Effort-Maximizing Contests

Wojciech Olszewski and Ron Siegel*

March 2016

Abstract

We introduce a new approach to analyzing contest design questions in settings that accommodate heterogenous prizes and many, possibly asymmetric contestants who may have private information about their ability or prize valuations.

We use the approach to derive the prize structure that maximizes contestants' effort. 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 makes it possible to derive 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 the audiences at Bocconi University, Boston College, Carnegie Mellon University, Duke University, Kyoto University, the University of Melbourne, the University of New South Wales, the University of Pittsburgh, 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, the Midwest Economic Theory Conference at Pennsylvania State University, the University of Hong Kong, the Conference on Economic Design in Istanbul, 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 presents a novel approach to the design of optimal contests with many contestants.[1] The goal of the approach is to address a central challenge in the analysis of contest design questions, namely, how to optimize over the set of feasible contests when solving even a single contest is difficult or impossible. The idea is to solve a different optimization problem, which is more manageable and whose solution approximates the optimal contest. This is made possible by the results of Olszewski and Siegel (2016) (henceforth: OS), who showed that for large contests players' equilibrium behavior and resulting allocation of prizes are approximated in a strong sense by a particular single-agent mechanism that allocates a continuum of prizes to a continuum of agent types. The problem of optimizing over a set of contests can be translated to optimizing over the set of mechanisms that approximate the equilibria of these contests, so that the solution to the mechanism design problem approximates the one to the contest design problem. The rich set of tools available in the mechanism design literature can be brought to bear on the mechanism design problem, which greatly increases its tractability.

The contest environment we consider has many contestants, who may be ex-ante asymmetric and may or may not have private information about their ability or prize valuations. 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. The combination of asymmetric players, incomplete information, and heterogeneous prizes makes most specifications of these contests impossible to solve, and thus particularly suitable for our contest design approach.

We apply the approach to investigate effort-maximizing contests. Our first result concerns settings in which the designer has some control over the composition of the pool of

[1] 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 (NSF) 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).

contestants who may participate in the contest. Given any distribution of prizes awarded in the contest, we show that a first-order stochastic dominance shift in the distribution of abilities in the player population always leads to an increase in their aggregate equilibrium effort. This is true regardless of the asymmetry among players or the structure of their private information. While this result may seem straightforward, it does not hold for asymmetric contests with a small number of players. For some small contests, such a shift in the ability distribution reduces the aggregate equilibrium effort. This is because in small contests there are countervailing effects. On the one hand, more able players tend to exert more effort. On the other hand, players may become more asymmetric, which discourages less able players, who exert less effort, and consequently also reduces the effort of more able players, who now face weaker competition. But in large contests the second effect disappears, because players effectively compete against those with similar abilities.

Our main results identify the prize distribution that maximizes players' aggregate equilibrium effort given a prize budget and a population of possibly asymmetric and privately informed players. Some results of this type exist in the literature, but they are often partial or limited to environments with ex-ante identical players or identical prizes, and rely on restrictive functional forms and informational assumptions.[2] 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 need not be (and often are not) identical. Our approach shows that the mechanism design problem corresponding to this contest design problem is to maximize the expected revenue in the mechanism across all prize distributions that satisfy a budget constraint. We show that this is in fact a calculus of variations problem, which can be solved by standard methods.

The solution to this problem demonstrates that the optimal prize distribution is closely linked to the curvature of players' effort cost and prize valuations. When effort costs are convex (and prize valuations are linear) or when prize valuations are concave (and effort costs are linear), which corresponds to risk-averse players when prizes are denominated in monetary terms, it is optimal to award numerous prizes whose value gradually decreases with players' ranking**.** This is true regardless of the structure of players' private information or the distribution of abilities in the player population. The ability distribution, effort costs, and prize valuations determine the precise optimal distribution of the prizes. In general, many players do not obtain a prize, but if the marginal cost of the first unit of effort is 0, it is optimal to award a prize to nearly all players, which induces participation by almost all

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

player types. This shows that despite their similarities, the optimal prize distribution and resulting player behavior when effort costs are linear and prize valuations are concave may be qualitatively different than when effort costs are convex and prize valuations are linear. We obtain a complete characterization of the prize distribution, which is described in closed form once functional forms for the ability distribution in the player population and effort costs or prize valuations are specified.[3] Awarding numerous prizes of different value is also optimal when players have a combination of convex costs and concave prize valuations.

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), a winner-take-all contest with a single grand prize is optimal. In particular, one does not need to know the precise effort costs, prize valuations, or distribution of abilities in the player population. Our qualitative characterization of the optimal prize structure, which applies to all equilibria, is 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.[4] A single grand prize remains optimal when players have a combination of concave costs and convex prize valuations.

We also consider how restricting the size of the highest prize that may be awarded affects the effort-maximizing prize distribution.[5] This restriction can be the result of policy, fairness considerations, or technological limitations.[6] The optimal prize distribution under this restriction is related to the one in the unrestricted case in a natural way: instead of a single grand prize, multiple maximal prizes are awarded, and instead of a range of prizes of different value, a range of prizes of different value along with multiple maximal prizes are awarded. In both cases, if the prize budget is large enough relative to the restriction, some of the budget will optimally remain unused. When this happens, even with concave prize utility it is optimal to award only multiple maximal prizes (just like with convex prize utility or concave costs), but with linear prize utility and convex costs in which the marginal cost of

[3] This is demonstrated by the examples in Sections 5.4 and 6.3.

[4] They showed that a single 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 6.2 provides a more detailed comparison of our results for linear prize valuations to those of MS.

[5] We derive the results in the previous paragraphs by solving the maximization problem with this constraint and taking the restriction on the highest possible prize to infinity.

[6] For example, several NSF categories have maximal awards. In a college admissions setting, it may be possible to fund additional admissions to existing universities, but creating another university to generate prizes whose value exceeds the value of admission to the best existing universities is likely infeasible.

the first unit of effort is 0 it remains optimal to award a range of heterogeneous prizes, along with multiple maximal prizes. This is another qualitative difference between the optimal prize distribution with concave prize utility and the one with convex effort costs. Regardless of the prize budget or the restriction on the maximal prize, it is never optimal to award multiple identical intermediate prizes - any identical prizes awarded must all be the highest possible prize.[7]

In addition to characterizing the effort-maximizing prize structure, our approach characterizes the equilibrium efforts of all players in the optimal contest. The approach also uncovers a novel connection between effort-maximizing contests with many players who have linear effort cost and Myerson's (1981) optimal auction with a single buyer. Myerson's optimal auction and the mechanism that approximates the optimal contest both implement monotone allocations that maximize the "virtual surplus" for each type, i.e., the allocation value minus the information rents accrued to higher types. The difference is in the constraints: the approximating mechanism is subject to an overall prize budget constraint, whereas Myerson's mechanism is subject to a capacity constraint type by type, since a single, exogenously given object is being auctioned.

The intuition why the approximating mechanism also maximizes the virtual surplus is as follows. In small contests increasing the value of a prize affects players in complicated ways, encouraging competition by some players and discouraging others (see MS). But in large contests, increasing the value of a prize has only two, clear-cut effects. First, it increases local competition for this prize by players with types close to the one who is allocated the prize in equilibrium, and the increased competition precisely exhausts the entire increase in the prize value. Second, it reduces the effort of all higher types, since they can now slack off and obtain the same utility by winning a slightly lower prize than they previously did. These two effect are captured by a "virtual effort" expression identical to Myerson's virtual surplus. This also helps to explain why our approach makes solving for the optimal prize structure a tractable problem, at least when costs are linear.

With convex costs, a similar expression describes the effect of increasing a prize on the aggregate cost of effort, but this does not provide a simple expression for the effect on the aggregate effort, because for every type the slope of the inverse cost function depends on that type's equilibrium effort. Nevertheless, this intuition helps to explain the qualitative differences between the optimal prize distribution with concave prize utility and linear effort

[7]Multiple identical intermediate prizes may be part of an effort-maximizing prize distribution when players' effort costs or prize valuations are not convex or concave. Our characterization of the conditions for the optimal prize distribution can be easily used to analyze such settings.

costs and the one with linear prize utility and convex costs in which the marginal cost of the first unit of effort is 0. In the former case, the virtual effort per unit of prize utility for types below a certain threshold is negative, so the number of prizes is always optimally restricted, whereas that of types above the threshold is positive, so any awarded prizes are increased as much as the budget and the restriction on the highest possible prize allow. In the latter case, the effect on the aggregate effort of awarding a small prize even to a low type is always positive, since for a small amount of prize utility a marginal effort cost of 0 implies that the level of effort that exhausts the prize utility is larger than the corresponding decrease in the effort of higher types who exert positive effort, because they have positive marginal effort costs.

Finally, our approach can be used to investigate contests that are optimal with respect to other goals.[8] One example, which we analyze in the appendix, is maximizing the highest efforts, rather than the aggregate, or average, effort. We show that the optimal prize distribution with many contestants is a small number of prizes of the highest possible value. This is true regardless of players' effort costs, prize utilities, private information, or distribution of abilities. This prize distribution is consistent with many crowd-sourcing contests, which involve a large pool of contestants and typically award a small number of prizes. Examples of other questions that can be investigated using our approach include how to maximize effort when prizes are costly (instead of having a fixed budget) and how to maximize social welfare when contestants' effort may be productive or wasteful. We leave these and other potential questions for future work.

The rest of the paper is structured as follows. Section 2 describes the contest environment. Section 3 describes the mechanism design framework and OS's approximation result. Section 5 analyzes the optimal prize structure when players have linear costs. Section 6 analyzes the optimal prize structure when players have linear prize valuations. Section 7 concludes. The appendix contains the proofs omitted from the main text.

2 Asymmetric contests

In the contests we consider, 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 $m \geq 1$.[9] Prize m is the highest possible prize. As discussed in the

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

[9] The restriction $m \geq 1$ is a convenient but inessential normalization.

introduction, such a prize restriction arises naturally in some settings. The results we obtain apply to settings in which such a restriction is not imposed by letting m diverge to infinity. 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, so it is without loss of generality to have the same number of prizes as players. 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 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 prize utility function h and cost function c are continuously differentiable and strictly increasing.[10] Notice that (1) can accommodate private information about ability by dividing each player's utility by x to obtain $h(y) - c(t)/x$, which has no effect on the results. We assume that sufficiently high effort levels are prohibitively costly, that is, for large enough t we have $h(1) < c(t)$, so no player chooses effort higher than $c^{-1}(h(1))$. 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 (2016)). 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 5.3 and 6.2 provide results also for convex h and concave c.

Our analysis will focus on large contests, that is, contests with a large n. We will consider sequences of contests, and refer to a contest with n players and n prizes as the "n-th contest" in the sequence. Every contest has at least one (mixed-strategy) Bayesian Nash equilibrium.[11]

[10] We also allow for $h'(0) = \infty$, where the derivative is a right-derivative.

[11] This follows from Corollary 5.2 in Reny (1999), because the mixed extension is better-reply secure.

3 Using mechanism design to study the equilibria of large 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 (2016)). Our approach to contest design builds on the technique for studying the equilibria of large contests, which was developed in OS. We now describe this technique, which allows us to approximate the equilibrium outcomes of large contests by considering the mechanism that implements a particular allocation of a continuum of prizes to a continuum of agent types.

3.1 Limit distributions

We first formalize a requirement that the contests in the sequence be "sufficiently similar" as the number of players n grows large. 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. 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 require that F^n converge in weak*-topology to a distribution F that has a continuous, strictly positive density f, and that G^n converge to some (not necessarily continuous) distribution G.[12] Notice that the restriction on F does not imply a similar restriction on distributions F_i^n of players' types, so these distributions may have gaps and atoms.[13]

In particular, the convergence of F^n and G^n to limit distributions F and G accommodates as a special, extreme case sequences of complete-information contests with asymmetric players, in which each player i's type distribution F_i^n in the n-th contest is a Dirac measure. A simple way to see this is to first choose the desired limit distributions F and G and then set player i's deterministic type in the n-th contest to be $x_i^n = F^{-1}(i/n)$ and prize j in the n-th contest to be $y_j^n = G^{-1}(j/n)$, where

$$G^{-1}(z) = \inf\{y \in [0,m] : G(y) \geq z\} \text{ for } 0 \leq z \leq 1.$$

[12] Convergence in weak*-topology can be defined as convergence of *cdf*s at points at which the limit *cdf* is continuous (see Billingsley (1995)).

[13] The restriction on F precludes a limit mass of players that have an atom at a particular type, as is the case when there is a non-vanishing fraction of identical players in a contest with complete information.

Then, the n-th contest is one of complete information, F^n converges to F, and G^n converges to G.

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 (positive) 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 heterogeneous prizes, and the value of prize j to player i is ij/n^2. These contests were studied by BL, who considered hospitals that have a common ranking for residents and compete for them by offering identity-independent wages.[14]

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

As will be stated in the next subsection, the mechanism that approximates the equilibrium outcomes of large contests implements the *assortative allocation*, which assigns to each type x prize $y^A(x) = G^{-1}(F(x))$. That is, the location in the prize distribution of the prize assigned to type x is the same as the location of type x in the type distribution. It is well known (see, for example, Myerson (1981)) that the unique incentive-compatible mechanism that implements the assortative allocation and gives type $x = 0$ a utility of 0 specifies for

[14] Xiao (2016) 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.

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 no prize to each type $x \leq 1-p$ and assigns a positive prize to each type $x > 1-p$, 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 large contests are approximated by the unique mechanism that implements the assortative allocation.

Theorem 1 *(OS) 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 $t^A(x_i^n)$.*

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

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

where f is the density of the limit aggregate type distribution F. More precisely, we refer to the aggregate expected 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).*

Corollary 1 implies that the structure of players' private information has a vanishing effect on the aggregate effort in large contests, since (3) depends on players' type distributions only through the limit distribution F.

4 Changing the distribution of players' types

We begin the contest design analysis by considering how changing the distribution of players' types affects their aggregate equilibrium effort. This is relevant to settings in which the designer has some control over the composition of the pool of contestants who may participate in the contest. A reasonable conjecture is that opting for players with higher types increases aggregate equilibrium effort. This is, however, not always the case in contests with a small number of players. To see this, consider a two-player all-pay auction with complete information and one prize. The prize is $y = 1$, the prize valuation function satisfies $h(1) = 1$, and the cost function is $c(t) = t$. Players' publicly observed types satisfy $0 < x_1 < x_2 < 1$. It is well known (Hillman and Riley (1989)) that in the unique equilibrium player 2 chooses a bid by mixing uniformly on the interval $[0, x_1]$ and player 1 bids 0 with probability $1 - x_1/x_2$ and with the remaining probability mixes uniformly on the interval $[0, x_1]$. The resulting aggregate effort is $x_1/2 + (x_1)^2 / (2x_2)$, which monotonically increases in x_1 and monotonically decreases in x_2. Thus, an increase in player 2's type, even when accompanied by a small increase in player 1's type, decreases aggregate effort. The intuition is that the increased asymmetry between the players, which discourages competition, outweighs the increase in their types, which encourage higher effort.

This example shows that a first-order stochastic dominance (FOSD) shift in players' aggregate type distribution may increase or decrease aggregate effort. Moreover, since solving for the equilibria of asymmetric contests with a small number of players and multiple prizes or incomplete information is difficult or impossible, it is likely difficult to obtain general conditions under which a FOSD shift increases aggregate effort in such contests.

The following result shows that in large contests a FOSD shift in players' aggregate type distribution always increases the aggregate equilibrium effort.

Proposition 1 *Take a sequence of contests with limit type distribution F, and for every contest in the sequence leave the set of prizes unchanged but change the set of players to obtain a new limit type distribution $\widetilde{F}$ that FOSD F. Then, for large enough N and any $n \geq N$, the aggregate effort in any equilibrium of the n-th contest in the modified sequence exceeds the aggregate effort in any equilibrium of the n-th contest in the original sequence.*

Proof: By Corollary 1, it suffices to show that a FOSD shift in F increases the value of (3). By (2) and the definition of y^A, this value is

$$\int_0^1 \left(c^{-1} \left(xh\left(G^{-1}\left(F\left(x\right)\right)\right) - \int_0^x h\left(G^{-1}\left(F\left(z\right)\right)\right) dz \right) \right) f\left(x\right) dx = \tag{4}$$

$$\int_0^1 \left(c^{-1} \left(xL\left(x\right) - \int_0^x L\left(z\right) dz \right) \right) f\left(x\right) dx, \tag{5}$$

where $L(x) = h\left(G^{-1}(F(x))\right)$. By looking at the areas below the graphs of L and L^{-1} in the rectangle $[0,x] \times [0, L(x)]$, we have that $\int_0^{L(x)} L^{-1}(r)dr + \int_0^x L(z)\,dz = xL(x)$.[15] Substituting into (5) we obtain

$$\int_0^1 \left(c^{-1}\left(\int_0^{L(x)} L^{-1}(r)\,dr\right)\right) f(x)\,dx = \int_0^1 \left(c^{-1}\left(\int_0^{L\left(F^{-1}(z)\right)} L^{-1}(r)\,dr\right)\right) dz, \tag{6}$$

where the equality follows from the change of variables $z = F(x)$.[16] Since a FOSD shift in F decreases F and therefore L pointwise, it increases F^{-1} and L^{-1} pointwise, and therefore increases (6).

The intuition for Proposition 1 is that in a large contest competition is "localized" in the sense that players compete against players with similar types.[17] Therefore, any decrease in local competition between some types resulting from a FOSD shift in players' type distribution is more than compensated for by an increase in local competition between some higher types.

5 Optimal prize distribution when players may not be risk neutral

We now investigate the prize distributions in large contests that maximize the aggregate effort subject to a budget constraint, taking the distribution of contestants' abilities as given. The budget 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.

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

[15] The definition of L^{-1} is $L^{-1}(r) = \inf\{x \in [0,1] : L(x) \geq r\}$.

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

[17] A discussion of this phenomenon appears in Bulow and Levin (2006).

maximal aggregate effort, subject to the budget constraint that the average prize does not exceed some value C^n that converges to C.[18]. Formally,

$$G_{\max}^n \in \arg\max_{G^n} \left\{ \sum_{i=1}^{n} \int_0^1 t_i^n(x)\, dF_i^n(x) : \begin{array}{c} t_i^n \text{ are an equilibrium of the contest with} \\ \text{type distributions } F_i^n \text{ and prizes } y_1^n \leq \cdots \leq y_n^n \\ \text{described by } G^n \text{ with } \left(\sum_{j=1}^n y_j^n\right)/n \leq C^n \end{array} \right\}.$$

Denote by $M_{\max}^n$ the corresponding maximal aggregate effort attained by $G_{\max}^n$.

We would like to characterize $G_{\max}^n$ and $M_{\max}^n$ for large n by solving for the corresponding objects in the limit setting. For this, denote by $\mathcal{M}$ the set of prize distributions that maximize (3) subject to the budget constraint $\int_0^m y dG(y) \leq C$. 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) subject to the budget constraint. Finally, consider any metrization of the weak*-topology on the space of prize distributions.

Proposition 2 *1. For any $\varepsilon > 0$, there is an N such that for every $n \geq N$, $G_{\max}^n$ 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 $G_{\max}^n$ converges to $G_{\max}$ in weak*-topology. 2. $M_{\max}^n/n$ 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 $M_{\max}^n/n$.*

Part 1 of Proposition 2 shows that the optimal 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 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 $y_j^n = G^{-1}(j/n)$ for $j = 1, ..., n$ generates, for large contests, aggregate

[18] 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 $G_{\max}^n$ 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.

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.[19]

For the remainder of the section, we consider players who have linear effort costs but may not be risk neutral, so $U(x,y,t) = xh(y) - t$. The curvature of h captures players' risk attitudes regarding lotteries over prizes. In Section 6, we obtain corresponding results for $U(x,y,t) = xy - c(t)$. Our analysis will show that different risk attitudes and curvatures of cost functions play a similar (but not identical) role in contest design. The beginning of Section 6 explains how our analysis extends to the general utility function (1).

5.1 Reduction to a calculus of variations problem

By Proposition 2, we can focus on solving the following problem:

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

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.

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:

$$\int_0^1 t^A(x) f(x)\, dx = \int_0^1 h\left(y^A(x)\right)\left(x - \frac{1-F(x)}{f(x)}\right) f(x)\, dx. \tag{7}$$

To gain some intuition for why (7) approximates the average effort in large contests, observe that (7) 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. This net increase in revenue, or "virtual value," also coincides with a monopolist's marginal revenue (Bulow and Roberts (1989)). In a large contest, increasing the prize that type x obtains also allows

[19] 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.

the designer to capture the entire increase in surplus for this type, because the higher prize increases this type's competition with nearby types until the gain from the higher prize is exhausted. But the prize increase also decreases the 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.[20] The assumption guarantees that the effort-maximizing functions G^{-1} we will derive are non-decreasing, without imposing a monotonicity constraint.

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

We rewrite (7) using the change of variables $z = F(x)$ to obtain

$$\int_0^1 h\left(G^{-1}(z)\right)\left(F^{-1}(z) - \frac{1-z}{f\left(F^{-1}(z)\right)}\right) dz = \int_0^1 h\left(G^{-1}(z)\right) k(z)\, dz, \tag{8}$$

where $k(z) = F^{-1}(z) - (1-z) / f\left(F^{-1}(z)\right)$.[21] The value $k(z)$ can be interpreted as the marginal "virtual effort." That is, the additional aggregate effort that can be extracted by a marginal increase in the prize utility h resulting from an increase in the prize assigned by the assortative allocation to the type whose percentile ranking is z in the limit aggregate type distribution. This additional effort is the combination of the increase in effort of the agent type with percentile ranking z and the decrease in effort of agents of all higher types. By Assumption 1, $k(z)$ strictly increases in z.

We now transform the budget constraint to obtain an equivalent constraint as a function of G^{-1}. 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

$$\int_0^1 G^{-1}(z)\, dz \leq C. \tag{9}$$

This is the desired form of our maximization problem, because maximizing (8) subject to (9) is a calculus of variations problem in variable G^{-1}. Because the allocation is assortative, the inverse prize distribution G^{-1} specifies for each location $z \in [0, 1]$ in the type distribution the prize $G^{-1}(z) \in [0, m]$ allocated to the type in that location, that is, to type x such that $F(x) = z$.

[20] The assumption is implied, for example, by a monotone hazard rate for distribution F.

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

5.2 Conditions describing the solution

We now derive the conditions that must be satisfied by an optimal inverse prize distribution G^{-1}. Because in the assortative allocation higher types obtain higher prizes, there are percentile locations $z_{\min} \leq z_{\max}$ in $[0, 1]$ such that types in percentile locations $z \leq z_{\min}$ in the limit distribution F are each allocated the prize $G^{-1}(z) = 0$ (no prize),[22] types in locations $z > z_{\max}$ are each allocated the highest possible prize $G^{-1}(z) = m$, and types in intermediate locations $z \in (z_{\min}, z_{\max})$ are allocated positive, non-maximal prizes $G^{-1}(z) \in (0, m)$.

If the range of intermediate types is non-trivial, that is, $z_{\min} < z_{\max}$, then the product of the marginal prize utility and the marginal virtual effort must be equal for all the intermediate types. Otherwise, a slight increase the prize allocated to an intermediate type with a higher product accompanied by an equal decrease in the prize allocated to an intermediate type with a lower product would increase the aggregate effort. The product must also be non-negative, otherwise slightly decreasing the prize for intermediate types would increase the aggregate effort. For similar reasons, weak inequalities hold for the lowest and highest types in this range. If there are no intermediate types, that is, $z_{\min} = z_{\max}$, then the product for the highest type allocated no prize must be weakly lower than the product for the lowest type allocated the highest possible prize. These two cases are formally summarized as follows (a rigorous proof is provided in the appendix).

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, $h'(0)\, k(z_{\min}) \leq h'(m)\, k(z_{\max})$.

If $h'(0) = \infty$, then similar arguments show that $k(z_{\min}) = 0$ holds instead of the inequalities in Case 1 and Case 2 that involve $h'(0)$.

5.3 Risk averse, risk neutral, and risk loving players

We now use the conditions of Section 5.2 to characterize the optimal prize distribution for risk averse, risk natural, and risk loving players, who have concave, linear, and convex prize utility functions h. We first show that players' risk attitude does not affect the optimal prize distribution when the budget constraint does not bind, and then show how players' risk attitudes affect the optimal prize distribution when the budget constraint binds.

Suppose that $C \geq m(1 - F(x^*))$, where $x^* \in (0, 1)$ is the unique type that satisfies $x^* - (1 - F(x^*))/f(x^*) = 0$. This is the type for whom the marginal virtual effort is 0.

[22] The inequality $z \leq z_{\min}$ is weak because G^{-1} is left-continuous as the inverse of a probability distribution.

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 $[0,1]$. Types $x < x^*$ have negative marginal virtual effort, so the value of the integrand in (7) is negative, and types $x > x^*$ have positive marginal virtual effort, so the value is positive. Denote by $z^* = F(x^*) \in (0,1)$ the percentile location in the type distribution of type x^*, so $k(z^*) = 0$. Then, since the marginal prize utility is always positive, optimizing the integrand in (8) separately for each $z \in [0,1]$ leads to assigning the lowest possible prize $G^{-1}(z) = 0$ to types in locations $z \leq z^*$, and assigning the highest possible prize m to types in locations $z > z^*$.[23] 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 3 *If $C \geq m(1 - F(x^*))$, then for any function h the optimal prize distribution consists of a mass $1 - F(x^*) \in (0,1)$ of the highest possible prize m and a mass $F(x^*)$ of prize 0.*

Proposition 3 shows that when the budget constraint does not bind, it is optimal to award a set of identical prizes, as in the all-pay auctions studied by CR, rather than, for example, heterogeneous prizes, as in the all-pay auctions studied by BL, or a combination of identical and heterogeneous prizes. Moreover, some of the budget is optimally left unused. This is analogous to a monopolist limiting the quantity sold. Relaxing the constraint on the highest possible prize (increasing m) optimally leads to an increase in the value of the awarded prizes, but does not change their quantity. Of course, for a fixed budget C, a sufficient increase in m would cause the budget constraint to bind.

We now consider a binding budget constraint, which seems more relevant in practice. So, for the remainder of the section we make the following assumption:

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

This assumption implies that for the optimal prize distribution the budget constraint (9) holds with equality. We now derive the form of the optimal prize distribution for convex and concave functions h, and show that the binding budget constraint leads to a qualitatively different optimal prize distribution in each case, in contrast to the finding with an unlimited budget. We first present the simpler result for convex functions h.

Proposition 4 *If h is weakly convex, so players are risk neutral or risk loving, then the optimal prize distribution consists of a mass C/m of the highest possible prize and a mass $1 - C/m$ of prize 0.*

[23] This corresponds to Case 2 in Section 5.2, with $z_{\min} = z_{\max} = z^*$.

Proof: Weak convexity implies that $z_{\min} = z_{\max}$. Indeed, since h' and G^{-1} are 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 4 shows that awarding identical prizes remains optimal when the budget constraint binds, provided that agents' marginal prize utility is non-decreasing. 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 5 *1. If h is weakly concave (but not linear on $[0, m]$), so players are weakly risk averse (but not risk neutral), then any optimal prize distribution assigns a positive mass to the set of intermediate prizes $(0, m)$. In addition, any optimal prize distribution may have atoms only at 0 (no prize) and m (the highest possible prize). 2. If h is strictly concave, then any optimal prize distribution awards all prizes up to the highest prize awarded. That is, G strictly increases on $[0, G^{-1}(1)]$.*

Proof: Weak concavity implies that $z_{\min} < z_{\max}$. 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 (9). This yields the first part of 1. For the second part, notice that $G^{-1}(z)$ strictly increases in z on interval $(z_{\min}, z_{\max})$, so G does not have atoms there. 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. For 2, the same observation shows that if h' is strictly decreasing, then G^{-1} is continuous on $(z_{\min}, z_{\max}]$. If G^{-1} were not right-continuous at $z_{\min}$, then the fact that $h'(0)\,k(z_{\min}) \leq \lambda$ and h' is strictly decreasing would violate $h'(G^{-1}(z))\,k(z) = \lambda$ on $(z_{\min}, z_{\max}]$.[24] Thus, G strictly increases on $[0, G^{-1}(z_{\max})]$. If $z_{\max} = 1$, we are done. If $z_{\max} < 1$, then $G^{-1}(z_{\max}) = G^{-1}(1) = m$, otherwise the fact that $h'(m)k(z_{\max}) \geq \lambda$ and h' is strictly decreases would violate $h'(G^{-1}(z_{\max}))\,k(z_{\max}) = \lambda$.

It may be tempting to attribute the qualitative difference between the optimal prize distributions with convex and concave prize valuations to the difference in players' risk attitudes: lotteries between no prize and the highest possible prize may be riskier than lotteries over a range of intermediate prizes, so the former can elicit more effort when players

[24] If $h'(0) = \infty$, then $k(z_{\min}) = 0$, so if G^{-1} were not right-continuous at $z_{\min}$ the product $h'\left(G^{-1}(z)\right)k(z)$ would be strictly positive for any $z \in (z_{\min}, z_{\max}]$ but approach 0 as $z \downarrow z_{\min}$, so could not be constant on $(z_{\min}, z_{\max}]$.

are risk loving, and the latter when they are risk averse. This intuition is misleading, however, because in large contests almost all player types are nearly certain of the prize they receive in equilibrium (Theorem 1).[25] Instead, what drives the qualitative difference is how the marginal prize utility changes as the prize increases. Because the marginal prize utility is always positive, absent the budget constraint it is optimal to award the lowest possible prize to types with negative marginal virtual effort and the highest possible prize to types with positive marginal virtual effort. The budget constraint introduces a tradeoff between the prizes allocated to different types. This tradeoff is optimally resolved by comparing the product of the marginal prize utility and the marginal virtual effort across types. Since the marginal virtual effort increases in type, what determines the comparison is whether the marginal prize utility increases or decreases in the prize, which correspond to convex and concave prize valuations. In the former case, increasing the prize increases the product, so it is optimal to allocate the highest possible prize to the highest types. In the latter case, increasing the prize decreases the product, so continuity of the marginal virtual effort implies that as we increase the prizes awarded to some types, it becomes increasingly attractive to award prizes to slightly lower types. The optimal prize distribution equates the product across all types allocated intermediate prizes. Such types exist, because the budget constraint implies that not all types can be awarded the highest possible prize.

Our next result, whose proof is in the appendix, shows that as the highest possible prize, m, becomes arbitrarily large, weak risk aversion (but not risk neutrality) implies that any optimal prize distribution becomes "continuous" on all non-zero prizes.

Proposition 6 *Let $G^m_{\max}$ be an optimal prize distribution when m is the highest possible prize. If h is weakly concave (but not linear on $[0, m]$), then when m diverges to infinity $G^m_{\max}$ converges to a distribution that may have an atom only at 0 (no prize).*

The proof of Proposition 6 also shows that $z_{\min}$ and $z_{\max}$ (weakly) increase as m increases; and for any $m' < m''$, we have $G^{m''}_{\max}(y) \geq G^{m'}_{\max}(y)$ for $y < m'$ and $G^{m''}_{\max}(y) \leq G^{m'}_{\max}(y)$ for $y \geq m'$.

We now show that when players are risk averse, so h is strictly concave, the constrained maximization problem has an explicit, closed-form solution. As shown in the proof of Proposition 6, $h'(0)\, k(z_{\min}) = \lambda$. Thus,

$$z_{\min} = k^{-1}(\lambda / h'(0)). \tag{10}$$

[25] In addition, if the intuition were correct, we would expect the optimal prize distribution to vary with players' risk attitudes also when the budget constraint does not bind, in contrast to the statement of Proposition 3.

Since $h'\left(G^{-1}\left(z_{\max}\right)\right)k\left(z_{\max}\right)=\lambda$ and h' is decreasing, $h'\left(m\right)k\left(z_{\max}\right)\leq\lambda$. If $z_{\max}<1$, then we also have $h'(m)k\left(z_{\max}\right)\geq\lambda$ (because we are in Case 1 of Section 5.2), so we obtain $h'\left(m\right)k\left(z_{\max}\right)=\lambda$. Thus,

$$z_{\max}=1\text{ or }k^{-1}(\lambda/h'\left(m\right)). \tag{11}$$

In addition,

$$G^{-1}\left(z\right)=\left(h'\right)^{-1}\left(\lambda/k\left(z\right)\right)\ \text{ for }z\in\left(z_{\min},z_{\max}\right] \tag{12}$$

and

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

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

5.4 Example of the optimal prize distribution for risk averse players

To illustrate our solution, let F be uniform and let $h\left(y\right)=\sqrt{y}$, so players are risk averse. Then $k\left(z\right)=2z-1$, $x^*=z^*=1/2$, $h'\left(0\right)=\infty$, $h'\left(s\right)=1/\left(2\sqrt{s}\right)$, and $\left(h'\right)^{-1}\left(r\right)=1/(4r^2)$. The budget constraint is binding if $C<m\left(1-F\left(x^*\right)\right)=m/2$. Since $h'\left(0\right)=\infty$, $z_{\min}=1/2$. Suppose first that $z_{\max}=1$. By (12) and the binding version of (9), we have $\int_{1/2}^{1}\left(2z-1\right)^2/(4\lambda^2)dz=C$. Solving for λ, we obtain $\lambda=1/\sqrt{24C}$. This yields $G^{-1}\left(z\right)=6C\left(2z-1\right)^2$; in particular, $z_{\max}=1$ implies $C\leq m/6$. Thus, we have that

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

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 (8), is $\sqrt{6C}/6$.

Suppose now that $z_{\max}<1$. By (11), $z_{\max}=\lambda\sqrt{m}+1/2$. The binding version of (9) implies that $\int_{1/2}^{\lambda\sqrt{m}+1/2}(\left(2z-1\right)^2/(4\lambda^2))dz+\int_{\lambda\sqrt{m}+1/2}^{1}mdz=C$. Solving for λ, we obtain $\lambda=(3m/4-3C/2)/\left(m\sqrt{m}\right)$. This implies that $G^{-1}\left(z\right)=0$ for $z\in\left[0,1/2\right]$, $G^{-1}\left(z\right)=4\left(2z-1\right)^2m^3/(3m-6C)^2$ for $z\in\left(1/2,(5m/4-3C/2)/m\right]$, and $G^{-1}\left(z\right)=m$ for $z\in\left((5m/4-3C/2)/m,1\right]$. Since $z_{\max}=\lambda\sqrt{m}+1/2<1$, we have $C>m/6$. Thus,

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

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 (8), is $(12C(m - C) + m^2)/(16m\sqrt{m})$.

The following figure depicts these results for $m = 1$.[26]

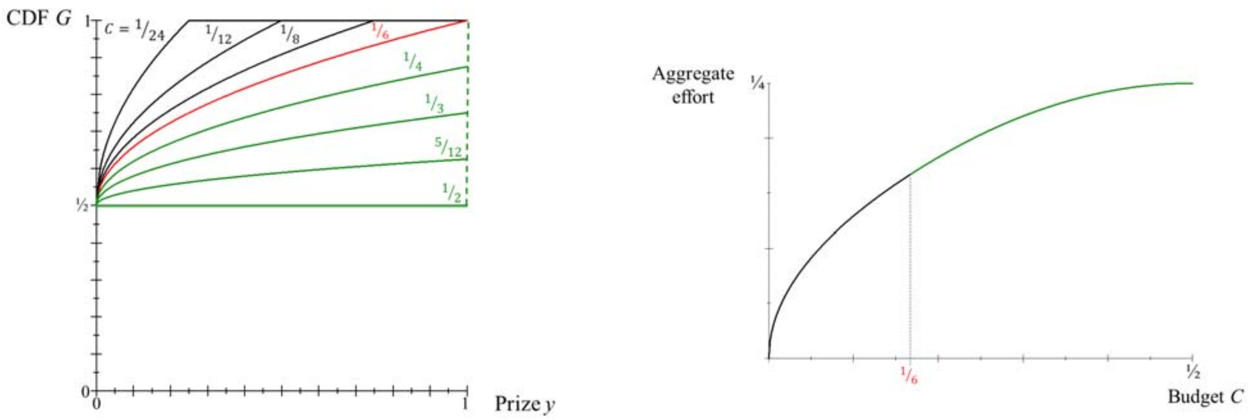

Figure 1: The optimal prize distribution as C increases from 0 to 1/2 (left), and the resulting aggregate effort (right)

To summarize, for $m \in (2C, 6C)$ the optimal prize distribution awards a mass of 0 prizes and a mass of the highest possible prize m, and is continuous and strictly increasing on the interval of prizes $y \in (0, m)$. As m increases, the mass of the highest possible prize awarded decreases (the same happens if C decreases). When $m = 6C$, the optimal prize distribution awards a mass of 0 prizes 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 $m \in (2C, 6C)$, which is the region in which the bound on the highest possible prize binds, and is constant in $m \geq 6C$.[27]

6 Optimal prize distribution when players may have non-linear costs

We now consider risk-neutral players with non-linear costs of effort, so $U(x, y, t) = xy - c(t)$, where $c(0) = 0$ and the cost function c is continuously differentiable and strictly increasing.

[26] While in this paper we do not consider the budget as an optimization variable, the right panel in Figure 1 illustrates that our analysis can be used to determine the optimal budget by equating the marginal aggregate effort to the marginal cost of increasing the budget.

[27] For $m \leq 2C$ the optimal prize distribution has atoms at $y = 0$ and $y = m$ of size $1/2$ each (by Proposition 3).

The discrete contests of MS correspond to this utility function. To simplify the analysis, we assume that the limit type distribution F is uniform, so type x is in percentile location $z = x$. The characterization we derive in Section 6.1 can be extended to general F and h as long as a condition that involves F and c (but not h) holds. This condition coincides with the monotone hazard rate condition when c is linear, and holds for any c if f is weakly decreasing.[28] In particular, the qualitative results we derive for concave costs also hold when players are in addition weakly risk loving, and the ones for convex costs also hold when players are in addition weakly risk averse.

6.1 Conditions describing the solution

The arguments from the beginning of Section 5, including Proposition 2, also apply to the present case. Thus, the effort-maximizing prize distributions in large contests are approximated by the prize distributions that maximize (3) subject to (9). But proceeding as in Section 5.1 to obtain a simple calculus of variations problem is not possible. A key step in Section 5.1 was to obtain (7). This step relied on the linearity of the cost function c to separate (5), obtained by substituting (2) into (3), into the sum of an integral and a double integral on which integration by parts is performed. This delivered a convenient expression for the virtual effort, which allowed us to consider type by type the effect of the prize allocated to this type on the aggregate effort, since this effect did not depend on the prizes allocated to other types. With a non-linear effort cost, this effect depends on the prizes allocated to other types, since these prizes determine each type's effort level, which in turn determines the marginal cost of effort. Put differently, even with a non-linear effort cost the prize allocated to a given type leads to utility increases for all higher types that are independent of all other types' prizes,[29] but the corresponding decreases in effort requires inverting the cost function, and therefore depend on the higher types' utilities, which are determined by other types' prizes.

To characterize the optimal prize distribution, we will use (6), which the proof of Proposition 1 shows is equivalent to (3). Since $L(x) = h\left(G^{-1}(F(x))\right) = G^{-1}(x)$, (6) is equal

[28] The condition is that

$$\frac{\int_z^1 \left(c^{-1}\right)'(l(r))\,dr}{f\left(F^{-1}(z)\right)}$$

weakly decreases in z, where

$$l(z) = \int_0^{L(x)} L^{-1}(r)\,dr.$$

[29] This follows from the derivation of (7) performed for the variable $c(t)$ instead of t.

to

$$\int_0^1 c^{-1}\left(\int_0^{G^{-1}(z)} G(y)dy\right)dz. \tag{13}$$

Notice that (13) contains G and G^{-1}, unlike (8), which contained only G^{-1}. In particular, (13) does not contain a simple expression for the virtual effort. Nevertheless, conditions similar to those in Section 5.2 can be provided in this case as well. For this, it is convenient to denote by $l(z) = \int_0^{G^{-1}(z)} G(y)dy$ the cost of the effort of type z. The objective function is then

$$\int_0^1 c^{-1}(l(z))\,dz \tag{14}$$

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

We now heuristically develop the conditions that the marginal virtual efforts must satisfy given an optimal prize distribution. To this end, recall that a marginal increase in the prize that a type z obtains increases his effort but decreases the effort of all higher types. From (2) we have that type z's utility in the mechanism is $\int_0^z G^{-1}(r)\,dr$, which is independent of his prize. Thus, a marginal increase in his prize must be accompanied by an increase in effort that precisely offsets the increase in prize utility. This gives

$$0 = \frac{d(zy - c(t(y)))}{dy} = z - c'(t(y))\,t'(y) \Rightarrow$$

$$t'(y) = \frac{z}{c'(t(y))} = z\left(c^{-1}\right)'(c(t(y))) = z\left(c^{-1}\right)'(l(z)),$$

where $t(y)$ is type z's effort as a function of his prize y, and the last equality follows from the definition of $l(z)$. The utility of a type $z' > z$ is $\int_0^{z'} G^{-1}(r)\,dr$, so a marginal increase in type z's prize $y = G^{-1}(z)$ increases type z''s utility by dr. We therefore have

$$dr = \frac{d(y'z' - c(\bar{t}(y)))}{dy} = -c'(\bar{t}(y))\,\bar{t}'(y) \Rightarrow$$

$$\bar{t}'(y) = -\frac{1}{c'(\bar{t}(y))}dr = -\left(c^{-1}\right)'(c(\bar{t}(y)))\,dr = -\left(c^{-1}\right)'(l(z'))\,dr,$$

where $\bar{t}(y)$ is type z''s effort as a function of type z's prize y. Consequently, the effect on the aggregate effort of a marginal increase in the prize of type z, that is, the marginal virtual effort, is

$$(c^{-1})'(l(z))z - \int_z^1 (c^{-1})'(l(r))dr. \tag{15}$$

Equipped with (15), we obtain the analogues of the two cases in Section 5.2, for precisely the same reasons, where $z_{\min}$ and $z_{\max}$ are the lower and upper bound of the range of types that obtain intermediate prizes. (A rigorous proof is provided in the appendix.)

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

$$(c^{-1})'(l(z))z - \int_z^1 (c^{-1})'(l(r))dr = \lambda \tag{16}$$

for $z \in (z_{\min}, z_{\max}]$; in addition,

$$(c^{-1})'(0)z_{\min} - \int_{z_{\min}}^1 (c^{-1})'(l(r))dr \leq \lambda, \text{ and } (c^{-1})'(l(1))(2z_{\max} - 1) \geq \lambda \text{ if } z_{\max} < 1. \tag{17}$$

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

$$(c^{-1})'(0) \leq (c^{-1})'(G(0)\, m). \tag{18}$$

If $c'(0) = 0$, so $(c^{-1})'(0) = \infty$, then Case 2 cannot arise, because the arguments that underlie Case 2 would require $G(0) = 0$, i.e., $z_{\min} = 0$, so every type would obtain the highest possible prize, which would lead to 0 aggregate effort. For Case 1, the proof of Proposition 8 below shows that the first inequality in (17) is replaced with the condition $z_{\min} = 0$.

6.2 Convex, linear, and concave functions c

We now characterize the optimal prize distribution for players with increasing or decreasing marginal effort cost. We will see that even when the budget constraint does not bind the optimal distribution depends on the curvature of the cost function, unlike in Section 5.3, where the curvature of the prize valuation did not affect the optimal prize distribution when the budget constraint did not bind.

We first present the result for concave costs, which shows that the optimal prize distribution coincides with the one for convex prize valuations (Propositions 3 and 4).

Proposition 7 *If players' cost function c is weakly concave, then the optimal prize distribution consists of a mass* $\min\{C/m, 1/2\}$ *of the highest possible prize and a mass* $1 - \min\{C/m, 1/2\}$ *of 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 (16) strictly increases in z. Thus, only the highest and lowest possible prizes are awarded. If the budget constraint binds, then a mass C/m of prize m is awarded.

If the budget constraint does not bind, then the monotonicity of the marginal virtual effort in z also implies that type $z = z_{\min} = z_{\max} > 0$ is determined by

$$(c^{-1})'(mz)z - \int_z^1 (c^{-1})'(mz)dr = 0,$$

because $l(r) = \int_0^{G^{-1}(r)} G(y)dy = \int_0^m G(0)dy = mz$ for any $r > z$. Since $(c^{-1})'(mz) > 0$, this is equivalent to $z - (1 - z) = 0$, or equivalently to $z = 1/2$.

Proposition 7 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 7 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 convex cost functions c.

Proposition 8 *1. Regardless of whether the budget constraint binds, if* c *is weakly convex but not linear on any interval with lower bound* 0*, then any optimal prize distribution assigns a positive mass to the set of intermediate prizes* $(0, m)$*. In addition, any optimal prize distribution may have atoms only at* 0 *(no prize) and* m *(the highest possible prize). 2. If* c *is strictly convex, then any optimal prize distribution awards all prizes up to the highest prize awarded. That is,* G *strictly increases on* $[0, G^{-1}(1)]$*. 3. If the marginal cost of the first unit of effort is* 0*, that is,* $c'(0) = 0$*, then* $z_{\min} = 0$*, so almost every type is awarded a positive prize.*

Proof: The first part of 1 is true because it follows from (18) that $z_{\min} < z_{\max}$. For the second part of 1, an atom at some intermediate prize would mean that $G^{-1}(\underline{z}) = G^{-1}(\overline{z})$ 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 (16) with $z = \overline{z}$ would be higher than with $z = \underline{z}$, which contradicts (16). For 2, notice that $l(z)$ increases discontinuously when $G^{-1}(z)$ does, so if $(c^{-1})'$ is strictly decreasing, a discontinuity in $G^{-1}(z)$ would leads to a discontinuous decrease in the left-hand side of (16). Thus, G^{-1} is continuous on $(z_{\min}, z_{\max}]$. If $z_{\min} = 0$, then $y_{\min} = \lim_{z \downarrow z_{\min}} G^{-1}(z) = 0$, otherwise $G^{-1}(z)$ for $z > 0$ can be "shifted down" to $G^{-1}(z) - y_{\min}$. This would reduce the cost of providing the prizes without affecting players' incentives, leading to the same aggregate equilibrium effort and relaxing the budget constraint, which would allow to increase the prizes for the highest types and increase aggregate effort. Suppose $z_{\min} > 0$.[30] If G^{-1} were not right-continuous at $z_{\min}$,

[30] The proof of part 3 shows that $z_{\min} = 0$ if $(c^{-1})'(0) = \infty$.

then the first inequality in (17) and the discontinuous decrease of $(c^{-1})'$ at $z_{\min}$ would violate (16). Thus, G strictly increases on $[0, G^{-1}(z_{\max})]$. If $z_{\max} = 1$, then 2 holds, because $G^{-1}(1)$ is the highest prize awarded. If $z_{\max} < 1$, then $G^{-1}(z_{\max}) = G^{-1}(1) = m$, otherwise the second inequality in (17), the discontinuity of $l(z)$ at $z = z_{\max}$, and the strict monotonicity of $(c^{-1})'$ would imply that (16) cannot be satisfied for z smaller than but close enough to $z_{\max}$. For 3, suppose that $z_{\min} > 0$. Then (16) implies that G^{-1} is discontinuous at $z_{\min}$.[31] But then the left-hand side of first inequality in (17) would equal ∞, and the arguments that underlie the first inequality in (17) for finite $(c^{-1})(0)$ would imply that G^{-1} is not effort maximizing.

This result highlights two differences between the optimal prize distribution for convex costs and the one for concave prize valuations. First, convex costs lead to intermediate prizes being awarded even if the budget constraint does not bind, unlike with concave prize valuations. This is because with convex costs a slight increase in a type's prize induces strictly more effort when his prize, and therefore effort, is 0 than when his prize is m, but such an increases reduces higher types' effort by the same amount. Thus, it cannot be optimal for some type to obtain prize 0 and a slightly higher type to obtain prize m. Second, if the marginal cost at 0 is 0, then almost every type is awarded a positive prize. This is because a marginal cost of 0 implies that a slight increase of a type's prize from 0 leads to an effort increase that outweighs the decrease in higher types' efforts. In other words, it is optimal to have almost every type put in positive effort, unlike with linear costs, which lead to negative marginal virtual effort for low types.

Proposition 8 is related to Proposition 5 in 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 distribution.

In addition, while Propositions 7 and 8 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 number of players. The players may be ex-ante symmetric or asymmetric, and may or may not have private information.

[31] If $\lim_{z \downarrow z_{\min}} G^{-1}(z) = 0$, then $\lim_{z \downarrow z_{\min}} (c^{-1})'(l(z)) = \infty$, so for any $\delta > \lambda$, for z slightly higher than $z_{\min}$ we have $(c^{-1})'(l(z))(z_{\min}/2) - \int_z^{z+z_{\min}/2}(c^{-1})'(l(r))dr > 0$ and $(c^{-1})'(l(z))(z_{\min}/2) - \int_{z+z_{\min}/2}^{1}(c^{-1})'(l(r))dr > \delta$, contradicting (16).

The next result shows that Proposition 6 extends to the setting with a convex cost function c.

Proposition 9 *Suppose the budget constraint binds, and let $G^m_{\max}$ be an optimal prize distribution 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).*

Similarly to Proposition 6, the proof of Proposition 9 also shows that $z_{\min}$ and $z_{\max}$ (weakly) increase as m increases; and for any $m' < m''$, we have $G^{m''}_{\max}(y) \geq G^{m'}_{\max}(y)$ for $y < m'$ and $G^{m''}_{\max}(y) \leq G^{m'}_{\max}(y)$ for $y \geq m'$.

6.3 Example of the optimal prize distribution for players with convex effort cost

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 8 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 $q(z) = (c^{-1})'(l(z))$. Plug $q(z)$ into (16), and differentiate with respect to z to obtain the differential equation $q\prime(z)z + 2q(z) = 0$ for $q(z)$.[32] Solving this equation, and substituting back into (16), 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 $(G^{-1})'(z) = (-2\lambda/z^4)(((c^{-1})')^{-1})'(\lambda/z^2)$.[33]

Since $c^{-1}(z) = \sqrt{z}$, $(c^{-1})'(z) = 1/(2\sqrt{z})$, $\left((c^{-1})'\right)^{-1}(z) = 1/(4z^2)$, 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 $(c^{-1})'(0) = \infty$, $z_{\min} = 0$. We must therefore have $y_{\min} = 0$, otherwise the same equilibrium effort can be achieved with a lower budget by "shifting down" $G^{-1}(z)$ for $z > 0$ to $G^{-1}(z) - y_{\min}$.

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

[33] 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.

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 the budget C, the prizes awarded increase continuously from the lowest prize $y = 0$, but in contrast to the case of concave h studied in Section 5.4, there is no atom at the lowest prize, so 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.

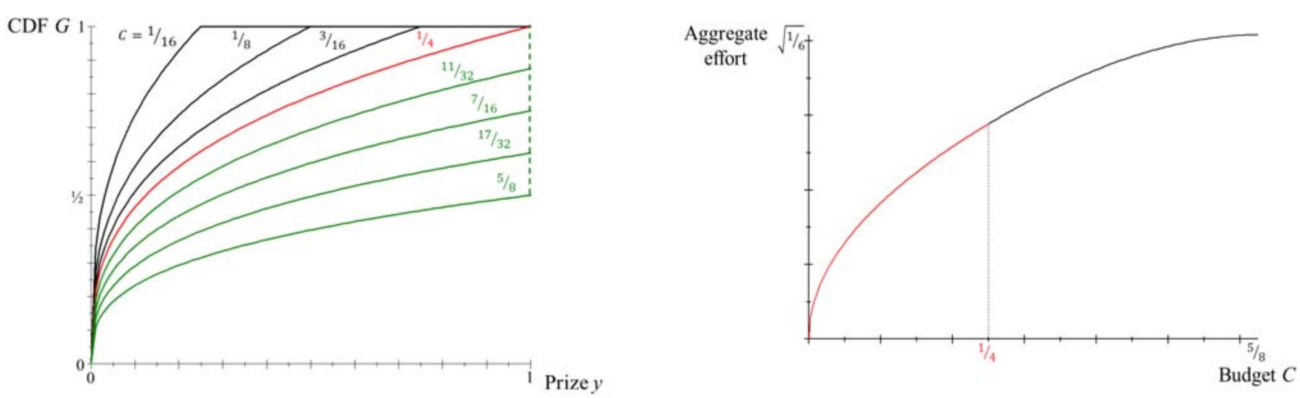

Figure 2: The optimal prize distribution as C increases from 0 to 5/8 (left), and the resulting aggregate effort (right)

The two solved examples, $h(x) = \sqrt{x}$ and $c(t) = t^2$, demonstrate the differences and the 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.

7 Conclusion

This paper introduced a new approach to analyzing contest design questions. The approach applies to contests with many contestants, and circumvents the difficult problem of solving for equilibria of contests with possibly asymmetric, privately informed players and heterogeneous prizes. We used the approach to investigate effort maximization 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 first showed that for any prize distribution a FOSD shift in the ability distribution increases aggregate effort. We then investigated the optimal prize distribution. Our key qualitative finding is that contestants' risk aversion and convex effort cost call for numerous prizes of different value. 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. The analysis enables deriving closed-form approximations of the effort-maximizing prize distributions for concrete utility functions and distributions of player types. This facilitates further analysis of large contests.

Our approach can be used to investigate many 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 distribution in such settings with many contestants is a small number of prizes of the highest possible value. This is true regardless of players' effort costs, prize valuations, private information, and type distributions.

8 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 $\displaystyle\int gdF^n\to_n\int gdF$ 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 2. 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 5.2. 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 $((G^n)^{-1})^{\infty}_{n=1}$ that satisfy the budget constraint and whose value of (8) 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 (8) 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, noticing that since k is increasing and continuous, the inequality $h'(m)\,k(z) \geq \lambda$ for $z > z_{\max}$ is equivalent to $h'(m)k(z_{\max}) \geq \lambda$.

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 (8) 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 (8) at least by $\left(\varepsilon/2^N\right)(I - \delta/3)$, and the latter change decreases (8) at most by $\left(\varepsilon/2^N\right)(S + \delta/3)$. This increases the value of (8) 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 (8) is higher for $(H^n)^{-1}$ than for $(\widetilde{H}^n)^{-1}$ because k is an increasing function.

Proof of Proposition 6. Assume first that $h'(0) < \infty$. Let $z_{\min}^m$, $z_{\max}^m$, and λ^m denote $z_{\min}$, $z_{\max}$, and λ for a given m. The proof of Proposition 5 shows that $z_{\min}^m < z_{\max}^m$ for all m. We claim that λ^m weakly increases with m. Suppose to the contrary that $\lambda^{m'} > \lambda^{m''}$ for some $m' < m''$.

Since $h'\left((G_{\max}^m)^{-1}(z)\right)k(z) = \lambda^m$ for all $z \in (z_{\min}^m, z_{\max}^m]$ and h' is decreasing, $h'(0)\,k(z) \geq \lambda^m$ for all $z \in (z_{\min}^m, z_{\max}^m]$, and since k is continuous, we have $h'(0)\,k(z_{\min}^m) \geq \lambda^m$. Since we also have $h'(0)\,k(z_{\min}^m) \leq \lambda^m$ (because we are in Case 1 of Section 5.2), 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): $(G^{m'}_{\max})^{-1}(z) = 0 \le (G^{m''}_{\max})^{-1}(z)$ for all $z \le z^{m'}_{\min}$, 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): $(G^{m'}_{\max})^{-1}(z) \le (G^{m''}_{\max})^{-1}(z)$ for all $z \in (z^{m'}_{\min}, \min\{z^{m'}_{\max}, z^{m''}_{\max}\}]$. If $z^{m'}_{\max} \ge z^{m''}_{\max}$, then we have (c): $(G^{m'}_{\max})^{-1}(z) \le (G^{m''}_{\max})^{-1}(z)$ for $z > \min\{z^{m'}_{\max}, z^{m''}_{\max}\}$, because $(G^{m'}_{\max})^{-1}(z) \le m'$ and $(G^{m''}_{\max})^{-1}(z) = m''$. If $z^{m'}_{\max} < z^{m''}_{\max}$, then (c) follows from (b). 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 $h'(0)\,k(z^m_{\min}) = \lambda^m$, 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 $h'(m)k(z^m_{\max}) \ge \lambda^m$ if $z_{\max} < 1$ would be violated for large enough m. In case (a), the monotonicity of $z^m_{\min}$ and λ^m implies that λ^m stabilizes at some λ for sufficiently large m,[34] and case (b) implies that $z^m_{\max}$ converges to 1 as m diverges to infinity, otherwise the budget constraint would be violated.[35] Therefore, in both these cases $G^m_{\max}$ converges to a distribution that may have an atom only at 0.

Suppose now that $h'(0) = \infty$. Then, $k(z^m_{\min}) = 0$ for all m; let $z^m_{\min} = z^{m'}_{\min} = z^{m''}_{\min}$. If $\lambda^{m'} > \lambda^{m''}$ for some $m' < m''$, then (a)-(c) hold, except that interval $(z^{m''}_{\min}, z^{m'}_{\min})$ on which we had strict inequality $(G^{m'}_{\max})^{-1}(z) < (G^{m''}_{\max})^{-1}(z)$ is now empty. However, we must have strict inequality $(G^{m'}_{\max})^{-1}(z) < (G^{m''}_{\max})^{-1}(z)$ for $z \in (z^m_{\min}, \min\{z^{m'}_{\max}, z^{m''}_{\max}\}]$, since otherwise $h'\left((G^m_{\max})^{-1}(z)\right) k(z) = \lambda^m$ could not be satisfied for both $m = m'$ and m''. The rest of the proof is the same as for $(c^{-1})'(0) < \infty$.

Proof for the conditions in Cases 1 and 2 from Section 6.1. 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 5.2. The argument is, however, more involved, because the objective function (14) 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

[34] If λ^m increases, then $z^m_{\min}$ increases, and $h'\left((G^m_{\max})^{-1}(z)\right) k(z) = \lambda^m$ implies that $(G^m_{\max})^{-1}(z)$ decreases for $z \in (z_{\min}, z_{\max})$. Thus, once $z^m_{\max}$ stabilizes at 1, λ^m can no longer increase, since (9) holds with equality.

[35] 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 length bounded away from 0, so $\int_0^1 G^{-1}(z)\,dz$ would diverge to infinity.

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.)

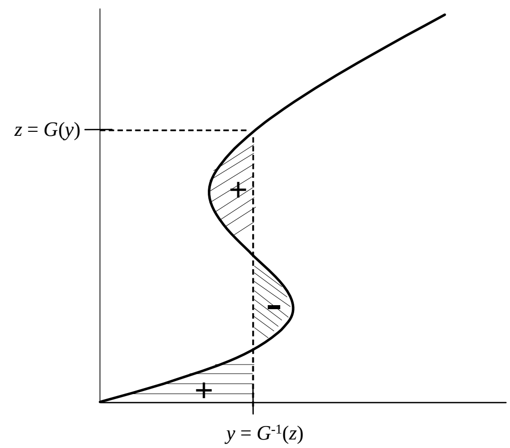

Figure 3: The definition of $l\,(z)$

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 (14) 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 (14) 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 (14) 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, (14) increases approximately by

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

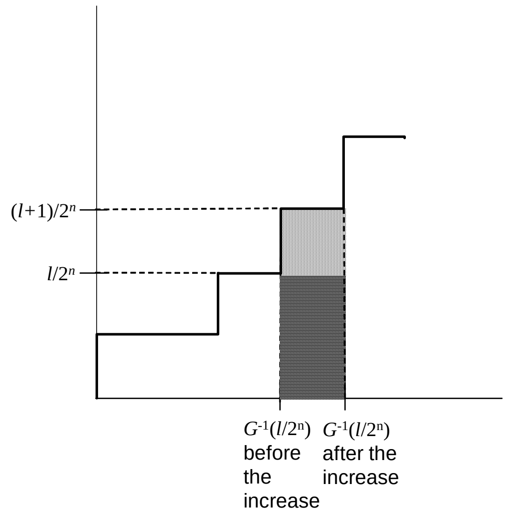

Figure 4: Increasing G^{-1}

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 5.2. 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 (14) 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 (14). 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 (14) on $(k/2^n, (l+1)/2^n]$. Finally, the value of $l(z)$ and the integrand of (14) on other intervals of the partition stay the same.

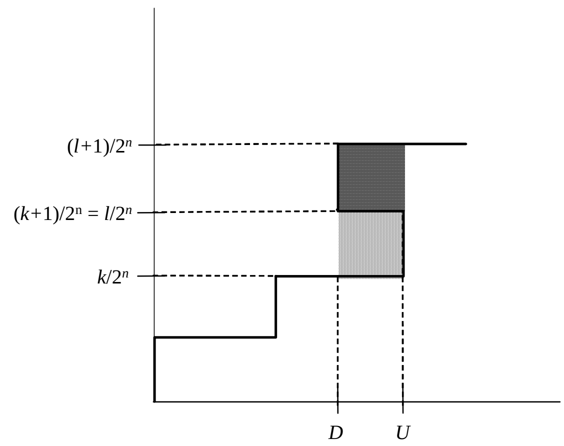

Figure 5: Making $(\widetilde{H}^n)^{-1}$ monotonic

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 (16) 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 (16) 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 $z_{\min} = z_{\max}$ implies $l\,(1) = \int_0^{G^{-1}(1)} G\,(r)\,dr = \int_0^m G\,(0)\,dr = G\,(0)\,m$.

Proof of Proposition 9. Assume first that $(c^{-1})'(0) < \infty$. Notice that by (16), the first part of **(17)**, and the assumption that $(c^{-1})'$ is decreasing, we have that

$$(c^{-1})'(0)z_{\min} - \int_{z_{\min}}^1 (c^{-1})'(l(r))dr = \lambda. \tag{19}$$

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 8, $z_{\min}^m < z_{\max}^m$ 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 (19), $z_{\min}^{m'} > z_{\min}^{m''}$. 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 (16) and the fact that the left-hand side of (16) 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): $(G_{\max}^{m'})^{-1}(z) \leq (G_{\max}^{m''})^{-1}(z)$ for $z > \min\{z_{\max}^{m'}, z_{\max}^{m''}\}$, because $(G_{\max}^{m'})^{-1}(z) \leq m'$ and $(G_{\max}^{m''})^{-1}(z) = m''$. If $z_{\max}^{m'} < z_{\max}^{m''} \leq 1$, then (c) follows from (b). 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.

By (19), 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 $(c^{-1})'(l(1))(2z^m_{\max} - 1) \geq \lambda^m$ if $z_{\max} < 1$ would be violated for large enough m. In case (a), the monotonicity of $z^m_{\min}$ and λ^m implies that λ^m stabilizes at some λ for sufficiently large m, and case (b) implies that $z^m_{\max}$ converges to 1 as m diverges to infinity, otherwise the budget constraint would be violated.[36] Therefore, in both these cases $G^m_{\max}$ converges to a distribution that may have an atom only at 0.

Suppose now that $(c^{-1})'(0) = \infty$. Then, $z^m_{\min} = 0$ for all m. If $\lambda^{m'} > \lambda^{m''}$ for some $m' < m''$, then (a)-(c) hold, except that interval $(z^{m''}_{\min}, z^{m'}_{\min})$ on which we had strict inequality $(G^{m'}_{\max})^{-1}(z) < (G^{m''}_{\max})^{-1}(z)$ is now empty. However, we must have strict inequality $(G^{m'}_{\max})^{-1}(z) < (G^{m''}_{\max})^{-1}(z)$ on some interval $\oslash \neq (z\prime, z'') \subset (0,1]$, since if $(G^{m'}_{\max})^{-1}(z)$ were equal to $(G^{m''}_{\max})^{-1}(z)$ for all z, (16) could not be satisfied for both m' and m''. The rest of the proof is the same as for $(c^{-1})'(0) < \infty$.

Proof of the claim from Section 7. 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

[36] The arguments are analogous to those used in the proof of Proposition 6.

$\int_{x^*}^1 c^{-1}(xh(y^A(x)) - \int_{x^*}^x h(y^A(z))dz)f(x)\,dx$. 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 $h'\left(y^A(x)\right)(xf(x)L - \varepsilon H)$.

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] **Bulow, Jeremy I., and John Roberts.** 1989. "The Simple Economics of Optimal Auctions." *Journal of Political Economy*, 97(5): 1060-1090.

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

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

[8] **Hillman, Arye L., and John G. Riley.** 1989. "Politically Contestable Rents and Transfers." *Economics and Politics*, 1(1): 17–39.

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

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

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

[12] **Olszewski, Wojciech, and Ron Siegel.** 2016. "Large Contests." *Econometrica*, 84(2), 835-854.

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

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

[15] **Xiao, Jun.** 2016. "Asymmetric All-Pay Contests with Heterogeneous Prizes." *Journal of Economic Theory*, 163: 178-221.