Effort-Maximizing Contests

Wojciech Olszewski and Ron Siegel*

August 2015

Abstract

We characterize the effort-maximizing prize structure in contests with many players and prizes that accommodate complete information, incomplete information, and ex-ante asymmetric players.

Awarding numerous prizes of different values is optimal when players are risk averse with linear effort cost, or risk neutral with convex effort cost. Awarding a small number of maximal prizes is optimal when players are risk loving with linear effort cost, or risk neutral with concave effort cost.

Our approach facilitates deriving closed-form approximations of the effort-maximizing prize structure for concrete utility functions and distributions of players' types. This facilitates further analysis of large contests.

*Department of Economics, Northwestern University, Evanston, IL 60208 (e-mail: wo@northwestern.edu) and The Pennsylvania State University, University Park, IL 16802 (e-mail: rus41@psu.edu). We thank Bruno Strulovici and audiences at the University of Melbourne, the University of New South Wales, the University of Queensland, the University of Wisconsin (Madison), the Conference in Honor of Sergiu Hart at Hebrew University, the Econometric Society World Congress in Montreal, and the Canadian Economic Theory Conference in London, Ontario for very helpful comments and suggestions. Financial support from the NSF (grant SES-1325968) is gratefully acknowledged.

1 Introduction

This paper investigates the following contest design question. Given a pool of (possibly) ex-ante heterogeneous contestants, who may or may not have private information about their abilities or prize valuations, and given a prize budget, what is the prize structure (the number and values of prizes) that maximizes the aggregate equilibrium effort of the contestants?

Some answers to this question have been provided in the existing literature, but they have often been partial or limited to environments with ex-ante identical players or identical prizes, and relied on restrictive functional forms and informational assumptions.[1] In reality, contestants are often ex-ante asymmetric, their costs of effort and prize valuations may not be linear, and they may have varying degrees of private information. When multiple prizes are awarded, they are often not identical. These features make even contests with a given prize structure difficult or impossible to solve, let alone solving and optimizing over the set of all possible prize structures.

We consider contests with a large number of contestants, who may be ex-ante asymmetric and may or may not have private information about their ability or prize valuations.[2] We model contests as multi-prize all-pay auctions, in which a player's bid represents her effort (or performance). Players' prize valuations and effort costs need not be linear, and the contest may award a combination of heterogeneous and identical prizes. Players' type distributions are independent, but need not be identical. Complete information is a special case. All players choose their effort simultaneously. The player with the highest effort obtains the highest prize, the player with the second-highest effort obtains the second-highest prize, etc.

To solve for the optimal prize structure we use the methods developed in Olszewski and Siegel (2015) (henceforth: OS). They showed that the equilibrium outcomes of contests with a large number of contestants and prizes can be approximated by certain incentive-compatible (IC) and individually rational (IR) mechanisms. The approximation applies uniformly across all equilibria, even when solving for a contest equilibrium may be difficult

[1] See, for example, Glazer and Hassin (1988), Barut and Kovenock (1998), Clark and Riis (1998), and Moldovanu and Sela (2001, 2006). Konrad (2007) provides an overview of the contest literature.

[2] Settings with a large number of contestants who compete for prizes by expending resources include college admissions (in 2012, 4-year colleges in the US received more than 8 mln applications and enrolled approximately 1.5 mln freshmen), grant competitions (in each of the last several years, the National Science Foundation received more than 40,000 grant applications and awarded more than 10,000 grants), sales competitions in large firms (Cisco, which has more than 15,000 partners in the US, holds several sales competitions among its partners), and certain sports competitions (between 2010 and 2012, Tokyo, London, New York, Chicago, and Sydney each hosted a marathon with more than 30,000 participants).

or impossible. This makes it possible to translate the contest design question to a tractable mechanism design problem, namely, to maximize the expected revenue in the mechanism across all prize structures that satisfy a budget constraint. We show that this is in fact a calculus of variations problem, which can be solved by standard methods.

Our main results characterize the optimal prize structure. When effort costs are convex (and prize valuations are linear) or when prize valuations are concave (and effort costs are linear), it is optimal to award numerous prizes of different values. In the former case, if the marginal cost of the first unit of effort is 0, it is optimal to award a prize to nearly all contestants. The latter case corresponds to risk-averse contestants when the prizes are denominated in monetary terms. When effort costs are linear or concave (and prize valuations are linear) or when prize valuations are linear or convex (and effort costs are linear), it is optimal to award only prizes of the highest possible value. As this value increases, the number of optimally awarded prizes decreases, which in the limit corresponds to a single grand prize. These results are in line with the findings of Moldovanu and Sela (2001) (henceforth: MS), who studied the symmetric equilibrium of contests with ex-ante symmetric contestants with incomplete information and linear prize valuations.[3]

In addition to characterizing the optimal prize structure, our approach characterizes the equilibrium efforts of all players in an optimal contest. The approach also uncovers a novel connection between optimal contests and Myerson's (1981) optimal auctions, because it uses a similar mechanism design formulation. The intuition behind this connection helps to explain why in our setting solving for the optimal prize structure is a tractable problem. In discrete contests, such as those of MS, increasing the value of a prize has countervailing effects: it increases competition among contestants who have a good chance of winning this prize, but discourages players with lower ability, who have a lower chance of winning this prize. This makes the problem challenging in a symmetric environment, and intractable in an asymmetric one. But the latter effect is absent in our mechanism design formulation, and this makes the problem tractable. Our characterization can also be used to derive the optimal prize structure in closed form once a functional form for the effort costs or prize valuations is specified.[4] Finally, our approach can be used to investigate contests that are optimal with respect to other goals, such as maximizing contestants' highest effort.[5]

[3]They showed that a grand prize is optimal when effort costs are linear or concave, but may be inferior to a set of two prizes of different values when effort costs are convex. Section 5.2 provides a more detailed comparison of our results for convex costs to those of MS.

[4]This is demonstrated by examples in Sections 4.4 and 5.3.

[5]Bodoh-Creed and Hickman (2015) study affirmative action by investigating a different large contest

Section 2 describes the contest environment. Section 3 describes the mechanism design framework and OS's approximation result. Section 4 analyzes the optimal prize structure when players have linear costs. Section 5 analyzes the optimal prize structure when players have linear prize valuations. Section 6 briefly discusses maximizing contestants' highest effort and concludes. The Appendix contains proofs of results not given in the text.

2 Asymmetric contests

In a contest, n players compete for n prizes of known value. Each player is characterized by a type $x \in X = [0, 1]$, and each prize is characterized by a number $y \in Y = [0, m]$ with $m \geq 1$. Prize 0 is "no prize." The prizes are denoted $y_1^n \leq y_2^n \leq \cdots \leq y_n^n$, some of which may be 0, i.e., no prize. Player i's privately known type x_i^n is distributed according to a *cdf* F_i^n, and these distributions are commonly known and independent across players. In the special case of complete information, each *cdf* corresponds to a Dirac (degenerate) measure.

In the contest, each player chooses his or her effort level t, the player with the highest effort obtains the highest prize, the player with the second-highest effort obtains the second-highest prize, and so on. Ties are resolved by a fair lottery. The utility of a type x player from exerting effort $t \geq 0$ and obtaining prize y is

$$U(x, y, t) = xh(y) - c(t), \tag{1}$$

where $h(0) = c(0) = 0$, and functions h and c are continuously differentiable and strictly increasing.[6] The functional form (1) and special cases thereof have been assumed in numerous existing papers (see, for example, Clark and Riis (1998), henceforth: CR, Bulow and Levin (2006), henceforth: BL, MS, and Xiao (2013)). We will be interested primarily in concave functions h and convex functions c, which capture risk aversion (assuming that y is a monetary prize, while t stands for effort, not for a monetary bid) and typical costs of effort. However, most of our analysis does not require these assumptions, and Sections 4.3 and 5.2 provide results also for convex h and concave c.

Since we study some limits of sequences of contests when n diverges to infinity, we refer to a contest with n players and n prizes as the "n-th contest." Every contest has at least one mixed-strategy Bayesian Nash equilibrium.[7]

model.

[6] Dividing each player's utility by x to obtain $h(y) - c(t)/x$, which accommodates private information about ability, has no effect on the results.

[7] This follows from Corollary 5.2 in Reny (1999).

3 Mechanism-design approach to studying contests

The optimal design of asymmetric contests of the kind described in Section 2 is difficult or impossible, because no method currently exists for characterizing their equilibria for most type and prize distributions. And even in the few cases for which a characterization exists, the equilibria have a complicated form, or can be derived only by means of algorithms (BL, Siegel (2010), and Xiao (2013)). Therefore, we will use the mechanism-design approach to studying the equilibria of large contests, which was developed in OS. We now describe this approach.

3.1 Limit distributions

Let $F^n = \left(\sum_{i=1}^n F_i^n\right)/n$, so $F^n(x)$ is the expected percentile ranking of type x in the n-th contest given the random vector of players' types. We denote by G^n the empirical prize distribution, which assigns a mass of $1/n$ to each prize y_j^n (recall that there is no uncertainty about the prizes). We assume that F^n converges in weak*-topology to a distribution F that has a continuous, strictly positive density f, and G^n converges to some (not necessarily continuous) distribution G.[8]

The convergence of F^n and G^n to limit distributions F and G accommodates as a special, extreme case complete-information contests with asymmetric players in which for some distributions F and G, player i's type in the n-th contest is $x_i^n = F^{-1}(i/n)$ and prize j is $y_j^n = G^{-1}(j/n)$, where

$$G^{-1}(r) = \inf\{z : G(z) \geq r\}.$$

One example is contests with identical prizes and players who differ in their valuations for a prize. For this, consider $h(y) = y$, F uniform, and G that has $G(y) = 1 - p$ for all $y \in [0, 1)$ and $G(1) = 1$, where $p \in (0, 1)$ is the limit ratio of the number of prizes to the number of players. Then $x_i^n = i/n$ and $y_j^n = 0$ if $j/n \leq 1 - p$ and $y_j^n = 1$ if $j/n > 1 - p$. The n-th contest is an all-pay auction with n players and $\ulcorner pn \urcorner$ identical (non-zero) prizes, and the value of a prize to player i is i/n. These contests were studied by CR, who considered competitions for promotions, rent seeking, and rationing by waiting in line.

Another example with complete information is contests with heterogeneous prizes and players who differ in their constant marginal valuation for a prize. For this, consider $h(y) = y$ and F and G uniform. Then $x_i^n = i/n$ and $y_j^n = j/n$. The n-th contest is an all-pay auction with n players and n prizes, and the value of prize j to player i is ij/n^2. These contests

[8] Convergence in weak*-topology can be defined as convergence of *cdf*s at points at which the limit *cdf* is continuous.

were studied by BL, who considered hospitals that have a common ranking for residents and compete for them by offering identity-independent wages.[9]

Many other complete-information contests with asymmetric players can be accommodated, including contests for which no equilibrium characterization exists. One example is contests with a combination of heterogeneous and identical prizes.

Another special, extreme case of the convergence of F^n and G^n is incomplete-information contests with ex-ante symmetric players that have the same *iid* type distributions $F_i^n = F$. This case includes the setting of MS. Beyond these extreme cases, our setting accommodates numerous incomplete-information contests with many ex-ante asymmetric players. No equilibrium characterization exists for such contests.

3.2 Assortative allocation and transfers

The *assortative allocation* assigns to each type x prize $y^A(x) = G^{-1}(F(x))$. It is well known that the unique incentive-compatible mechanism that implements the assortative allocation and gives type $x = 0$ a utility of 0 specifies for every type x effort

$$t^A(x) = c^{-1}\left(xh\left(y^A(x)\right) - \int_0^x h\left(y^A(z)\right) dz\right). \tag{2}$$

For example, in the setting corresponding to CR the assortative allocation assigns a prize to each type higher than $1-p$, and the associated efforts are $t^A(x) = 0$ for $x \leq 1-p$ and $t^A(x) = 1-p$ for $x > 1-p$. In the setting corresponding to BL, the assortative allocation assigns prize x to type x, and the associated efforts are $t^A(x) = x^2/2$.

3.3 The approximation result

Corollary 2 in OS, which we state as Theorem 1 below, shows that the equilibria of contests with many players and prizes are approximated by the unique mechanism that implements the assortative allocation.

Theorem 1 *For any $\varepsilon > 0$ there is an N such that for all $n \geq N$, in any equilibrium of the n-th contest each of a fraction of at least $1-\varepsilon$ of the players i obtains with probability at least $1-\varepsilon$ a prize that differs by at most ε from $y^A(x_i^n)$, and chooses effort that is with probability at least $1-\varepsilon$ within ε of $t^A(x_i^n)$.*

[9] Xiao (2013) presented another model with complete information and heterogenous prizes, in which players have increasing marginal utility for a prize. He considered quadratic and exponential specifications, which are obtained in our model by setting $h(y) = y^2$ and $h(y) = e^y$ and F and G uniform.

Theorem 1 implies that the aggregate effort in large contests can be approximated by

$$\int_0^1 t^A(x) f(x) dx. \tag{3}$$

More precisely, we refer to the aggregate effort divided by n in an equilibrium of the n-th contest as the *average effort.* We then have the following corollary of Theorem **1.**

Corollary 1 *For any $\varepsilon > 0$ there is an N such that for all $n \geq N$, in any equilibrium of the n-th contest the average effort is within ε of (3).*

4 Optimal contests for risk-averse players

We now investigate the prize structures in large contests that maximize the aggregate effort subject to a budget constraint. The constraint says that the average prize, or the prize per contestant, cannot exceed a certain value. We will present the main steps of our analysis, and relegate the technical details to the Appendix. In this section, we assume that $U(x, y, t) = xh(y) - t$, i.e., that $c(t) = t$. In the next section, we obtain corresponding results for $U(x, y, t) = xy - c(t)$. This will show that different risk attitudes and curvatures of cost functions play a similar role in optimal contest design.[10]

Our first result shows that in order to solve the design problem for large contests it is enough to identify the prize distributions that maximize (3) in the limit setting subject to the budget constraint that the expected prize does not exceed a certain value $C > 0$. To formulate this result, consider a sequence of contests whose corresponding sequence of average type distributions F^n converges to a distribution F with a continuous, strictly positive density f. The corresponding empirical prize distributions $G^n_{\max}$ are ones that maximize the aggregate effort. That is, $G^n_{\max}$ describes a set of n prizes that lead to some equilibrium with maximal aggregate effort, subject to the budget constraint that the average prize does not exceed some value C^n that converges to C.[11] Denote the corresponding maximal aggregate

[10] We conjecture that we would obtain a similar result for our general utility function (1), but the analysis would be more complicated and no new interesting insight would be obtained by combining different risk attitudes and curvatures of cost functions.

[11] That a maximizing set of prizes exists can be shown by a straightforward upper hemi-continuity argument of the kind used, for example, to prove Corollary 2 in Siegel (2009). We note, however, that our results do not depend on the existence of such a maximizing set of prizes. For example, none of the analysis changes if $G^n_{\max}$ is instead chosen to correspond to a set of n prizes that lead to some equilibrium with aggregate equilibrium effort that is within $1/n$ of the supremum of the aggregate equilibrium efforts over all sets of n prizes (subject to the budget constraint) and all equilibria for any given set of prizes.

effort by $M^n_{\max}$.

For the limit setting, denote by $\mathcal{M}$ the set of prize distributions that maximize (3) subject to the budget constraint that the expected prize does not exceed C. An upper hemi-continuity argument, given in the Appendix, shows that $\mathcal{M}$ is not empty. Denote by M the corresponding maximal value of (3). Finally, consider any metrization of the weak*-topology on the space of prize distributions.

Proposition 1 *1. For any $\varepsilon > 0$, there is an N such that for every $n \geq N$, $G^n_{\max}$ is within ε (in the metrization) of some distribution in $\mathcal{M}$. In particular, if there is a unique prize distribution $G_{\max}$ that maximizes (3) subject to the budget constraint, then $G^n_{\max}$ converges to $G_{\max}$ in weak*-topology. 2. $M^n_{\max}/n$ converges to M. 3. For any $\varepsilon > 0$ and any G in $\mathcal{M}$, there are an N and a $\delta > 0$ such that for any $n \geq N$ and any empirical prize distribution G^n of n prizes that is within δ of G, the average effort in any equilibrium of the n-th contest with empirical prize distribution G^n is within ε of $M^n_{\max}/n$.*

Part 1 of Proposition 1 shows that the optimal empirical prize distributions in large contests are approximated by the prize distributions that maximize (3) subject to the budget constraint. Part 2 shows that the maximal aggregate equilibrium effort is approximated by the maximal value of (3) subject to the budget constraint. Part 3 shows that any empirical prize distribution that is close to a prize distribution that maximizes (3) subject to the budget constraint generates aggregate equilibrium effort (in any equilibrium) that is close to maximal. For example, given a prize distribution G that maximizes (3) subject to the budget constraint, the set of n prizes defined by $y^n_j = G^{-1}(j/n)$ for $j = 1, ..., n$ generates, for large contests, aggregate equilibrium effort that is close to maximal; moreover, the average prize C^n for the so defined distributions G^n converges to the expected prize C for the distribution G.[12]

4.1 Reduction to a calculus of variations problem

By Proposition 1, we can focus on solving the following problem:

$$\begin{array}{l} \max_G \int_0^1 t^A(x) f(x)\, dx \\ \text{s.t. } \int_0^m y dG(y) \leq C. \end{array}$$

The parameter $C > 0$ should be interpreted as the budget per contestant, denominated in units of the prize $y = 1$. Similarly, prizes are denominated in units of the prize $y = 1$, that is, prize y costs y. Thus, the expected prize cannot exceed C.

[12] It is easy to see that for any G, distributions G^n close to G can always be chosen so that C^n does not exceed C.

To solve this problem, we will show that it is equivalent to a calculus of variations problem in variable G^{-1}. For this, we first transform the objective function. By substituting (2) into (3) and integrating by parts, we obtain the following expression for the aggregate effort in the mechanism that implements the assortative allocation:

$$\int_0^1 t^A(x) f(x)\, dx = \int_0^1 h\left(y^A(x)\right)\left(x - \frac{1-F(x)}{f(x)}\right) f(x)\, dx. \tag{4}$$

To gain some intuition for why (4) approximates the average effort in large contests, observe that (4) coincides with the expected revenue from a bidder in a single-object independent private-value auction if we replace $h\left(y^A(x)\right)$ with the probability that the bidder wins the object when his type is x (Myerson (1981)). In the auction setting, increasing the probability that type x obtains the object along with the price the type is charged allows the auctioneer to capture the entire increase in surplus for this type, but requires a decrease in the price that higher types are charged to maintain incentive compatibility. In a large contest, increasing the prize that type x obtains also allows the auctioneer to capture the entire increase in surplus for this type, because the higher prize increases the competition with slightly lower types until the gain from the higher prize is exhausted. But the prize increase also decreases the effort and competition of higher types for their prizes, since the prize of type x becomes more attractive to them.

For the remainder of the analysis, we make the following assumption, which is standard in the mechanism design literature:[13]

Assumption 1. $x - (1-F(x))/f(x)$ strictly increases in x.

We rewrite (4) using the change of variables $z = F(x)$ to obtain

$$\int_0^1 h\left(G^{-1}(z)\right)\left(F^{-1}(z) - \frac{1-z}{f\left(F^{-1}(z)\right)}\right) dz = \int_0^1 h\left(G^{-1}(z)\right) k(z)\, dz, \tag{5}$$

where $k(z) = F^{-1}(z) - (1-z)/f\left(F^{-1}(z)\right)$.[14] By Assumption 1, $k(z)$ strictly increases in z.

We now transform the budget constraint. Since G is a probability distribution on $[0,m]$, we have $\int_0^m y dG(y) = m - \int_0^m G(y)\, dy$ (by integrating by parts) and $\int_0^m G(y)\, dy + \int_0^1 G^{-1}(z)\, dz = m$ (by looking at the areas below the graphs of G and G^{-1} in the square $[0,m] \times [0,1]$). Thus, the budget constraint can be rewritten as

$$\int_0^1 G^{-1}(z)\, dz \le C. \tag{6}$$

[13] The assumption is implied, for example, by a monotone hazard rate.

[14] Even though G^{-1} may be discontinuous, it is monotonic, so the change of variables applies.

This is the desired form of our maximization problem, because maximizing (5) subject to (6) is a calculus of variations problem in variable G^{-1}.

4.2 Conditions describing the solution

In this section, we derive the conditions that must be satisfied by the optimal G^{-1}. Consider an optimal G^{-1}. Because it is non-decreasing, left-continuous, and takes values in $[0, m]$, there are $z_{\min} \leq z_{\max}$ in $[0,1]$ such that $G^{-1}(z) = 0$ for $z \leq z_{\min}$, $G^{-1}(z) = m$ for $z > z_{\max}$, and $G^{-1}(z) \in (0, m)$ for $z \in (z_{\min}, z_{\max})$.

There are two cases:

Case 1 ($z_{\min} < z_{\max}$): Then, there exists a $\lambda \geq 0$ such that $h'(G^{-1}(z))\, k(z) = \lambda$ for $z \in (z_{\min}, z_{\max}]$; in addition, $h'(0)\, k(z_{\min}) \leq \lambda$, and $h'(m) k(z_{\max}) \geq \lambda$ if $z_{\max} < 1$.

Case 2 ($z_{\min} = z_{\max}$): Then, $h'(0)\, k(z_{\min}) \leq h'(m)\, k(z_{\max})$.

A rigorous proof that G^{-1} satisfies the conditions described in the two cases is provided in the Appendix. To gain some intuition for Case 1, note that $h'(G^{-1}(z))\, k(z)$ is the derivative of the integrand of (5) with respect to $G^{-1}(z)$ for a given z. Thus, if $h'(G^{-1}(\underline{z}))\, k(\underline{z}) < h'(G^{-1}(\overline{z}))\, k(\overline{z})$ for some $\underline{z}, \overline{z} \in (z_{\min}, z_{\max})$, then an infinitesimal increase in $G^{-1}(\overline{z})$ accompanied by a simultaneous decrease in $G^{-1}(\underline{z})$ of the same infinitesimal size would raise the value of the objective function (5) without affecting the budget constraint (6). At $z = z_{\min}$ or $z > z_{\max}$, we have only inequalities, as the value of $G^{-1}$is 0 or m, respectively, and cannot be decreased or increased. Since k is increasing and continuous, the inequality $h'(m)\, k(z) \geq \lambda$ for $z > z_{\max}$ is equivalent to $h'(m) k(z_{\max}) \geq \lambda$.[15] The intuition for Case 2 is analogous.

4.3 Concave and convex functions h

Denote by $x^* \in (0,1)$ the unique type that satisfies $x^* - (1 - F(x^*))/f(x^*) = 0$. Such a type exists because, by Assumption 1, $x - (1 - F(x))/f(x)$ strictly increases in x, and f is continuous and strictly positive on $[0,1]$. For types $x < x^*$, the value of the integrand in (4) is negative, and for $x > x^*$ the value is positive. Let $z^* = F(x^*) \in (0,1)$, so $k(z^*) = 0$. Then, optimizing the integrand in (5) leads to $G^{-1}(z) = 0$ if $z \leq z^*$ and, if the budget allows,

[15] Finally, $h'\left(G^{-1}(z)\right) k(z) = \lambda$ at $z = z_{\max}$ by left-continuity of G^{-1} and continuity of h' and k.

to $G^{-1}(z) = m$ if $z > z^*$.[16] This G^{-1} is left-continuous and monotonic, so the corresponding G is a prize distribution and is therefore optimal. We thus obtain the following result.

Proposition 2 *If $C \geq m(1 - F(x^*))$, then for any function h the optimal prize distribution assigns mass $1 - F(x^*) \in (0, 1)$ to the highest possible prize m and mass $F(x^*)$ to prize 0.*

Proposition 2 shows that an all-pay auction with identical prizes, as studied by CR, is optimal when the budget constraint is not binding. Moreover, it shows that when the budget is large some of the budget is optimally left unused. This is analogous to a monopolist limiting the quantity sold.

Of course, a binding budget constraint seems more relevant in practice. So, for the remainder of our analysis we make the following assumption:

Assumption 2. $C < m(1 - F(x^*))$.

This assumption implies that for the optimal prize distribution the budget constraint (6) holds with equality. We now derive the form of the optimal prize distribution for convex and concave functions h by using the conditions derived in Section 4.2. We first present the simpler, although perhaps less interesting, result for convex functions h.

Proposition 3 *If h is weakly convex, then the optimal prize distribution assigns mass C/m to the highest possible prize and mass $1 - C/m$ to prize 0.*

Proof: In this case, we have $z_{\min} = z_{\max}$. Indeed, since h' is weakly increasing and k is strictly increasing, for any $z' < z''$ in $(z_{\min}, z_{\max})$ we would have $h'(G^{-1}(z'))\,k(z') < h'(G^{-1}(z''))\,k(z'')$, which would violate $h'(G^{-1}(z'))\,k(z') = h'(G^{-1}(z''))\,k(z'') = \lambda$.

Proposition 3 shows that an all-pay auction with identical prizes remains optimal when the budget constraint is binding, provided that agents' marginal prize utility is non-decreasing. If prizes are monetary, this condition says that agents are risk neutral or risk loving. If the highest possible prize, m, is increased, fewer maximal prizes are optimally awarded. The limit as m grows arbitrarily large corresponds to a "single grand prize."

We now consider concave functions h.

Proposition 4 *If h is weakly concave (but not linear on $[0, m]$), then any optimal prize distribution assigns positive mass to intermediate prizes $y \in (0, m)$. An optimal prize distribution may have atoms only at 0 (no prize) and m (the highest possible prize).*

[16] Notice that this corresponds to Case 2 in Section 4.2, with $z_{\min} = z_{\max} = z^*$.

Proof: In this case, we have $z_{\min} < z_{\max}$. Indeed, since $h'(0) > h'(m)$, we cannot have that $z_{\min} = z_{\max}$ and $h'(0)\,k(z_{\min}) \leq h'(m)\,k(z_{\max})$, unless $k(z_{\min}) = k(z_{\max}) \leq 0$. But $k(z_{\max}) \leq 0$ implies that $z_{\max} \leq z^*$, so $G^{-1}(z) = 1$ for $z > z_{\max}$ violates the budget constraint (6). This yields the first part of the result. To obtain the second part, notice that $G^{-1}(z)$ strictly increases in z on interval $(z_{\min}, z_{\max})$. This follows from the fact that $h'(G^{-1}(z))\,k(z) = \lambda$ on $(z_{\min}, z_{\max}]$ and the fact that $k(z)$ strictly increases in z.

Our next result, whose proof is in the Appendix, shows that as the highest possible prize, m, becomes arbitrarily large, the optimal prize distribution is "continuous" on all non-zero prizes.

Proposition 5 *Let $G^m_{\max}$ be a distribution that maximizes (5) subject to the budget constraint when m is the highest possible prize. If h is weakly concave (but not linear $[0, m]$), then when m diverges to infinity $G^m_{\max}$ converges to a distribution that may have an atom only at* 0 *(no prize).*

We now show that when h is strictly concave the constrained maximization problem has an explicit, closed-form solution. As shown in the proof of Proposition 5, $h'(0)\,k(z_{\min}) = \lambda$. Thus,

$$z_{\min} = k^{-1}(\lambda/h'(0)). \tag{7}$$

Since $h'(G^{-1}(z_{\max}))\,k(z_{\max}) = \lambda$ and h' is decreasing, $h'(m)\,k(z_{\max}) \leq \lambda$. If $z_{\max} < 1$, then we also have $h'(m)k(z_{\max}) \geq \lambda$ (because we are in Case 1 of Section 4.2), so we obtain $h'(m)\,k(z_{\max}) = \lambda$. Thus,

$$z_{\max} = 1 \text{ or } k^{-1}(\lambda/h'(m)). \tag{8}$$

In addition,

$$G^{-1}(z) = (h')^{-1}(\lambda/k(z)) \text{ for } z \in (z_{\min}, z_{\max}] \tag{9}$$

and

$$G^{-1}(z) = \begin{cases} 0 & z \leq z_{\min} \\ m & z > z_{\max} \end{cases}.$$

Thus, G^{-1} is pinned down by λ. The value of λ is determined by the binding budget constraint.

4.4 Example

To illustrate our solution, let F be uniform and let $h(y) = \sqrt{y}$.[17] Then $k(z) = 2z - 1$, $x^* = z^* = 1/2$, $h'(0) = \infty$, $h'(s) = 1/(2\sqrt{s})$, and $(h')^{-1}(r) = 1/(4r^2)$. The budget constraint is binding if $C < m(1 - F(x^*)) = m/2$. By (7), $z_{\min} = 1/2$. Suppose first that $z_{\max} = 1$. By (9) and the binding version of (6), we have $\int_{1/2}^{1} (2z-1)^2/(4\lambda^2)dz = C$. Solving for λ, we obtain $\lambda = 1/\sqrt{24C}$. This yields $G^{-1}(z) = 6C(2z-1)^2$; in particular, $z_{\max} = 1$ implies $C \leq m/6$. Thus, we have that

$$G(y) = \begin{cases} \frac{1}{2} + \sqrt{\frac{y}{24C}} & y \in [0, 6C] \\ 1 & y \in [6C, m] \end{cases}.$$

This is a continuous distribution over an interval of positive intermediate prizes (along with a mass 1/2 of "no prize"). The corresponding aggregate effort, given by (5), is $\sqrt{6C}/6$.

Suppose now that $z_{\max} < 1$. By (8), $z_{\max} = \lambda\sqrt{m} + 1/2$. The binding version of (6) implies that $\int_{1/2}^{\lambda\sqrt{m}+1/2}((2z-1)^2/(4\lambda^2))dz + \int_{\lambda\sqrt{m}+1/2}^{1} m dz = C$. Solving for λ, we obtain $\lambda = (3m/4 - 3C/2)/(m\sqrt{m})$. This implies that $G^{-1}(z) = 0$ for $z \in [0, 1/2]$, $G^{-1}(z) = 4(2z-1)^2 m^3/(3m - 6C)^2$ for $z \in (1/2, (5m/4 - 3C/2)/m]$, and $G^{-1}(z) = m$ for $z \in ((5m/4 - 3C/2)/m, 1]$. Since $z_{\max} = \lambda\sqrt{m} + 1/2 < 1$, we have $C > m/6$. Thus,

$$G(y) = \begin{cases} \frac{1}{2} + \sqrt{\frac{y(3m-6C)^2}{16m^3}} & y \in [0, m) \\ 1 & y = m \end{cases}.$$

This is a continuous distribution over an interval of positive intermediate prizes, along with a mass $(6C - m)/4m$ of the highest possible prize (and a mass 1/2 of "no prize"). The corresponding aggregate effort, given by (5), is $(12C(m - C) + m^2)/(16m\sqrt{m})$.

[17] Although $h'(0) = \infty$, it is straightforward to show that a slight modification of our characterization of the solution from the previous subsection still applies.

The following figure depicts these results for $m = 1$.

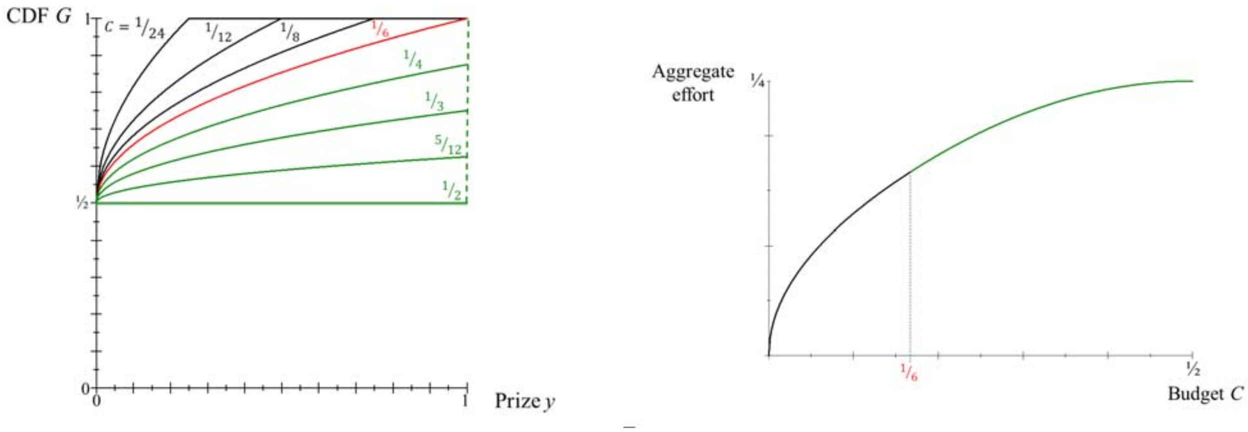

Figure 1: The optimal prize distribution as C increases from 0 to 1/2 (left), and the resulting aggregate effort (right)

To summarize, for $m \in (2C, 6C)$ the optimal prize distribution has atoms at $y = 0$ and $y = m$, and is continuous and strictly increasing on the interval of prizes $y \in (0, m)$. When $m = 6C$, the optimal prize distribution has only an atom at $y = 0$, and is continuous and strictly increasing on the interval of prizes $y \in (0, 6C)$. The same distribution is also the optimal one for any $m > 6C$. The resulting aggregate effort increases in $m \in (2C, 6C)$, and is constant in $m \geq 6C$.[18]

5 Optimal Contests for Agents with Convex Costs

We now suppose that $U(x, y, t) = xy - c(t)$, where $c(0) = 0$ and c is continuously differentiable and strictly increasing. The discrete contests of MS correspond to this utility function. To simplify the analysis, we assume that the limit type distribution F is uniform.[19]

5.1 Conditions describing the solution

The arguments from the beginning of Section 4, including Proposition 1, also apply to the present case. Thus, (2) and (3) imply that the effort-maximizing prize structures in large contests are approximated by the prize distributions that solve the following problem:

[18] For $m \leq 2C$ the optimal prize distribution has atoms at $y = 0$ and $y = m$ of size 1/2 each (by Proposition 2).

[19] Our analysis can be extended to general F and h (instead of $h(y) = y$) without any conceptual difficulty, but such an extension requires more involved notation and calculations.

$$\max_{G^{-1}} \left\{ \int_0^1 c^{-1} \left(zG^{-1}(z) - \int_0^z G^{-1}(r)\, dr \right) dz \right\}$$

$$\text{subject to } \int_0^1 G^{-1}(z)\, dz \leq C.$$

We first transform the objective function to a more convenient form. By looking at the areas below the graphs of G and G^{-1} in the rectangle $[0, G^{-1}(z)] \times [0, z]$, we have that $\int_0^{G^{-1}(z)} G(y)dy + \int_0^z G^{-1}(r)\, dr = zG^{-1}(z)$. Thus, the objective function can be rewritten as

$$\int_0^1 c^{-1} \left(\int_0^{G^{-1}(z)} G(y)dy \right) dz. \tag{10}$$

The difficulty with (10) is that it contains G and G^{-1}, unlike (5) which contained only G^{-1}. Nevertheless, conditions similar to those in Section 4.2 can be provided in this case as well.

Consider an optimal G^{-1}. There exist $z_{\min} \leq z_{\max}$ in $[0,1]$ such that $G^{-1}(z) = 0$ for $z \leq z_{\min}$, $G^{-1}(z) = m$ for $z > z_{\max}$, and $G^{-1}(z) \in (0, m)$ for $z \in (z_{\min}, z_{\max})$.

There are two cases:[20]

Case 1 ($z_{\min} < z_{\max}$): Then, there exists a $\lambda \geq 0$ such that

$$(c^{-1})'(l(z))z - \int_z^1 (c^{-1})'(l(r))dr = \lambda, \tag{11}$$

where $l(z) = \int_0^{G^{-1}(z)} G(y)dy$, for $z \in (z_{\min}, z_{\max}]$; in addition,

$$(c^{-1})'(0)z_{\min} - \int_{z_{\min}}^1 (c^{-1})'(l(r))dr \leq \lambda \text{ and } (c^{-1})'(l(1))(2z_{\max} - 1) \geq \lambda. \tag{12}$$

Case 2 ($z_{\min} = z_{\max}$): Then,

$$(c^{-1})'(0) \leq (c^{-1})'(l(1)). \tag{13}$$

5.2 Convex and concave functions c

As in Section 4.3, we first present the simpler, although perhaps less interesting, result for concave functions c.

[20] The intuition for these cases is similar to the one given in Section 4.2. The details are in the Appendix.

Proposition 6 *If the budget constraint binds and c^{-1} is weakly convex, then the optimal prize distribution assigns mass C/m to the highest possible prize m and mass $1 - C/m$ to prize* 0.

Proof: In this case, we have $z_{\min} = z_{\max}$. Indeed, since $(c^{-1})'$ and l are weakly increasing, $(c^{-1})'(l(z))z$ strictly increases in z; in turn, $\int_z^1 (c^{-1})'(l(r))dr$ weakly decreases in z. Therefore, the left-hand side of (11) strictly increases in z.

Proposition 6 mirrors Propositions 2 and 4 in MS, which show that when the cost function is linear or concave it is optimal to award the entire budget as a single prize. The discrepancy between MS's single prize and the mass of identical highest prizes prescribed by Proposition 6 arises because MS do not impose a bound on the highest possible prize. Increasing the highest possible prize, m, in our setting leads to optimally awarding a smaller mass of this prize. This corresponds, in the limit, to "awarding the entire budget as a single prize."

We now turn to the analysis of convex cost functions c. The results hold under, and are formulated for, conditions somewhat weaker than convexity.

Proposition 7 *If $(c^{-1})'(0) > (c^{-1})'(r)$ for all $r > 0$, then the optimal prize distribution assigns a positive mass to intermediate prizes $y \in (0, 1)$. If $(c^{-1})'(r) > 0$ for all r, then the optimal prize distribution may have atoms only at* 0 *(no prize) and m (the highest possible prize).*

Proof: The first part is true because it follows from (13) that $z_{\min} < z_{\max}$. For the second part, an atom at some intermediate prize would mean that Case 1 must hold and $G^{-1}(\underline{z}) = G^{-1}(\overline{z})$ for some $z_{\min} < \underline{z} < \overline{z} < z_{\max}$. Then, however, $l(\underline{z}) = l(\overline{z})$, so $(c^{-1})'(l(\underline{z}))\underline{z} < (c^{-1})'(l(\overline{z}))\overline{z}$; in turn, $\int_{\underline{z}}^1 (c^{-1})'(l(r))dr \geq \int_{\overline{z}}^1 (c^{-1})'(l(r))dr$. Thus, the left-hand side of (11) with $z = \overline{z}$ would be higher than with $z = \underline{z}$, which contradicts (11).

Proposition 7 is related to Proposition 5 of MS, which shows that with a convex cost function splitting the budget into two prizes is sometimes better than awarding the entire budget as a single prize. Our results go beyond showing that it may not be optimal to award the entire budget as a single prize, and instead characterize the optimal prize structure.

In addition, while Propositions 6 and 7 are related to the results in MS, the set of contests and equilibria to which they apply are different from those studied by MS. While MS studied contests with any finite number of players, the players were restricted to being ex-ante symmetric and having private information about their cost, and the analysis focused on the symmetric equilibrium. Our results apply to **all equilibria of** contests with a large, but finite, number of players. The players may be ex-ante symmetric or asymmetric, and may or may not have private information.

The next result shows that Proposition 5 extends to the setting with convex cost function c.

Proposition 8 *Let $G^m_{\max}$ be the distribution that maximizes (10) subject to the budget constraint when m is the highest possible prize. If c is weakly convex (but not linear on any interval with lower bound 0), then as m diverges to infinity, $G^m_{\max}$ converges to a distribution that may have an atom only at 0 (no prize).*

5.3 Example

To illustrate our solution, and also show that for specific utility functions we can derive the corresponding optimal prize distributions G in closed form, let F be uniform and let $c(t) = t^2$. Proposition 7 shows that $z_{\min} < z_{\max}$ and the optimal prize distribution G may have atoms only at 0 and m. For simplicity, let $m = 1$.

Define an auxiliary function $m(z) = (c^{-1})'(l(z))$. Plug $m(z)$ into (11), and differentiate with respect to z to obtain the differential equation $m\prime(z)z + 2m(z) = 0$ for $m(z)$.[21] Solving this equation, and substituting back into (11), we obtain $(c^{-1})'(l(z)) = \lambda/z^2$. By the definition of $l(z)$, and using the equality $\int_0^{G^{-1}(z)} G(y)dy = zG^{-1}(z) - \int_0^z G^{-1}(r)\,dr$, we obtain $((c^{-1})')^{-1}(\lambda/z^2) = zG^{-1}(z) - \int_0^z G^{-1}(r)\,dr$. Assuming differentiability of G^{-1}, we obtain $(G^{-1})'(z) = (-2\lambda/z^4)(((c^{-1})')^{-1})'(\lambda/z^2)$.[22]

Since $c^{-1}(z) = \sqrt{z}$, $(c^{-1})'(z) = 1/(2\sqrt{z})$,[23] $\left((c^{-1})'\right)^{-1}(z) = 1/(4z^2)$, and $\left(\left((c^{-1})'\right)^{-1}\right)'(z) = -1/(2z^3)$. Thus, $G^{-1}(z) = z^3/(3\lambda^2) + y_{\min}$, where $y_{\min}$ is the "lowest prize" awarded. Since $(c^{-1})'(0) = \infty$, the first inequality in (12) implies that we have $z_{\min} = 0$. We must therefore have $y_{\min} = 0$, as otherwise $G^{-1}(z)$ can be "shifted down" to $G^{-1}(z) - y_{\min}$. This would reduce by $y_{\min}$ the cost of providing the prizes without affecting players' incentives, leading to the same aggregate equilibrium effort and relaxing the budget constraint.

---

[21] The solution can be verified to be differentiable.

[22] We will show that an optimal prize distribution G with differentiable inverse G^{-1} exists. No other prize distribution will lead to higher aggregate effort, since the aggregate effort corresponding to any prize distribution can be approximated arbitrarily closely by the aggregate effort corresponding to a prize distribution with a differentiable inverse.

[23] Although $\left(c^{-1}\right)'(0) = \infty$, it is straightforward to show that a slight modification of our characterization of the solution from Section 5.1 still applies.

Suppose first that $z_{\max} = 1$. Substituting the expression for $G^{-1}(z)$ into the binding budget constraint, we obtain $\lambda = 1/\sqrt{12C}$, which gives $G^{-1}(z) = 4z^3C$, so $C \leq 1/4$. Substituting G^{-1} into the target function, the aggregate effort is $\sqrt{3C}/3$, which increase in the budget C.

Now suppose that $z_{\max} < 1$. The binding budget constraint gives $\lambda = z_{\max}^2/(12(C - 1 + z_{\max}))^{1/2}$, so $1 = G^{-1}(z_{\max}) = 12(C - 1 + z_{\max})/(3z_{\max})$, which implies that $z_{\max} = 4(1 - C)/3$. Since $z_{\max} < 1$, we must have that $C > 1/4$. Substituting the expression for $z_{\max}$ into the expression for λ, and substituting the resulting expression into the expression for G^{-1}, gives $G^{-1}(z) = 27z^3/\left(64(1 - C)^3\right)$. Substituting into the target function, the aggregate effort is $\sqrt{(1 - C)}\,(1 - 8(1 - C)/9)$. This expression increases for C in $(1/4, 5/8]$, and decreases for C in $[5/8, 1]$. Notice that the value of this expression at $C = 1/4$ coincides with the one for $z_{\max} = 1$.

We therefore conclude that the budget constraint binds for $C \leq 5/8$. For $C \leq 1/4$ the maximal aggregate effort is $\sqrt{3C}/3$, and the optimal prize distribution is $G(y) = \sqrt[3]{y/4C}$ for $y \leq 4C$ and $G(y) = 1$ for $y > 4C$. For C in $(1/4, 5/8]$ the maximal aggregate effort is $\sqrt{(1 - C)}\,(1 - 8(1 - C)/9)$, and the optimal prize distribution is $G(y) = \sqrt[3]{y}4(1 - C)/3$ for $y < 1$ and $G(y) = 1$ for $y = 1$.

Notice that regardless of C, the prizes awarded begin with the lowest prize $y = 0$, but in contrast to the case of concave h studied in Section 4.4, there is no atom at the lowest prizes, i.e., every player gets a positive prize. This is because the marginal cost of effort at 0 is 0 (so $(c^{-1})'(0) = \infty$). In addition, even when the budget constraint does not bind, there is a positive mass of intermediate prizes.

The following figure depicts these results.

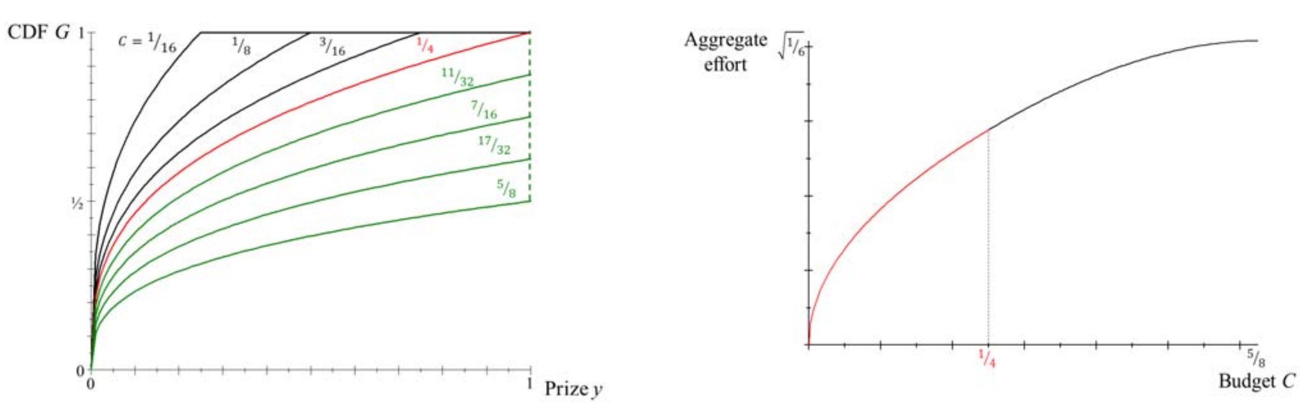

Figure 2: The optimal prize distribution as C increases from 0 to 5/8 (left), and the resulting aggregate effort (right)

The two solved examples, $h(x) = \sqrt{x}$ and $c(t) = t^2$, demonstrate that there are differences as well as similarities between concave valuations and convex costs. One difference is that with t^2 every player obtains a positive prize, while with $\sqrt{x}$ there is a mass 1/2 of players who get a prize of 0. Another difference is that as the prize budget increases, with $\sqrt{x}$ the prize distribution approaches a mass 1/2 of the highest possible prize, whereas with t^2 the prize distribution with an unrestricted budget is still a range of prizes, plus an atom at the highest possible prize.

6 Conclusion

This paper investigated the effort-maximizing prize structure in all-pay contests with asymmetric players and heterogeneous prizes. Such contests are often difficult or impossible to solve when the number of players is small, which makes contest design intractable. We employ the methods developed in OS to solve the design question in a general contest setting with a large number of players. Our key insight is that contestants' risk aversion and convex effort cost call for numerous prizes of different values. In contrast, risk neutrality or love call for a small number of prizes of the highest possible value, or a single grand prize. The same is true for concave effort costs, which is in line with the results of MS.

Our methods also allow for deriving closed-form approximations of the effort-maximizing prize structure for concrete utility functions and distributions of players. This facilitates further analysis of large contests.

Finally, our framework can also be used to investigate other contest design questions. One example is identifying the prize structure that maximizes the highest efforts, rather than the aggregate, or average, effort. This is relevant, for example, in innovation contests whose goal is to generate the best inventions, products, or technologies. In the Appendix we show that the optimal prize structure in such settings with many contestants is a small number of prizes of the highest possible value. This is true regardless of contestants' risk attitude and cost curvature.

7 Appendix

Proof of Corollary 1. Theorem 1 shows that for large n, in any equilibrium of the n-th contest the average effort is within $\varepsilon/2$ of

$$\frac{\sum_{i=1}^{n} \int_0^1 t^A(x)\, dF_i^n(x)}{n} = \int_0^1 t^A(x)\, dF^n(x),$$

where the equality follows from the definition of F^n. In addition,

$$\int_0^1 t^A(x)\, dF^n(x) \to_n \int_0^1 t^A(x)\, dF(x),$$

which follows from the fact that t^A is monotonic and the assumption that F is continuous, because $\int g dF^n \to_n \int g dF$ for any bounded and measurable function g for which distribution F assigns measure 0 to the set of points at which function g is discontinuous. (This fact is established as the first claim of the proof of Theorem 25.8 in Billingsley (1995).) Thus, for large n, $\int_0^1 t^A(x)\, dF^n(x)$ is within $\varepsilon/2$ of $\int_0^1 t^A(x)\, dF(x)$.

Proof that $\mathcal{M} \neq \varnothing$. Let $(G^n)_{n=1}^\infty$ be a sequence on which (3) converges to its supremum, and which satisfies the budget constraint. By passing to a convergent subsequence (in the weak*-topology) if necessary, assume that G^n converges to some G. We will show below that $(G^n)^{-1}$ converges almost surely to G^{-1}. This will imply that $(y^n)^A(x) = (G^n)^{-1}(F(x))$ converges almost surely to $y^A(x) = G^{-1}(F(x))$, and since functions h and c^{-1} are continuous, also that $(t^n)^A(x)$ given by (2) with G replaced with G^n converges almost surely to $t^A(x)$ given by (2). This will in turn imply that the value of (3) with $(G^n)^{-1}$ instead of G^{-1} converges to the value of (3). Finally, as G^n satisfies the budget constraint with C^n, and C^n converges to C, we have that G satisfies the budget constraint with C. Indeed, the budget constraints are integrals of a continuous function (mapping y to y) with respect to distributions G and $G^n_{\max}$, respectively, and weak*-topology may be alternatively defined by convergence of integrals over continuous functions.

Thus, it suffices to show that $(G^n)^{-1}$ converges to G^{-1}, except perhaps on the (at most) countable set $R = \{r \in [0,1] : \text{there exist } y' < y'' \text{ such that } G(y) = r \text{ for } y \in (y', y'')\}$.

Suppose first that for some $r \in [0,1]$ and $\delta > 0$ we have that $(G^n)^{-1}(r) \leq G^{-1}(r) - \delta$ for arbitrarily large n. Passing to a subsequence if necessary, assume that the inequality holds for all n, and that $(G^n)^{-1}(r)$ converges to some $y \leq G^{-1}(r) - \delta$. Then, there exists a prize z such that $y < z < G^{-1}(r)$ and G is continuous at z. We cannot have that $G(z) = r$, since this would imply that $G^{-1}(r) \leq z$. Thus, $G(z) < r$. Since $G^n(z)$ converges to $G(z)$, as G is continuous at z, we have that $G^n(z) < r$ for large enough n. This yields $z \leq (G^n)^{-1}(r)$, contradicting the assumption that $(G^n)^{-1}(r)$ converges to $y < z$.

Suppose now that for some $r \in [0,1] - R$ and $\delta > 0$ we have that $(G^n)^{-1}(r) \geq G^{-1}(r) + \delta$ for arbitrarily large n. Passing to a subsequence if necessary, assume that the inequality holds for all n, and that $(G^n)^{-1}(r)$ converges to some $y \geq G^{-1}(r) + \delta$. Then, there exists a prize z such that $G^{-1}(r) < z < y$ and G is continuous at z. We have that $r < G(z)$, as $r \notin R$. Since $G^n(z)$ converges to $G(z)$, as G is continuous at z, we have that $r \leq G^n(z)$

for large enough n. This yields $(G^n)^{-1}(r) \leq z$, contradicting the assumption that $(G^n)^{-1}(r)$ converges to $y > z$.

Proof of Proposition 1. Since every sequence of distributions has a converging subsequence in weak*-topology, suppose without loss of generality that $G^n_{\max}$ converges to some distribution G. Denote the value of (3) under distribution G by V. If Part 1 is false, then $G \notin \mathcal{M}$, so $V < M$. The distribution G satisfies the budget constraint, since distributions $G^n_{\max}$ satisfy the budget constraint.

Consider a distribution $G_{\max} \in \mathcal{M}$, and for every n consider an empirical distribution G^n of a set of n prizes, such that G^n converges to $G_{\max}$ in weak*-topology. For example, such a set of n prizes is defined by $y^n_j = G^{-1}_{\max}(j/n)$ for $j = 1, ..., n$.

Corollary **1** shows that for large n the average effort in any equilibrium of the n-th contest with empirical prize distribution G^n exceeds $2(V + M)/3$. On the other hand, Corollary **1** also shows that for large n the average effort in any equilibrium of the n-th contest with empirical prize distribution $G^n_{\max}$ falls below $(V + M)/3$. This contradicts the definition of $G^n_{\max}$ for large n.

For Part 2, Corollary **1** applied to the sequence G^n defined above implies that $\liminf M^n_{\max}/n \geq M$. If $\limsup M^n_{\max}/n > M$, then there is a corresponding subsequence of $G^n_{\max}$. A converging subsequence of this subsequence has a limit G. For this G, the value of (3) is by Corollary **1** strictly larger than M, a contradiction.

Part 3 follows from part 2 and the fact that Corollary **1** shows that the average effort in any equilibrium of the n-th contest with empirical prize distribution G^n converges to M.

Proof for the conditions in Cases 1 and 2 from Section 4.2. To simplify notation, we assume that $m = 1$.

We will show that in Case 1 the condition $h'(G^{-1}(z))\,k(z) = h'(G^{-1}(z'))\,k(z')$ holds for all $z, z' \in (z_{\min}, z_{\max})$. For this, we first approximate G^{-1} by a sequence of inverse distribution functions $((G^n)^{-1})_{n=1}^{\infty}$ that satisfy the budget constraint and whose value of (5) converges to that for G^{-1}. We then show that if the condition fails there exists a sequence of inverse distribution functions $((H^n)^{-1})_{n=1}^{\infty}$ that satisfy the budget constraint such that for large n the value of (5) for $(H^n)^{-1}$ exceeds that for $(G^n)^{-1}$ by a positive constant independent of n, and therefore improves upon G^{-1}. The second condition in Case 1 and the condition in Case 2 are obtained by analogous arguments.

To define $(G^n)^{-1}$, partition interval $[0, 1]$ into intervals of size $1/2^n$, and set the value of $(G^n)^{-1}$ on interval $(j/2^n, (j+1)/2^n]$ to be constant and equal to the highest number in the set $\{0, 1/2^n, 2/2^n, ..., (2^n - 1)/2^n, 1\}$ that is no higher than $G^{-1}(j/2^n)$. By left-continuity of G^{-1}, $(G^n)^{-1}$ converges pointwise to G^{-1}, so the value of (5) for $(G^n)^{-1}$ converges to that for G^{-1}.

Suppose that $h'(G^{-1}(\underline{z}))\, k(\underline{z}) < h'(G^{-1}(\overline{z}))\, k(\overline{z})$ for some $\underline{z}, \overline{z} \in (z_{\min}, z_{\max})$. By left-continuity of G^{-1}, and continuity of h' and k, the previous inequality also holds for points slightly smaller than $\underline{z}$ and $\overline{z}$. Thus, there are $\delta > 0$, N, and intervals $(j/2^N, (j+1)/2^N]$ and $(l/2^N, (l+1)/2^N]$, such that for every $n \geq N$ we have $h'((G^n)^{-1}(z))\, k(z) - h'((G^n)^{-1}(z'))k(z') > \delta$ for any $z \in (j/2^N, (l+1)/2^N]$ and $z' \in (l/2^N, (j+1)/2^N]$.

Denote the infimum of the values $h'((G^n)^{-1}(z))k(z)$ for $n \geq N$ and z in the former interval by I, and the supremum of the values $h'((G^n)^{-1}(z))k(z)$ for $n \geq N$ and z in the latter interval by S. Now, define functions $(\widetilde{H}^n)^{-1}$ by increasing the value of $(G^n)^{-1}$ on $(j/2^N, (j+1)/2^N]$ by ε, and decreasing the value of $(G^n)^{-1}$ on $(l/2^N, (l+1)/2^N]$ by ε, so the budget constraint is maintained. For sufficiently small $\varepsilon > 0$, the former change increases (5) at least by $(\varepsilon/2^N)(I - \delta/3)$, and the latter change decreases (5) at most by $(\varepsilon/2^N)(S + \delta/3)$. This increases the value of (5) by at least $\delta\varepsilon/2^N$ (for all $n \geq N$).

If functions $(\widetilde{H}^n)^{-1}$ are monotonic, they are inverse distribution functions, so it suffices to set $(H^n)^{-1} = (\widetilde{H}^n)^{-1}$. Otherwise, define $(H^n)^{-1}$ by setting its value on interval $(0, 1/2^n]$ to the lowest value of $(\widetilde{H}^n)^{-1}$ over intervals $(0, 1/2^n], (1/2^n, 2/2^n], ..., ((2^n - 1)/2^n, 1]$, setting its value on interval $(1/2^n, 2/2^n]$ to the second lowest value of $(\widetilde{H}^n)^{-1}$ on these intervals, etc. The value of (5) is higher for $(H^n)^{-1}$ than for $(\widetilde{H}^n)^{-1}$ because k is an increasing function.

Proof of Proposition 5. Let $z^m_{\min}$, $z^m_{\max}$, and λ^m denote $z_{\min}$, $z_{\max}$, and λ for a given m. The proof of Proposition 4 shows that $z^m_{\min} < z^m_{\max}$ for all m. We claim that λ^m weakly increases with m. Suppose to the contrary that $\lambda^{m'} > \lambda^{m''}$ for some $m' < m''$.

Since $h'\left((G^m_{\max})^{-1}(z)\right) k(z) = \lambda^m$ for all $z \in (z^m_{\min}, z^m_{\max}]$ and h' is decreasing, $h'(0)\, k(z) \geq \lambda^m$ for all $z \in (z^m_{\min}, z^m_{\max}]$, and since k is continuous, we have $h'(0)\, k(z^m_{\min}) \geq \lambda^m$. Since we also have $h'(0)\, k(z^m_{\min}) \leq \lambda^m$ (because we are in Case 1 of Section 4.2), we obtain $h'(0)\, k(z^m_{\min}) = \lambda^m$. Since k is increasing, this implies that $z^{m'}_{\min} > z^{m''}_{\min}$. In particular, we have (a): $(G^{m'}_{\max})^{-1}(z) = 0 \leq (G^{m''}_{\max})^{-1}(z)$ for all $z \leq z^{m'}_{\min}$, and the inequality is strict for $z \in (z^{m''}_{\min}, z^{m'}_{\min})$. Since $h'\left((G^m_{\max})^{-1}(z)\right) k(z) = \lambda^m$ for all $z \in (z^m_{\min}, z^m_{\max}]$ and h' is decreasing, we have (b): $(G^{m'}_{\max})^{-1}(z) \leq (G^{m''}_{\max})^{-1}(z)$ for all $z \in (z^{m'}_{\min}, \min\{z^{m'}_{\max}, z^{m''}_{\max}\}]$. If $z^{m'}_{\max} \geq z^{m''}_{\max}$, then we have (c): $(G^{m'}_{\max})^{-1}(z) \leq m' < (G^{m''}_{\max})^{-1}(z) = m''$ for $z > \min\{z^{m'}_{\max}, z^{m''}_{\max}\}$. If $z^{m'}_{\max} < z^{m''}_{\max} \leq 1$, then $h'(m')k\left(z^{m'}_{\max}\right) \geq \lambda^{m'}$ (because we are in Case 1 of Section 4.2). But $h'\left(\left(G^{m''}_{\max}\right)^{-1}\left(z^{m'}_{\max}\right)\right) k\left(z^{m'}_{\max}\right) = \lambda^{m''}$, so $\lambda^{m'} > \lambda^{m''}$ implies that $\left(G^{m''}_{\max}\right)^{-1}\left(z^{m'}_{\max}\right) \geq m'$. Thus, as the inverse of any CDF is increasing, we again obtain (c), except $(G^{m''}_{\max})^{-1}(z) \leq m''$.

Now, (a), (b), and (c) imply that the budget constraint cannot be satisfied with equality by both $G^{m'}_{\max}$ and $G^{m''}_{\max}$, which completes the argument.

By $h'(0)\,k\left(z^m_{\min}\right) = \lambda^m$, we obtain that $z^m_{\min}$ also weakly increases with m.

Notice now that either (a) $z^m_{\max} = 1$ for sufficiently large m, or (b) λ^m converges to some λ as m diverges to infinity. Otherwise, the condition that $h'(m)k\left(z^m_{\max}\right) \geq \lambda^m$ if $z_{\max} < 1$ would be violated. In case (a), the monotonicity of $z^m_{\min}$ and λ^m implies that λ^m stabilizes at some λ for sufficiently large m,[24] and case (b) implies that $z^m_{\max}$ converges to 1 as m diverges to infinity, otherwise the budget constraint would be violated.[25] Therefore, in both these cases $G^m_{\max}$ converges to a distribution that may have an atom only at 0.

Proof for the conditions in Cases 1 and 2 from Section 5.1. To simplify notation, we again assume that $m = 1$.

The proof that G^{-1} satisfies the conditions described in the two cases is analogous to that for the conditions in Section 4.2. The argument is, however, more involved, because the objective function (10) depends on G as well as on G^{-1}. For the argument, it is convenient to extend the functional $l(z)$ to functions G^{-1} that are not monotonic. We define $l(z)$ by adding with the plus sign the area above the graph of G^{-1} between 0 and z and below the line $y = G^{-1}(z)$, and with the minus sign the area below the graph of G^{-1} between 0 and z and above the line $y = G^{-1}(z)$. (This is illustrated in Figure 3, where $l(z)$ is equal to the sum of the shaded areas taken with the signs marked on them.)

[24] If λ^m increases, then $z^m_{\min}$ increases, and $h'\left(\left(G^m_{\max}\right)^{-1}(z)\right)k(z) = \lambda^m$ implies that $\left(G^m_{\max}\right)^{-1}(z)$ decreases for $z \in (z_{\min}, z_{\max})$. Thus, once $z^m_{\max}$ stabilizes at 1, λ^m can no longer increase, since (6) holds with equality.

[25] If $z^m_{\max}$ were bounded away from 1 for sufficiently large m, $G^{-1}(z)$ would be equal to m on an interval of lenth bounded away from 0, so $\int_0^1 G^{-1}(z)\,dz$ would diverge to infinity.

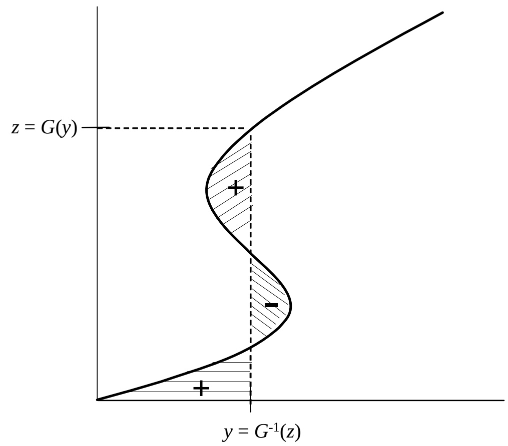

Figure 3: The definition of $l\,(z)$

To derive the first condition in Case 1, consider some inverse distribution function G^{-1} that takes values only in the set $\{0, 1/2^n, 2/2^n, ..., (2^n - 1)/2^n, 1\}$, and is constant on each interval $(0, 1/2^n], (1/2^n, 2/2^n], ..., ((2^n-1)/2^n, 1]$. Suppose that we increase the value of $(G)^{-1}$ on an interval $(l/2^n, (l+1)/2^n]$ by $\varepsilon > 0$. (That is, we move the graph of $(G)^{-1}$ in Figure 4 to the right, by the shaded square.) This change does not affect the integrand in (10) on intervals $(k/2^n, (k+1)/2^n]$ for $k < l$. It increases $\int_0^{G^{-1}(z)} G(y)dy$ for $z \in (l/2^n, (l+1)/2^n]$ by $\varepsilon(l/2^n)$ (the darkened rectangle in Figure 4), so to a first-order approximation it increases the integrand in (10) on $(l/2^n, (l+1)/2^n]$ by $(c^{-1})'(l(z))\varepsilon(l/2^n)$. For any $k > l$, it decreases $\int_0^{G^{-1}(z)} G(y)dy$ by $\varepsilon(1/2^n)$ (the shaded square in Figure 4) on $(k/2^n, (k+1)/2^n]$, so to a first-order approximation it decreases the integrand in (10) on $(k/2^n, (k+1)/2^n]$ (for all $k > l$) by $(c^{-1})'(l(z))\varepsilon(1/2^n)$. Letting $z = l/2^n$, we have that, in total, (10) increases approximately by

$$\varepsilon(1/2^n)\left[(c^{-1})'(l(z))z - \int_z^1 (c^{-1})'(l(r))dr\right].$$

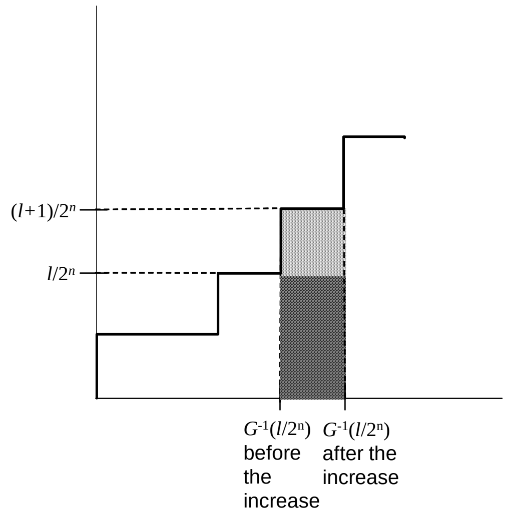

Figure 4: Increasing G^{-1}

Thus, if the first condition in Case 1 is violated for an optimal G^{-1}, we could construct functions $(G^n)^{-1}$ that converge to G^{-1} and functions $(\widetilde{H}^n)^{-1}$, as in the proof for the conditions from Section 4.2. If functions $(\widetilde{H}^n)^{-1}$ are monotonic, we would obtain a contradiction to the optimality of G^{-1}.

If a $(\widetilde{H}^n)^{-1}$ is not monotonic, then there is a monotonic $(H^n)^{-1}$ whose value of (10) is higher than that for $(\widetilde{H}^n)^{-1}$. Indeed, consider two adjacent intervals $(k/2^n, (k+1)/2^n]$ and $(l/2^n, (l+1)/2^n]$ (that is, $k+1 = l$) such that $(\widetilde{H}^n)^{-1}(z) = U$ on $(k/2^n, (k+1)/2^n]$ and $(\widetilde{H}^n)^{-1}(z) = D$ on $(l/2^n, (l+1)/2^n]$, where $D < U$. By changing the value of $(\widetilde{H}^n)^{-1}$ on $(k/2^n, (k+1)/2^n]$ to D, and changing the value of $(\widetilde{H}^n)^{-1}$ on $(l/2^n, (l+1)/2^n]$ to U, we raise the value of (10). This is easy to see in Figure 5, in which the graph of $(H^n)^{-1}$ is obtained from the graph of $(\widetilde{H}^n)^{-1}$ by moving it to the left by the shaded square, and moving it to the right by the darkened square. This makes the value of $l(z)$ on $(k/2^n, (k+1)/2^n]$ higher than its previous value on $(l/2^n, (l+1)/2^n]$ by the shaded area. Similarly, the value of $l(z)$ on $(l/2^n, (l+1)/2^n]$ becomes higher than its previous value on $(k/2^n, (k+1)/2^n]$ by the shaded area. This increases the integrand of (10) on $(k/2^n, (l+1)/2^n]$. Finally, the value of $l(z)$ and the integrand of (10) on other intervals of the partition stay the same.

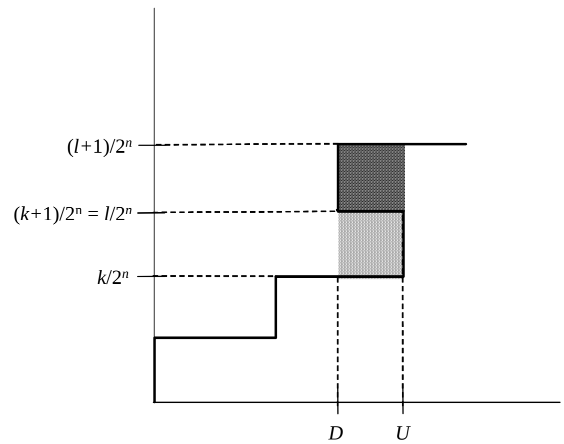

Figure 5: Making $(\widetilde{H}^n)^{-1}$ monotonic

For the second condition in Case 1, notice that the inequality $(c^{-1})'(l(z))z - \int_z^1 (c^{-1})'(l(r))dr \geq \lambda$ for $z > z_{\max}$ reduces to $(c^{-1})'(l(1))(2z_{\max} - 1) \geq \lambda$ by taking the limit as z tends to $z_{\max}$. For Case 2, notice that the left-hand side of (11) for $z = z_{\min}$ is equal to $(c^{-1})'(0)z_{\min} - \int_{z_{\min}}^1 (c^{-1})'(l(r))dr$, and the limit of the left-hand side of (11) as z tends to $z_{\max}$ is $(c^{-1})'l(1)\, z_{\max} - \int_{z_{\max}}^1 (c^{-1})'(l(r))dr$. This yields the condition in Case 2, as $z_{\min} = z_{\max}$.

Proof of Proposition 8. Notice that by (11) and the first part of (12), and the assumption that $(c^{-1})'$ is decreasing, we have that

$$(c^{-1})'(0)z_{\min} - \int_{z_{\min}}^1 (c^{-1})'(l(r))dr = \lambda. \tag{14}$$

Let $z_{\min}^m$, $z_{\max}^m$, and λ^m denote $z_{\min}$, $z_{\max}$, and λ for a given m. As argued in the proof of Proposition 7, $z_{\min}^m < z_{\max}^m$ for all m. We claim that λ^m weakly increases with m. Suppose to the contrary that $\lambda^{m'} > \lambda^{m''}$ for some $m' < m''$.

By (14), $z_{\min}^{m'} > z_{\min}^{m''}$. In particular, we have (a): $(G_{\max}^{m'})^{-1}(z) = 0 \leq (G_{\max}^{m''})^{-1}(z)$ for all $z \leq z_{\min}^{m'}$, and the inequality is strict for $z \in (z_{\min}^{m''}, z_{\min}^{m'})$. By (11) and the fact that the left-hand side of (11) is decreasing in $G^{-1}(z)$, we have (b): $(G_{\max}^{m'})^{-1}(z) \leq (G_{\max}^{m''})^{-1}(z)$ for all $z \in (z_{\min}^{m'}, \min\{z_{\max}^{m'}, z_{\max}^{m''}\}]$. If $z_{\max}^{m'} \geq z_{\max}^{m''}$, then we have (c): $(G_{\max}^{m'})^{-1}(z) \leq m' < (G_{\max}^{m''})^{-1}(z) = m''$ for $z > \min\{z_{\max}^{m'}, z_{\max}^{m''}\}$. If $z_{\max}^{m'} < z_{\max}^{m''} \leq 1$, then $(c^{-1})'(l'(1))(2z_{\max}^{m'} - 1) \geq \lambda^{m'}$ (by the second part of (12)). But (11) implies that $(c^{-1})'\left(l''\left(z_{\max}^{m'}\right)\right)(2z_{\max}^{m'} - 1) \leq \lambda^{m''}$. Thus, $l''\left(z_{\max}^{m'}\right) \geq l'(1)$, and this is possible only when (c) is satisfied (except

$(G_{\max}^{m''})^{-1}(z) \leq m''$). Now, (a), (b), and (c) imply that the budget constraint cannot be satisfied with equality by both $G_{\max}^{m'}$ and $G_{\max}^{m''}$, which completes the argument.

Notice now that λ^m converges to some λ as m diverges to infinity, because otherwise the condition $(c^{-1})'(l(1))(2z_{\max}^m - 1) \geq \lambda^m$ would be violated. Thus, $G_{\max}^m$ converges to a distribution G. This G may have an atom only at 0, because otherwise the budget constraint would be violated for sufficiently large m.[26]

Proof of the claim from Section 6. To formalize the problem, we consider the prize structure that maximizes the expected aggregate effort of the fraction ε of the players with the highest efforts. We then take ε to 0. For any ε, it is straightforward to show that the optimal prize distribution in the limit setting approximates the optimal prize structures in large contests. Thus, it suffices to consider the optimal prize distributions in the limit setting.

In this setting, we known that the measure ε of agents with the highest efforts are those with the highest types, i.e., those with types x for which $F(x) \geq 1-\varepsilon$. Consider first linear costs, i.e., $U(x,y,t) = xh(y) - t$. From (2) we obtain that the aggregate effort of the measure ε of the agents with the highest efforts are $\int_{x^*}^1 \left(xh\left(y^A(x)\right) - \int_0^x h\left(y^A(z)\right)dz\right) f(x)\,dx$, where $F(x^*) = 1-\varepsilon$. From this, it is clear that setting $y^A(x) = G^{-1}(F(x)) = 0$ for $x < x^*$ is optimal. We can therefore rewrite the target function as $\int_{x^*}^1 (xh\left(y^A(x)\right) - \int_{x^*}^x h\left(y^A(z)\right)dz) f(x)\,dx$. From this, it is easy to see that for sufficiently small ε it is optimal to set $y^A(x) = G^{-1}(F(x)) = 1$ for $x \geq x^*$. Indeed, changing the order of integration gives $\int_{x^*}^1 h(y^A(x))(xf(x) - (1-F(x)))dx$, so increasing $y^A(x)$ slightly increases the integrand by at least $h'(y^A(x))(xf(x) - \varepsilon)$, and $f(x)$ is assumed continuous and positive on $[0,1]$ and is therefore bounded away from 0.

Now consider non-linear costs, i.e., $U(x,y,t) = xh(y) - c(t)$. It is again optimal to set $y^A(x) = G^{-1}(F(x)) = 0$ for $x < x^*$, so we can again rewrite the target function as $\int_{x^*}^1 c^{-1}(xh(y^A(x)) - \int_{x^*}^x h(y^A(z))dz) f(x)\,dx$. Assuming that (the positive and continuous) $(c^{-1}(z))'$ lies in an interval $[L,H]$, $L>0$, for z in $[0, h(1)]$, we can apply the same intuition and conclude that for sufficiently small ε it is optimal to set $y^A(x) = G^{-1}(F(x)) = 1$ for $x \geq x^*$. This is because increasing $y^A(x)$ slightly increases the target function by at least $h'\left(y^A(x)\right)(xf(x)L - \varepsilon H)$.

[26]See footnote 25.

References

[1] **Barut, Yasar, and Dan Kovenock.** 1998. "The Symmetric Multiple Prize All-Pay Auction with Complete Information." *European Journal of Political Economy*, 14(4): 627-644.

[2] **Bodoh-Creed, Aaron, and Brent Hickman.** 2015. "College Assignment as a Large Contest." Mimeo.

[3] **Billingsley, Patrick.** 1995. "Probability and Measure." John Wiley and Sons, Inc.

[4] **Bulow, Jeremy I., and Jonathan Levin.** 2006. "Matching and Price Competition." *American Economic Review*, 96(3): 652-68.

[5] **Clark, Derek J., and Christian Riis.** 1998. "Competition over More than One Prize." *American Economic Review*, 88(1): 276-289.

[6] **Glazer, Amihai and Refael Hassin.** 1988. "Optimal Contests." *Economic Inquiry*, 26(1): 133-143.

[7] **Konrad, Kai A.** 2007. "Strategy in Contests - an Introduction." Berlin Wissenschaftszentrum Discussion Paper SP II 2007 – 01.

[8] **Moldovanu, Benny, and Aner Sela.** 2001. "The Optimal Allocation of Prizes in Contests." *American Economic Review*, 91(3): 542-58.

[9] **Moldovanu, Benny, and Aner Sela.** 2006. "Contest Architecture." *Journal of Economic Theory*, 126(1): 70-96.

[10] **Myerson, Roger B.** 1981. "Optimal Auction Design." *Mathematics of Operations Research*, 6(1): 58-73.

[11] **Olszewski, Wojciech, and Ron Siegel.** 2015. "Large Contests." *Econometrica*, forthcoming.

[12] **Reny, Philip J.** 1999. "On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Games." *Econometrica*, 67(5): 1029-1056.

[13] **Siegel, Ron.** 2010. "Asymmetric Contests with Conditional Investments." *American Economic Review*, 100(5): 2230-2260.

[14] **Xiao, Jun.** 2013. "Asymmetric All-Pay Contests with Heterogeneous Prizes." Mimeo.