails, the designer finds it optimal to utilize lotteries to allocate the prize and the optimal prize allocation rule is the same as described in Theorem 2.

4.3 Expected Maximum Effort

In this section, 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 all-pay-contest and chooses the prize allocation rule Q. Note that the payoff equivalence result (Proposition 1) I obtained earlier holds regardless of the designer's objective. So the agent's effort level still equals

$$a(\theta) = G\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right),$$

which is the same as the effort level obtained in Section 3.2.

As is shown in Section 3.2, the expected total effort is given by

$$\begin{aligned}\pi_A(Q) &= n\int_0^1 a(\theta)dF(\theta)\\ &= n\int_0^1 G\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)dF(\theta).\end{aligned}$$

Recall that in any implementable mechanism, Q must be non-decreasing which implies that a must also be non-decreasing. So 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. So the expected maximum effort is given by

$$\begin{aligned}\pi_M(Q) &= n\int_0^1 a(\theta)F^{n-1}(\theta)dF(\theta)\\ &= n\int_0^1 G\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)F^{n-1}(\theta)dF(\theta).\end{aligned}$$

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. Then the designer's objective is to maximize

$$\begin{aligned}\pi(Q) &= n\alpha\int_0^1 G\left(\theta Q(\theta) - \int_{\theta^*}^\theta Q(t)dt\right)dF(\theta)\\ &\quad +n(1-\alpha)\int_0^1 G\left(\theta Q(\theta) - \int_{\theta^*}^\theta Q(t)dt\right)F^{n-1}(\theta)dF(\theta)\\ &= n\int_0^1 \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right]G\left(\theta Q(\theta) - \int_0^\theta Q(t)dt\right)dF(\theta)\end{aligned}$$

Define

$$H^\alpha(\theta, Q(\theta), I(\theta)) = \left[\alpha + (1-\alpha)F^{n-1}(\theta)\right]G\left(\theta Q(\theta) - I(\theta)\right)$$

Then 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

$$J^\alpha(\theta \mid Q) = \left[\frac{\partial H^\alpha(\theta, Q(\theta), I(\theta))}{\partial Q(\theta)} + \frac{\int_\theta^1 \frac{\partial H^\alpha(s,Q(s),I(s))}{\partial I(s)}dF(s)}{f(\theta)}\right]$$

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.[17]

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

Further, I show that, with convex value distribution, the condition is more likely to hold when α is smaller, i.e. when the designer puts more weight on the expected maximum effort.

Proposition 4. *Suppose F is convex. For any $\alpha_1 > \alpha_2$, if J^{α_1} is non-decreasing, then J^{α_2} is also non-decreasing.*

Before concluding this section, I would like to note that, however, a static all-pay-contest may not be the optimal contest format any more. In fact, if the designer only cares about the expected maximum effort, then a winner-pay contest may work better than any all-pay contests.

Definition 3. *An implementable direct mechanism $(\boldsymbol{q}, \boldsymbol{b})$ is a winner-pay-mechanism with deterministic payments if and only 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 5. *Suppose $\alpha = 0$. For any implementable,symmetric all-pay 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). So the designer may still choose to conduct a static all-pay-contest, especially when she puts a relatively large weight on the expected total effort. The above proposition, however, provides a potential rational for why elimination contests, where contestants pays more when they earns a larger prize, are sometimes used.

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

5 Conclusion

In this paper, I study the optimal design of contests when the contestants have private information and convex effort costs, and the designer wants to maximize the expected total effort. Instead of tackling the contest design problem directly, I first demonstrate that the contest problem can be transformed into an auction design problem with risk averse bidders. I then solve the latter problem by employing a variational approach, in particular, the infinite-dimensional version of the Kuhn-Tucker theorem. Lastly, I verify that the solution to the risk averse bidder problem is indeed a solution to the original contest problem.

The solution is twofold: First, it is always optimal for the designer to employ the most competitive contest format: a static, grand all-pay-contest with as many as possible contestants. 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 structure 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 greater heterogeneity in contestant pool, when the task is more skill-sensitive, when the cost function is less convex, and when the designer puts an additional weight on the expected maximum effort.

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.

[2019] Fang, D., Noe, T., & Strack, P. (2019). Turning up the heat: The discouraging effect of competition in contest, *Journal of Political Economy*, Forthcoming.

[2020] Gershkov, A., Moldovanu, B., Strack, P. and Zhang, M. (2020). Optimal Auctions for Dual Risk Averse Bidders: Myerson meets Yaari. 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.

[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.

[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.

[2019] Olszewski, W., & Siegel, R. (2019). Performance-maximizing large contests, *Theoretical Economics*, Forthcoming.

[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.

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

[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.

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 agents are assumed to be symmetric.

Proof for Proposition 1: (1) (Necessity) The truthtelling constraint requires that for any $\theta_i > \theta_i'$,

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

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 $\delta \in (0, 1 - \theta']$,

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

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 truthtelling 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,

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

So 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 all-pay-mechanism with deterministic payment $(\boldsymbol{q}, \boldsymbol{a})$, where $a_i(\theta_i, \boldsymbol{\theta_{-i}}, r) = a_i(\theta_i)$ is the solution to

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

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

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

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

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

which further implies

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

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). So 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 (M):

$$
\max_Q \int_0^1 H^\alpha(\theta, Q(\theta), I(\theta))dF(\theta)
$$

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 have $I(0) = 0$, but $I(1)$ is not fixed. Lastly, let

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

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

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

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

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

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

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

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

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

(3) $\lambda(\theta) \geq 0$.

(4) The complementary slack condition

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

Now let us return to the previously mentioned Euler-Lagrange conditions. Here I have two such conditions:[18]

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

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

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,

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

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

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

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

Recall that by definition,

$$J^\alpha(\theta \mid Q) = \left[\frac{\partial H^\alpha(\theta, Q(\theta), I(\theta))}{\partial Q(\theta)} + \frac{\int_\theta^1 \frac{\partial H^\alpha(s,Q(s),I(s))}{\partial I(s)} dF(s)}{f(\theta)}\right]$$

and

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

So

$$\frac{dJ^\alpha(\theta \mid Q)}{d\theta} = \frac{d\frac{\partial H^\alpha}{\partial Q}}{d\theta} - \frac{\partial H^\alpha}{\partial I}(\theta) - \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 \frac{\partial H^\alpha}{\partial I}(s)dF(s)$$

which equals the left hand side of Equation (1). Thus the assumption that $J^\alpha(\theta)$ is non-decreasing almost everywhere on $[0,1]$ directly implies that the left hand side of Equation (1), 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. Moreover, the corresponding pair of (x^*, I^*) is on the boundary of the convex feasible set of (x, I) such that

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

and

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

In other words, the inequality constrant

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

is binding everywhere on $[0, 1]$. Thus (x^*, I^*) must be (at least) a local maximizer for $\mathcal{L}$.

To further prove its global optimality, note that the concavity of G implies the functional Y is concave in (x, I) . To see this, I obtain:

$$\frac{\partial^2 h^\alpha}{\partial^2 x'} = \left[\alpha + (1-\alpha) F^{n-1}(\theta)\right] \frac{\theta^2}{f^2(\theta)} G'' \left(-\theta \frac{x'(\theta)}{f(\theta)} - I(\theta)\right) \leq 0,$$

$$\frac{\partial^2 h^\alpha}{\partial^2 I} = \left[\alpha + (1-\alpha) F^{n-1}(\theta)\right] G'' \left(-\theta \frac{x'(\theta)}{f(\theta)} - I(\theta)\right) \leq 0,$$

and

$$\frac{\partial^2 h^\alpha}{\partial x' \partial I} = -\left[\alpha + (1-\alpha) F^{n-1}(\theta)\right] \frac{\theta}{f(\theta)} G'' \left(-\theta \frac{x'(\theta)}{f(\theta)} - I(\theta)\right) \geq 0.$$

Moreover,

$$\frac{\partial^2 h^\alpha}{\partial^2 x'} \frac{\partial^2 h^\alpha}{\partial^2 I} - \frac{\partial^2 h^\alpha}{\partial x' \partial I} \frac{\partial^2 h^\alpha}{\partial x' \partial I} = 0.$$

So the Hessian matrix corresponds to (x', I) is negative semi-definite. It follows that, for any (x_1, I_1), (x_2, I_2), and any $\gamma \in [0, 1]$,

$$\begin{aligned} Y(\gamma(x_1, I_1) + (1-\gamma)(x_2, I_2)) &= \int_0^1 h^\alpha(\theta, \gamma x_1'(\theta) + (1-\gamma) x_2'(\theta), \gamma I_1(\theta) + (1-\gamma) I_2(\theta)) d\theta \\ &\geq \gamma \int_0^1 h^\alpha(\theta, x_1'(\theta), I_1(\theta)) d\theta + (1-\gamma) \int_0^1 h^\alpha(\theta, x_2'(\theta), I_2(\theta)) d\theta \\ &= \gamma Y(x_1, I_1) + (1-\gamma) Y(x_2, I_2) \end{aligned}$$

Thus Y is concave by definition. Then by Proposition 1 in Luenberger (1969), Section 7.8, Q^* is also a global maximizer for the designer's problem, as desired.

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, I either have

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

or $\lambda(\theta) > 0$, which further implies $x(\theta) - \int_\theta^1 F^{n-1}(t)dF(t) = 0$. 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.

Moreover, by the same arguments in the proof for Theorem 3, I can show if such Q^* exists, then it must be the global maximizer for the designer's problem.

Proof for Proposition 2: Take any two concave functions G_1, G_2, and define $J_{G_1}(\theta \mid Q)$ and $J_{G_2}(\theta \mid Q)$ accordingly. Suppose there exists a concave function $K(\cdot)$ such that $G_1 = K(G_2)$. Note that as both G_1 and G_2 are increasing, K must also be increasing. So K' is positive and decreasing.

To prove Proposition 2, it suffices to show $\frac{dJ_{c_1}(\theta|Q)}{d\theta} \geq 0$ implies $\frac{dJ_{c_2}(\theta|Q)}{d\theta} \geq 0$ for any Q. Or equivalently, $\frac{dJ_{c_2}(\theta|Q)}{d\theta} < 0$ implies $\frac{dJ_{G_1}(\theta|Q)}{d\theta} < 0$ for any Q.

Recall that by definition, I have

$$\begin{aligned}
J(\theta \mid Q) =& \frac{\partial H(\theta, Q(\theta), I(\theta))}{\partial Q} + \frac{\int_\theta^1 \frac{\partial H(s,Q(s),I(s))}{\partial I(s)} dF(s)}{f(\theta)} \\
=& \theta G'\left(\theta Q(\theta) - I(\theta)\right)' - \frac{\int_\theta^1 G'\left(sQ(s) - I(s)\right) dF(s)}{f(\theta)} \\
=& \theta G'\left(\theta Q(\theta) - I(\theta)\right)' - \int_\theta^1 G'\left(sQ(s) - I(s)\right) ds
\end{aligned}$$

Then:

$$\begin{aligned}\frac{dJ_{G_1}(\theta \mid Q)}{d\theta} =& \frac{d\left[\theta G_1'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta} \\ &+ G_1'\left(\theta Q(\theta) - I(\theta)\right)) \\ =& \frac{d\left[\theta K'\left(G_2\left(\theta Q(\theta) - I(\theta)\right)\right) G_2'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta} \\ &+ K'\left(G_2\left(\theta Q(\theta) - I(\theta)\right)\right) G_2'\left(\theta Q(\theta) - I(\theta)\right)\end{aligned}$$

In the above formula,

$$\begin{aligned}&\frac{d\left[\theta K'\left(G_2\left(\theta Q(\theta) - I(\theta)\right)\right) G_2'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta} \\ \leq& K'\left(G_2\left(\theta Q(\theta) - I(\theta)\right)\right) \frac{d\left[\theta G_2'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta}\end{aligned}$$

as K is assumed to be concave (an thus $K'' \leq 0$). Thus I obtain, for any Q,

$$\begin{aligned}\frac{dJ_{G_1}(\theta \mid Q)}{d\theta} &\leq K'\left(G_2\left(\theta Q(\theta) - I(\theta)\right)\right)\left[fracd\left[\theta G_2'\left(\theta Q(\theta) - I(\theta)\right)\right]d\theta + G_2'\left(\theta Q(\theta) - I(\theta)\right)\right] \\ &= K'\left(G_2\left(\theta Q(\theta) - I(\theta)\right)\right) \frac{dJ_{G_2}(\theta \mid Q)}{d\theta}.\end{aligned}$$

It is then clear that, for any Q, $\frac{dJ_{G_2}(\theta|Q)}{d\theta} < 0$ would imply $\frac{dJ_{G_1}(\theta|Q)}{d\theta} < 0$, as desired.

Proof for Proposition 3: For this case, computation yields

$$H(\theta, Q(\theta), I(\theta)) = \sqrt{\theta Q(\theta) - I(\theta)},$$

and

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

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

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

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. So 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 4: By definition, I obtain:

$$
\begin{aligned}
J^{\alpha}(\theta\mid Q)=&\left[\frac{\partial H^{\alpha}(\theta,Q(\theta),I(\theta))}{\partial Q(\theta)}+\frac{\int_\theta^1\frac{\partial H^{\alpha}(s,Q(s),I(s))}{\partial I(s)}dF(s)}{f(\theta)}\right]\\
=&\left[\alpha+(1-\alpha)F^{n-1}(\theta)\right]\theta G'\left(\theta Q(\theta)-I(\theta)\right)\\
&-\frac{\int_\theta^1[\alpha+(1-\alpha)F^{n-1}(s)]G'\left(sQ(s)-I(s)\right)dF(s)}{f(\theta)}\\
=&\alpha J^1(\theta\mid Q)+(1-\alpha)J^0(\theta\mid Q)
\end{aligned}
$$

where

$$J^1(\theta\mid Q)=\theta G'\left(\theta Q(\theta)-I(\theta)\right)-\frac{\int_\theta^1 G'\left(sQ(s)-I(s)\right)dF(s)}{f(\theta)}$$

and

$$J^0(\theta\mid Q)=F^{n-1}\theta G'\left(\theta Q(\theta)-I(\theta)\right)-\frac{\int_\theta^1 F^{n-1}(s)G'\left(sQ(s)-I(s)\right)dF(s)}{f(\theta)}.$$

Moreover,

$$\begin{aligned}
\frac{dJ^1(\theta \mid Q)}{d\theta} =& \frac{d\left[\theta G'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta} \\
& + G'\left(\theta Q(\theta) - I(\theta)\right) - \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 G'\left(sQ(s) - I(s)\right) dF(s)
\end{aligned}$$

and

$$\begin{aligned}
\frac{dJ^0(\theta \mid Q)}{d\theta} &= \left(F^{n-1}(\theta)\right)' \theta G'\left(\theta Q(\theta) - I(\theta)\right) + F^{n-1}(\theta) \frac{d\left[\theta G'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta} \\
&\quad + F^{n-1}(\theta) G'\left(\theta Q(\theta) - I(\theta)\right) - \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 F^{n-1}(s) G'\left(sQ(s) - I(s)\right) dF(s) \\
&= \left(F^{n-1}(\theta)\right)' \theta G'\left(\theta Q(\theta) - I(\theta)\right) + F^{n-1}(\theta) \frac{d\left[\theta G'\left(\theta Q(\theta) - I(\theta)\right)\right]}{d\theta} \\
&\quad + F^{n-1}(\theta) G'\left(\theta Q(\theta) - I(\theta)\right) \\
&\quad - F^{n-1}(\theta) \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 \frac{F^{n-1}(s)}{F^{n-1}(\theta)} G'\left(sQ(s) - I(s)\right) dF(s) \\
&\geq F^{n-1}(\theta) \frac{dJ^1(\theta \mid Q)}{d\theta} - F^{n-1}(\theta) \left(\frac{1}{f(\theta)}\right)' \int_\theta^1 \left[\frac{F^{n-1}(s)}{F^{n-1}(\theta)} - 1\right] G'\left(sQ(s) - I(s)\right) dF(s) \\
&\geq F^{n-1}(\theta) \frac{dJ^1(\theta \mid Q)}{d\theta} \qquad (2)
\end{aligned}$$

I obtain the first inequality by using $\left(F^{n-1}(\theta)\right)' \theta G'\left(\theta Q(\theta) - I(\theta)\right) \geq 0$; I obtain the second inequality by using $-\left(\frac{1}{f(\theta)}\right)' \geq 0$ as the distribution is convex, $G'(\cdot) > 0$ and $\frac{F^{n-1}(s)}{F^{n-1}(\theta)} - 1 \geq 0$ for any $s \geq \theta$.

Fix any $\alpha_1 > \alpha_2$. Suppose $\frac{dJ^{\alpha_1}(\theta|Q)}{d\theta} > 0$, then either (1) both $\frac{dJ^0(\theta|Q)}{d\theta}$ and $\frac{dJ^1(\theta|Q)}{d\theta}$ are positive, or (2) $\frac{dJ^0(\theta|Q)}{d\theta} > 0$ and $\frac{dJ^1(\theta|Q)}{d\theta} \leq 0$ (Inequality (2) rules out the other possibility where $\frac{dJ^0(\theta|Q)}{d\theta} \leq 0$ and $\frac{dJ^1(\theta|Q)}{d\theta} > 0$). In the former case, $\frac{dJ^{\alpha_2}(\theta|Q)}{d\theta} > 0$ clearly holds. In the latter case, I have

$$\frac{dJ^{\alpha_2}(\theta)}{d\theta} - \frac{dJ^{\alpha_1}(\theta)}{d\theta} = (\alpha_1 - \alpha_2)\left(\frac{dJ^0(\theta \mid Q)}{d\theta} - \frac{dJ^1(\theta \mid Q)}{d\theta}\right) > 0$$

which implies

$$\frac{dJ^{\alpha_2}(\theta \mid Q)}{d\theta} > \frac{dJ^{\alpha_1}(\theta \mid Q)}{d\theta} > 0.$$

That is, if J^{α_1} is non-decreasing, then J^{α_2} is also non-decreasing, as desired.

Proof for Proposition 5: As shown in Section 4.3, if the designer uses all-pay-mechanisms with deterministic effort, then the expected maximum effort is given by

$$\pi_M(Q) = n \int_0^1 F^{n-1}(\theta) G\left(\theta Q(\theta) - I(\theta)\right) dF(\theta).$$

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

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

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

$$\begin{aligned}
\pi_M^w(Q) &= n \int_0^1 F^{n-1}(\theta) b(\theta) dF(\theta) \\
&= n \int_0^1 F^{n-1}(\theta) G\left(\frac{\theta Q(\theta) - I(\theta)}{F^{n-1}(\theta)}\right) dF(\theta) \\
&\geq n \int_0^1 F^{n-1}(\theta) G\left(Q(\theta)\theta - I(\theta)\right) dF(\theta) = \pi_M(Q)
\end{aligned}$$

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