Contest Architecture

Benny Moldovanu and Aner Sela*

October 28, 2004

Abstract

A contest architecture specifies how the contestants are split among several subcontests whose winners compete against each other (while other players are eliminated).

We compare the performance of such dynamic schemes to that of static winner-take-all
contests from the point of view of a designer who maximizes either the expected total
effort or the expected highest effort. For the case of a linear cost of effort, our main
results are: 1) If the designer maximizes expected total effort, the optimal architecture
is a single grand static contest. 2) If the designer maximizes the expected highest effort,
and if there are sufficiently many competitors, it is optimal to split the competitors
in two divisions, and to have a final among the two divisional winners. Finally, if the
effort cost functions are convex, the designer may benefit by splitting the contestants
into several sub-contests, or by awarding prizes to all finalists.

1 Introduction

Contests are situations in which agents spend resources in order to win one or more prizes. Independently of success, all contestants bear the cost of their "bids". Numerous applications

^{*}We wish to thank an associate editor and two anonymous referees for very helpful comments. We are also grateful to Dan Kovenock who discussed this paper's presentation at ESEM 2002, and to Lones Smith who provided us a very valuable reference. Moldovanu: Department of Economics, University of Bonn, mold@uni-bonn.de; Sela: Department of Economics, Ben-Gurion University, anersela@bgumail.bgu.ac.il

of winner-take-all grand contests have been made to rent-seeking and lobbying in organizations, R&D races, political contests, promotions in labor markets, trade wars, military and biological wars of attrition. Due to the pervasive nature of such competitions (which are either designed or arise naturally), there exist large scientific and popular literatures¹ on the subject. Most of the scientific attention has focused on contests where a unique prize is awarded, and where all contestants compete against each other in a simultaneous grand contest.

A casual survey of designed real-life contests reveals that often contestants do not compete in an "each against all" fashion, but rather that they are first divided into several subcontests whose winners compete against each other at later stages (while other players are eliminated). Such designs are very popular in the world of sport but they are also often observed in instances such as: 1) The organization of internal labor markets in large firms and public agencies: the sub-contests are usually regional or divisional, and the prizes are promotions to well-defined (and usually equally-paid²) positions on the next rung of the hierarchy-ladder; 2) Political competition (e.g., for the US presidency) where candidates first spend resources to secure their party's nomination, and later, if they are nominated, spend more resources to get elected. 3) Science contests among university or high-school students, e.g., the Mathematics Olympiad.

A main question that we want to address here is whether there are intrinsic advantages for designs with multiple elimination rounds, or whether such architectures result solely from external constraints (e.g., in many sports matches must be bilateral, and the number of matches needs to be restricted due to time limits).

In the present paper we first study the properties of contests where the designer splits the contestants into several parallel sub-contests, each with an equal number of contestants. Then we study two-stage contests where the winners of the sub-contests at the first stage compete among themselves in a second stage. Our analysis relies on several tools (borrowed

¹An entertaining popular book is Frank and Cook (1995).

²Equal benefits for similar positions in the hierarchy (independent of the occupants' abilities) can be also rationalized by the designer's wish to minimize influence costs. Inderst et al. (2001) show that a hierarchical contest consisting of several sub-contests may be beneficial for such a designer.

from mathematical statistics) that can yield important insights for general studies of complex contests.

In our basic static model n contestants exert effort in order to win one of p prizes. Each contestant i exerts an observable effort. The contestant with the highest effort wins the first prize, the contestant with the second-highest effort wins the second prize, and so on until all the prizes are allocated. All contestants (including those that did not win any prize) incur a cost that is a strictly increasing function of their effort. The contestants have private information³ about a parameter ("ability") that affects their effort cost function. Cost functions are assumed to be strictly increasing in effort. The function governing the distribution of abilities in the population is common knowledge, and abilities are drawn independently of each other.

We focus on two goals⁴ for the contest designer and some relations among them: 1) Maximization of total expected effort⁵; 2) Maximization of expected highest effort.

The static model⁶ for the grand contest is isomorphic to a "private values" all-pay auction with several prizes⁷. Hence, for the case of linear cost functions and total effort maximization

⁴Since the equilibrium effort functions only dependent on the contest architecture, the methods of this paper can be used to analyze other goals arising in various applications. For example, in rent seeking contests one may wish to minimize effort.

⁵In Moldovanu and Sela (2001) we focused on splits in the prize sum in a one-stage grand contest and we calculated the optimal ratio of prizes for a designer that maximizes the expected total effort. Except for the case where contestants have an increasing marginal cost of effort, the value of the second and lower prizes should be zero. Related questions have been addressed in various models by Glazer and Hassin (1988), Barut and Kovenock (1998) and Krishna and Morgan (1998).

⁶A different model emphasizes the use of contests in order to extract effort under "moral hazard" conditions (see Lazear and Rosen (1981), Green and Stokey (1983), Nalebuff and Stiglitz (1983), and Rosen (1986)) In that literature agents usually have the same known ability, but output is a stochastic function of the unobservable effort. The identity of the most productive agent is determined by an external shock

⁷All-pay auction models with linear cost functions and incomplete information about the prize's value include Weber (1985), Hillman and Riley (1989), Krishna and Morgan (1997). Contest models with several

³Contest models with complete information about the value of a unique prize include, among others: Tullock (1980), Varian (1980), Moulin (1986), Dasgupta (1986), Hillman and Samet (1987), Dixit (1987), Baye et. al. (1993). Baye et. al. (1996) offers a complete characterization of equilibrium behavior in the complete information all-pay auction with one prize.

(which parallels the goal of revenue maximization in standard auctions⁸), we can apply some insights derived from the Payoff Equivalence Theorem. But there are no analogous payoff equivalence results for highest effort maximization or for non-linear cost functions.

In the dynamic, two-stage model⁹, the winners of the sub-contests organized at the first stage compete against each other in a final at the second stage. In order to avoid signalling effects due to possibly sophisticated information manipulation across stages, we assume here that the contestants at stage two do take into account the fact that their opponents won at a previous stage, but they do not observe their opponents' past actions.

Our main results can be summarized as follows:

- 1. If the designer wishes to maximize the expected total effort, and if the cost function is linear, the optimal architecture is a grand static contest for a unique large prize. A contest with several stages is equivalent to the grand static contest if a unique prize is awarded to a final winner, and the designer cannot increase her payoff by awarding prizes to finalists other than the ultimate winner.
- 2. If the designer wishes to maximize the expected value of the highest effort, and if the cost function is linear, splitting the contestants in several static sub-contests is not beneficial. But, for a sufficiently high number of contestants, two-stage contests with elimination dominate the static grand contest. In particular, the best performance is achieved by splitting the original competitors in exactly two divisions, which leads to a final among the two divisional winners.

identical prizes include Clark and Riis (1998), who compare simultaneous versus sequential designs under complete information, and Bulow and Klemperer (1999) who study a war of attrition.

⁸Analogously to reserve price in auctions, the introduction of minimum effort requirements is clearly beneficial (see Moldovanu and Sela, 2001). But note that for the calculation of an optimal minimum effort requirement the designer needs precise information (about the distribution of abilities in the population.

⁹Amegashie (1999) determines the optimal number of finalists for a designer that minimizes expenditures in a complete information, two-stage contest à la Tullock with homogenous contestants. The comparison between one-stage and two-stage contests is ambiguous there and depends on the parameters of the Tullock success functions at each stage.

3. If the effort cost function is convex, a designer that maximizes total effort can benefit by splitting the contestants in several sub-contests¹⁰. In particular, if the grand contest is dominated by a split one for a designer who maximizes the highest effort, then it is also dominated for a designer who maximizes the total effort. Finally, for the case of convex cost functions, we show that the designer may benefit from awarding prizes to finalists (besides the winner) in the first stage of a two-stage contest.

The rest of the paper is organized as follows: In Section 2 we present the one-stage and two-stage contest models. In Section 3 we derive symmetric equilibrium effort functions and we show that they are determined by differences among densities of successive order statistics. Section 4 contains all results for linear cost functions. We also display there novel single-crossing properties that arise when we vary the number of prizes and competitors. Section 5 deals with the case of convex costs. Section 6 concludes. In Appendix A we set up the analytical tools that are necessary in order to study stochastic dominance relations among functions of order statistics. Appendix B contains several proofs that would otherwise interrupt the flow of the argument.

2 The Model

2.1 One-Stage Contests

We consider a contest with a fixed total prize sum equal to 1. The designer determines the number and value of prizes, and these variables are made common knowledge. In case that p prizes are awarded, the value of the j-th prize is denoted by V_j , where $V_1 \geq V_2 \geq ... \geq V_p \geq 0$, and where $\sum_{i=1}^p V_i = 1$.

The set of contestants is $N = \{1, 2, ..., n\}$, where $n \ge 2$ and n > p. Each player i makes an effort x_i . These efforts are submitted simultaneously. An effort x_i causes a cost denoted

¹⁰If the grand contest is not optimal for a given cost function, then it continues to be dominated by split contests also for all cost functions that are more convex (i.e., for all functions with a higher Arrow-Pratt curvature index). The benefit of splitting increases in the degree of convexity.

by $c_i \gamma(x_i)$, where $\gamma: R_+ \to R_+$ is a strictly increasing function with $\gamma(0) = 0$, and where $c_i > 0$ is an ability parameter.¹¹ Note that a **low** c_i means that i has a **high** ability and vice-versa. We denote by g the inverse function γ^{-1} .

The ability (or type) of contestant i is private information to i. Abilities are drawn independently of each other from an interval [m, 1] according to a distribution function F which is common knowledge. We assume that F has a continuous density dF > 0. In order to avoid infinite bids caused by zero costs, we assume that m, the type with the highest possible ability, is strictly positive.¹²

The contestant with the highest effort wins the first prize V_1 . The contestant with the second highest effort wins the second prize V_2 , and so on until all the prizes are allocated.¹³ That is, the payoff of contestant i who has ability c_i and exerts an effort x_i is either $V_j - c_i \gamma(x_i)$ if i wins prize j, or $-c_i \gamma(x_i)$ if i does not win a prize. In the case of p equal prizes, the contestants with the p highest efforts win the available prizes.

Denote by $C_1, C_2, ... C_n$ the identical, independently distributed random variables governing the distribution of the contestants' abilities. Denote by $C_{(1,n)}, C_{(2,n)}, ..., C_{(n,n)}$ the corresponding order statistics, and by $F_{(1,n)}, F_{(2,n)}, ..., F_{(n,n)}$ their respective distribution functions. In Appendix A we list the explicit formulas for the distributions and densities of order statistics.

Each contestant i chooses his effort in order to maximize expected utility (given the other competitors' actions and the values of the prizes). The contest designer can organize one grand contest or she can split the contestants into t parallel sub-contests, each with $\frac{n}{t}$ participants.

We consider here two forms of utility for the designer: 1) The designer maximizes the expected value of total effort $E(\sum_{i=1}^{n} x_i)$ and 2) The designer maximizes the expected value of the highest effort $E(x_{\text{max}})$.

¹¹The treatment of the case in which i's cost function is $\delta(c_i)\gamma(x_i)$, where δ is strictly monotonically increasing, is completely analogous. The main assumption here is the separability of ability and effort.

¹²The case where m=0 can be treated as well, but requires slightly different methods.

The choice of the interval [m, 1] is a normalization.

¹³If h > 1 agents tie for a prize, each one of them gets the respective prize with probability $\frac{1}{h}$.

2.2 Two-Stage Contests

We now extend the previous model to a dynamic setting where the contest takes the form of familiar elimination tournaments: In the first stage, the n contestants are split into $t \geq 2$ parallel and symmetric sub-contests (or divisions), each with $k = \frac{n}{t}$ contestants. The designer determines the number of prizes $r < \frac{n}{t}$, and the prize sum α , $0 \leq \alpha \leq \frac{1}{t}$ in each division. In the second stage, the t first-stage winners (i.e., the players who exerted the highest effort in each division) compete against each other for s < t prizes that add up to the remaining prize sum of $1 - t\alpha$. All other players are eliminated after the first stage.

In this paper we focus on the case where only one prize is awarded in the final (second stage of the elimination tournament) and in each division at the first stage. By the results in Moldovanu and Sela (2001), this is optimal for a contest designer having the goals studied here that faces contestants with linear cost functions.

Each contestant i chooses his effort(s) in order to maximize expected utility. We assume that a contestant with type c that exerts efforts x^1 and x^2 at stages 1 and 2, respectively, bears a total cost of $c(\gamma(x^1) + \gamma(x^2))$. This separability assumption fits situations where the effect of the first-stage effort is not lasting¹⁴.

First-stage contestants know that they may need to exert extra effort at the second stage in order to win the entire tournament. The remaining contestants at stage two know that their opponents won at the previous stage, but we assume that they do not observe their opponents' past actions. Thus, if F denotes the distribution of contestants' abilities in the first stage, and if we assume that a separating equilibrium is played at the first stage (this will be the case below) each player at the second stage perceives the abilities of his opponents as being drawn from the distribution function $G = F_{(1,k)}$. That is, G is a contestant's updated belief about his t-1 opponents, given that each one of them won already a contest against k rivals.

The alternative assumption of a total cost of $c\gamma(x^1 + x^2)$ acts as a "budget constraint" and creates additional technical difficulties. Obviously, there is a difference between the two formulations only for non-linear cost-functions.

3 Equilibrium Characterization

Proposition 1 Consider a one-stage contest with n contestants where the designer awards p < n prizes, $V_1 \ge V_2 \ge ... \ge V_p \ge 0$. In a symmetric equilibrium, each contestant makes an effort according to the strictly decreasing function¹⁵

$$b(c) = g\left[\sum_{i=1}^{p} V_i \int_{c}^{1} \frac{1}{s} (dF_{(i,n-1)}(s) - dF_{(i-1,n-1)}(s))\right]$$
 (1)

Proof: In Moldovanu and Sela (2001) we used the first-order maximization condition in order to obtain a differential equation involving the equilibrium effort function and its derivative. That condition involves the different probabilities with which an agent expects to win each of the p prizes. We proved that the symmetric equilibrium effort function is given by

$$b(c) = g\left[\sum_{i=1}^{p} V_i \int_{c}^{1} -\frac{1}{s} dF_i^n(s)\right]$$
 (2)

where $F_i^n(s), 1 \leq i \leq n$, denotes the probability that an agent with type s meets n-1 competitors such that i-1 of them have lower types and n-i have higher types. The representation in the statement of the Proposition follows by relations 1,2 of Lemma 3 in Appendix A. \Box

The above new representation is very useful since it allows us to directly employ various stochastic dominance results among order statistics and functions thereof. For the case of equal prizes, which is the focus of the this paper¹⁶, we obtain an even more compact characterization:

Corollary 1 Consider a one-stage contest with n contestants where the designer awards p < n equal prizes, each worth $\frac{1}{p}$. The symmetric equilibrium effort function is given by

$$b_{n,p}(c) = g\left[\frac{1}{p} \int_{c}^{1} \frac{1}{s} dF_{(p,n-1)}(s)\right]$$
(3)

¹⁵We use the convention $dF_{(0,n-1)}(s) \equiv 0$.

¹⁶In each basic (sub)contest there will be equal prizes. But note that a two-stage elimination tournament where a portion of the prize sum is divided among the first-stage winners is, in fact, a contest with several, unequal prizes.

Proof: The result follows by the telescopic nature of the equilibrium effort function in Proposition 1. ⊡

The equilibrium in the two-stage contest where a prize of α is awarded to the winner of each division, while the additional prize $1 - t\alpha$ is awarded to the final winner is given by:

Corollary 2 The symmetric sub-game-perfect equilibrium in a two-stage tournament is as follows: In the second stage the effort function is given by

$$b_2(c) = g \left[\int_c^1 (1 - t\alpha) \frac{1}{s} dG_{(1,t-1)}(s) \right]$$
 (4)

In the first stage the effort function is given by

$$b_1(c) = g \left[\int_c^1 \left[\alpha + (1 - G(c))^{t-1} (1 - t\alpha) - c\gamma(b_2(c)) \right] \frac{1}{s} dF_{(1,k-1)}(s) \right]$$
 (5)

Proof. The claim for the second stage is clear by Proposition 1. A contestant's value of winning in the first stage is given by the sum of the prize in the first stage, α , and of the expected payoff in the second stage, which is given by:

$$(1 - G(c))^{t-1}(1 - t\alpha) - c\gamma((b_2(c)))$$

$$= (1 - t\alpha)[(1 - G(c))^{t-1} - c\int_c^1 \frac{1}{s} dG_{(1,t-1)}(s)]$$

The important thing to notice is that the expected payoff in the second stage contest is higher for higher ability (or lower c). Thus, in the first stage, we have a contest where the value of the prize is higher for higher ability types, and the equilibrium bid is derived analogously to the derivation in Proposition 1 \Box

4 Linear Cost Functions

It is clear from the above results that the equilibrium for a strictly increasing cost function γ is obtained by applying the inverse function $\gamma^{-1} = g$ to the equilibrium obtained for the linear cost function $\gamma(x) = x$. The equilibrium properties in the linear cost case are therefore central to our analysis, and we next focus on this case.

4.1 One-Stage Contests

Our first structural results make explicit the trade-offs induced by varying the number of contestants and the number of prizes: For a fixed number of contestants, increasing the number of prizes has a negative effect on the equilibrium effort of high ability contestants and a positive effect on the equilibrium effort of low ability contestants¹⁷. In contrast, for a fixed number of prizes, increasing the number of contestants has exactly opposite effects: a positive one on the equilibrium effort of high ability contestants and a negative one on the equilibrium effort of low ability. In both cases there exists exactly one type whose equilibrium effort is unaffected by the change. Besides the intrinsic interest, these newly identified single-crossing¹⁸ properties are used in the analysis of the convex cost case.

Lemma 1 Consider a contest with n contestants. For any number of prizes p, r such that n > r > p, the equilibrium effort functions $b_{n,r}(c)$ and $b_{n,p}(c)$ are single-crossing. That is, there exists a unique $c^* = c^*(n, r, p) \in (m, 1)$ such that:

- 1. $b_{n,r}(c^*) = b_{n,p}(c^*);$
- 2. $b_{n,p}(c) > b_{n,r}(c)$ for all $c \in [m, c^*)$;
- 3. $b_{n,p}(c) < b_{n,r}(c) \text{ for all } c \in (c^*, 1).$

Proof. See Appendix B \Box

The intuition for the proof is as follows: When the number of prizes increases (while keeping the prize sum fixed) the agent with the highest ability exerts less effort since the value of her prize (which she gets for sure) decreases, while agents with very low ability exert

¹⁷Another operation that induces a similar change is the imposition of a bid cap. The effects of this operation in an all-pay auction with a unique prize whose value is private information are studied by Gavious et al. (2003).

¹⁸The reader may be familiar with single-crossing as an assumed property on utility functions (see for example Athey, 2000, who uses it in order to establish monotone comparative statics in the theory of one-person decision making under risk). In contrast, single-crossing is here an endogenously arising property of equilibrium effort functions.

higher efforts since now there are additional prizes to be won. Thus, the two bidding functions $b_{n,p}$ and $b_{n,r}$ must cross at least once in the interval $(m,1)^{19}$. Uniqueness is established by noting that the derivatives $b'_{n,p}$ and $b'_{n,r}$ are equal for a unique type in the interior of the interval (m,1).

Lemma 2 Consider a contest with p prizes. For any numbers of contestants n, k such that n > k > p, the equilibrium effort functions $b_{n,p}(c)$ and $b_{k,p}(c)$ are single-crossing. That is, there exists a unique $c^* = c^*(n, k, p) \in (m, 1)$ such that:

- 1. $b_{n,p}(c^*) = b_{k,p}(c^*);$
- 2. $b_{n,p}(c) > b_{k,p}(c)$ for all $c \in [m, c^*)$;
- 3. $b_{n,p}(c) < b_{k,p}(c)$ for all $c \in (c^*, 1)$.

Proof: The proof uses exactly the same steps and arguments analogous to those presented in the proof of Lemma 1. \Box

Lemmas 1 and 2 pointed out that high and low ability contestants are affected by changes in the number of contestants and prizes in exactly opposite ways. Naturally, the interesting question is: What is the aggregate effect of variations in the number of prizes and the number contestants on the designer's payoff? This is answered by the following result:

Proposition 2 Assume that the designer's payoff is the expected value of total effort. Then the following hold:

- 1. The designer's payoff increases in the number of contestants;
- 2. The designer's payoff decreases in the number of prizes.

Proof: See Appendix B. \Box

¹⁹Note that they are also equal at c = 1.

The proof first shows that the designer's payoff is given by the expectation of functions of order statistics (of the random variables governing abilities). The monotonicity results follow easily by stochastic dominance relations among order statistics.

Consider now the following two types of static contest architectures for a given group of n contestants:

- 1. In the *Grand Architecture* (GA) the entire group of n contestants competes for one prize worth 1.
- 2. In the t-Parallel-Architecture (t-PA) there are t > 1 separate divisions, each consisting of $\frac{n}{t}$ contestants, competing for one prize worth $\frac{1}{t}$.

The next result shows that splitting the contestants into several sub-contests is not beneficial for a designer who maximizes the total effort²⁰.

Theorem 1 Assume that the designer's payoff is given by the expected value of total effort. Then the designer's payoff in the Grand Architecture (GA) is larger than the respective payoff in any Parallel (t-PA) Architecture. This result extends to the case of concave cost functions.

Proof: In the linear-cost case, the designer's payoff is given by $R_{n,p} = n \int_m^1 b_{n,p}(c) dF(c)$. In the GA the designer's payoffs is given by $R_{n,1}$ and by $t \cdot \frac{1}{t} R_{\frac{n}{t},1} = R_{\frac{n}{t},1}$ in the t-PA (see the proof of Proposition 2). The result follows by repeated applications of Proposition 2. The proof for concave-cost case appears in Appendix B. \Box

The above result is closely related to insights about the benefit of bundling in standard auctions: consider for example a seller that uses second-price auctions in order to sell to n buyers a divisible object either as a single unit, or in t separate auctions, each for an equal portion of the object worth $\frac{1}{t}$ of the value of the entire unit. In the bundled auction, revenue equals the expectation of the second highest value out of n. If, in the second alternative, each buyer can compete in at most one auction, and there are $\frac{n}{t}$ buyers competing for each

²⁰From the above result (or from our previous work) it also follows that having multiple prizes in each contest is here detrimental for the designer.

portion, the total revenue to the seller is t times the expected second highest value for a portion of size $\frac{1}{t}$ out of $\frac{n}{t}$. This equals the expectation of the second highest value out of $\frac{n}{t}$.

We now consider a principal whose payoff is given by the expected value of the highest effort. The comparative statics with respect to changes in the number of prizes is similar to that obtained for the maximization of total effort: the designer's payoff decreases in the number of prizes. But, the dependence on the number of contestants is more subtle now. By the formula obtained in the proof of Proposition 2, if the number of contestants n tends to infinity, the payoff of a principal who maximizes the expected value of the total effort monotonically increases (and converges to $\frac{1}{m}$). Using Lemma 2, there are four effects at play when we increase the number of contestants. On the positive side: 1) The effort of high ability contestants goes up; 2) We add up the effort of more contestants. On the negative side: 3) The effort of low ability contestants goes down; 4) The measure of types whose effort goes down increases²¹. Roughly speaking, we showed that effects 1 and 2 are together stronger than effects 3 and 4.

In the present situation, we completely lose effect 2, but effect 1 has more weight since we are only interested in the highest effort, which naturally comes from high ability contestants.

Proposition 3 Assume that the designer's payoff is given by the expected value of the highest effort. As the number of contestants tends to infinity, the designer's payoff in the Grand Architecture converges to $\frac{1}{2m}$. But, the convergence need not be monotonic and the designer's payoff may be maximized for a number of contestants $n^* < \infty$.

Proof: For the first part, see Appendix B and also the intuition following Theorem 4 in the next subsection. For the second part, see example below. \Box

Example 1 The expected value of the highest effort in the Grand Architecture is given by $P_{n,1} = \int_m^1 b_{n,1}(c) dF_{(1,n)}(c)$. Assume that abilities are drawn from the interval [0.5, 1] according to the distribution function $F(s) = (2s-1)^{0.1}$. If there are two contestants, we obtain that the designer's payoff is $P_{2,1} = 1.2116$. Since $P_{2,1} > \lim_{n \to \infty} P_{n,1} = 1$, the designer's payoff

 $^{^{21}}$ This last effect was not formally proven, but it easily follows from the proof of Lemma 2.

cannot be monotonically increasing. A simple numerical calculation reveals that n=2 is in fact the optimal number of contestants. For the uniform distribution F(s)=2s-1, the designer's payoff is monotonically increasing, hence $n^*=\infty$.

The above example demonstrates the advantages of restricting entry²² if the designer maximizes the expected highest performance²³.

Recall that a designer who maximizes the expected value of total effort preferred the Grand Architecture to any Parallel Architecture (Theorem 1). The result was an immediate consequence of the monotonicity of the designer's payoff in the number of contestants (see Proposition 2-1). Whenever the designer's payoff is not monotonically increasing in the number of contestants (as it may happen here), a parallel design may, in principle, be advantageous. But, in a parallel design, the prize awarded in each sub-contest is only a fraction of the total prize, thus causing a decrease in the effort of high ability contestants. Our next result uses the single-crossing property exhibited in Lemma 2 to show that the second effect always dominates:

Theorem 2 Assume that the designer's payoff is given by the expected value of the highest effort. Then her payoff in the Grand Architecture is larger than her payoff in any Parallel Architecture.

Proof: See Appendix B.

²²We displayed an example where the optimal number of contestants is two. Interestingly, many competitions for US defense procurement involved only two firms: the F-15 and F-16 engine competition (General Electric versus Pratt and Whitney), the Sparrow air-to-air missile competition (General Dynamics versus Raytheon), the SSN-688 submarine competition (Electric Boat versus Newport News).

²³Taylor (1995), and Fullerton and McAfee (1999) assume positive fixed costs of entry in a tournament, and also show that restricting entry is optimal. Che and Gale (2003) study a complete-information model where firms invest in the quality of an innovation and offer it to a buyer for a certain price (thus "prizes" are endogenous in their model). This has the flavor of an all-pay auction since all investments are sunk. Che and Gale show that a first-price auction among two firms is the optimal contest design in this framework.

4.2 Two-stage Contests

Consider first a two-stage tournament where the whole prize is awarded in the final, i.e., where $\alpha = 0$. Then, no matter what types contestants have, the contestant with the highest ability wins her respective sub-contest, and then goes on to win the entire tournament cashing the whole prize. But this outcome is the same as the outcome in the one-stage, Grand Architecture contest. By the Payoff Equivalence Theorem, the payoff of a designer that wishes to maximize total effort must be the same in both situations. This insight generalizes of course to dynamic tournaments with more than two-stages.

Our next result shows that the designer cannot increase her payoff by awarding prizes to finalists other than the ultimate winner. The intuition is as follows: By Proposition 2 we know that a one stage contest with t prizes where the prizes go to the t most able competitors is dominated by a one-stage contest with a unique prize, and thus, by Payoff Equivalence, by an equivalent two-stage contest with a unique prize for the winner. In a two-stage contest where prizes are awarded to the divisional winners, it is not necessarily the case that the prizes go to the (overall) t most able contestants²⁴. Since efficiency is lost, such a contest yields even less total effort than a contest where the t prizes indeed go to the t most able contestants.

Proposition 4 Consider a two-stage tournament where the designer's payoff is given by the expected value of total effort. Then, it is optimal for the contest designer to award the entire prize sum in the second stage. In other words, $\alpha = 0$ is optimal.

Proof. See Appendix B. \Box

The above discussion and Proposition 4 yield:

Theorem 3 Assume that designer's payoff is given by the expected value of total effort.

²⁴For example, it can happen that the two contestants having the highest ability compete against each other in the same sub-contest at stage one. Thus, while the most able competitor surely wins the tournament, the second most able may not qualify for the final, while a lesser competitor may do so if he won another sub-contests.

Then, the designer's payoff in the one-stage Grand Architecture contest is larger or equal than the respective payoff in any two-stage contest.

The previous theorem shows that, in some cases, the justification for elimination tournaments may be of external nature. But, at least, elimination tournaments with a unique prize do not perform worse than a grand static contest. In marked contrast to these results, we next find that, from the point of view of a designer that maximizes total effort, the one-stage Grand Architecture contest is dominated by two-stage contests if there are enough contestants:

Theorem 4 Assume that the designer's payoff is given by the expected value of the highest effort. Then, for a sufficiently high number of competitors, the designer's payoff in any two-stage tournament where the whole prize is awarded to the final winner is larger than the designer's payoff in the one-stage Grand Architecture. In particular, the highest payoff is obtained in a contest with two divisions in the first stage.

Proof: See Appendix B.⊡

The proof of the above result is concise, but not very intuitive. Here is a more intuitive argument, with some interest of its own: In a two-stage tournament with t > 2 divisions at the first stage and t finalists at the second stage, each finalist knows that the other contestants already won against $\frac{n}{t} - 1$ competitors. Thus, the updated belief on contestants' abilities follows the distribution $1 - (1 - F(s))^{\frac{n}{t}}$. When n gets large, this distribution puts more and more mass on very high abilities, and it converges to a Dirac distribution concentrated on the highest ability. Thus, for large n, the second stage contest approximates a **complete information**²⁵ contest among t contestants with known ability m. In the symmetric equilibrium of such a complete information contest²⁶, all t contestants randomize their effort on

²⁵Similar "purification" arguments (a la Harsanyi) for an all-pay auction with two contestants have been analyzed by Amman and Leininger (1996). It is worth noting that the incomplete information structure that approaches complete information naturally arises in our two-stage contest with many contestants, whereas it is arbitrarily constructed in Amman and Leininger's paper (and generally in Harsanyi's approach).

 $^{^{26}}$ See Baye et. al (1996) . They also derive asymmetric equilibria.

the interval $[0,\frac{1}{m}]$ according to the distribution function $H(x)=(mx)^{\frac{1}{t-1}}$. Hence, the distribution of the highest effort in the complete information contest is the distribution of the highest order statistic, $H_{(t,t)}(x)=(mx)^{\frac{t}{t-1}}$. It is easily calculated that $E[H_{(t,t)}(x)]=\frac{1}{m}\frac{t}{2t-1}$. This function is decreasing in t, and thus it attains a maximum at $t=2^{27}$. Finally, note that if the number of divisions t grows out of bounds²⁸, the expected value of the highest effort converges to $\frac{1}{2m}$, confirming the earlier result obtained in Proposition 3.

5 Convex Cost Functions

In the previous sections we have looked at several scenarios that have opposite effects on high ability and low ability types, respectively, and we have studied the ensuing trade-offs. If the cost functions are convex, instruments that extract additional efforts from high-ability types are, roughly speaking, less potent, since these types already exert large efforts and face high marginal costs. Thus, with convex effort costs, the effect of an instrument that decreases the effort of high ability types while increasing the effort of low ability types gets amplified (and more so if the cost function is more convex), and total effort may increase when such an instrument is used, even if it failed to do so when effort costs were linear. In this section we briefly illustrate this phenomenon for two instruments²⁹: splits of contestants in several divisions in a static contest, and the award of interim prizes to all finalists in a contest with two rounds.

Our first example shows that, for convex cost functions, splitting the contestants into several sub-contests may be beneficial for a designer maximizing total effort ³⁰.

²⁷Recall here the optimality of precisely two contestants obtained in Example 1 for abilities that are drawn from the interval [0.5, 1] according to the distribution function $F(s) = (2s-1)^{0.1}$. Note that this distribution is also such that most of the mass is concentrated near the type with highest ability. Thus, also in that example, the complete information case yields the right insight.

²⁸In order to use the previous steps, we need to also insure that $\frac{n}{t}$ gets larger than any bound. For example, consider $n = a^2$, $a \in N$, and let $t = \sqrt{n}$.

 $^{^{29}}$ See also Moldovanu and Sela (2001) who focus on splits in the prize sum.

³⁰If the cost function is convex enough, it is possible that a joint split of contestants and prizes in each sub-contest is even more beneficial than the simple splits. See Moldovanu and Sela (2001) for the effect of

Example 2 Let n=6, and let abilities be uniformly distributed on the interval $\left[\frac{1}{2},1\right]$. Consider the convex cost function $\gamma(x)=x^2$. Total effort in the Grand-Architecture is

$$R_{6,1} = 6 \int_{0.5}^{1} 2\left(\int_{c}^{1} (6-1) \frac{1}{s} (1 - (2s-1))^{6-2} 2\right)^{0.5} = 2.1014$$
 (6)

Total effort the 2-Parallel-Architecture is

$$2 \cdot R_{3,1} = 2 \cdot 3 \int_{0.5}^{1} 2\left(\int_{c}^{1} \frac{1}{2} (3-1) \frac{1}{s} (1 - (2s-1))^{3-2} 2\right)^{0.5} = 2.4299 \tag{7}$$

The relations between the designer's payoffs in the various architectures depend on the precise relations between the function governing the distribution of abilities and the convex cost function. There is no general ranking of architectures. But we can offer a comparative statics result by varying the degree of convexity. As usual, we say that an increasing convex function β is "more convex" than another increasing convex function α if there exists a strictly monotonic increasing and convex function μ such that $\beta = \mu \circ \alpha$. ³¹

Proposition 5 Assume that a designer facing contestants with a given convex cost function γ prefers a Parallel architecture to the Grand Architecture from the point of view of total effort maximization. Then this preference extends to any situation where contestants have a more convex cost function δ .

Proof: See Appendix B.

The same methods as above generally show that an increase in the convexity of the cost function makes a given split more advantageous. Hence, apart from the bilaterality that is built in many sports, a standard design where only two contestants compete against each other and where the winner gets the prize (which may be the ability of competing at the next stage) can be rationalized by strongly increasing marginal costs of effort³².

prize splits.

³¹This partial order is equivalent to the one obtained by comparing curvatures according to the Arrow-Pratt index. This index is used for comparing the risk-aversion of agents with increasing and concave utility functions. The same logic applies of course to increasing convex functions.

³²Note that the sub-contests in Rosen's (1986) pioneering paper on the subject, have two competitors and a unique prize by assumption.

Our next result connects the two designer's goals by showing that in any instance where the parallel design is beneficial to a maximizer of the expected value of highest effort, it is also beneficial to a maximizer of the expected total effort. Recall that a split design increases the effort of low ability types while lowering the effort of high ability types. Intuitively, the benefit of a split for a designer who only focuses on the highest effort is amplified for a designer that maximizes total effort (and hence takes into account all types).

Proposition 6 Assume that cost functions are convex, and assume that a Parallel Architecture dominates the Grand Architecture for a designer who maximizes the expected value of the highest effort. Then the same preference extends to a designer who maximizes the expected value of total effort.

Proof: See Appendix B.

Our last example shows that a designer that maximizes total effort may benefit from awarding interim prizes to all finalists in the first stage of a two-stage contest.

Example 3 Consider a tournament with four contestants such that the winners of two semi-finals meet in the final. The prize sum is equal to one. Assume that contestants' abilities are uniformly distributed on the interval [0.5,1], and that the cost function is $\gamma(x) = x^2$.

The distribution of abilities in stages 1 and 2 are given, respectively, by

$$F(c) = 2c - 1$$

 $G(c) = F_{(1,2)}(c) = (2c - 1)^2 + 2(2c - 1)(2 - 2c) = -4c^2 + 8c - 3$

Let α be the prize for a winner of a semifinal. The equilibrium bids in stages 1 and 2 are given, respectively, by:

$$b_1(c) = \left[\int_c^1 \left[\alpha + (1 - 2\alpha)(1 - G(c) - c(b_2(c))^2 \right] \frac{1}{s} dF(s) \right]^{0.5}$$

$$b_2(c) = \left[\int_c^1 (1 - 2\alpha) \frac{1}{s} dG(s) \right]^{0.5}$$

Total effort is given by

$$R(\alpha) = 2 \int_{m}^{1} b_2(c) dG(c) + 4 \int_{m}^{1} b_1(c) dF(c)$$

If the entire prize is awarded to the winner of the final, total effort is R(0) = 2.23. But, if the winner of the final gets a prize of 0.5 while each of the two first-stage winners gets a prize of 0.25 in the first stage, total effort is R(0.25) = 2.66.

6 Conclusion

We have compared the performance of robust architectures in contests with privately informed agents. We first analyzed the properties of contests where the designer splits the contestants into several parallel sub-contests (each with an equal number of contestants). Then we analyzed two-stage contests where the winners of the sub-contests at the first stage compete among themselves in a final while the first-round losers are eliminated.

We studied here the performance of contest designs that can be implemented by a designer without detailed knowledge about the underlying situation. We did not perform here a "mechanism design analysis" in the usual sense, since the basic contest structure remains fixed: we only allow for a restricted set of design variations. In particular, none of our main results relied on properties of the underlying distribution of abilities in the population (which is unlikely to be known to the designer in each instance where the contest is used). It seems intuitive that our "architectures" are relatively robust to non-dramatic changes in the environment, and that these architectures can be ex-ante specified and implemented, without any precise knowledge about the particularities in each application.

While our model is such that the information released in the dynamic contest has a very simple structure, we believe that an interesting avenue is to focus on the role of information in contests with multiple rounds. Another important avenue for future research is the embedding the present analysis in a model of competition among contest designers. Models of competing mechanism designers are rare (either because they are notoriously difficult, or because they immediately lead to "Bertrand paradoxes"), but we think that significant progress can be made by studying the realistic and relevant scenario where only the contest architecture may be varied (but not other features). Since the contest architecture influences the expected payoffs of the participating agents, it is interesting to analyze which agents en-

gage in which contests. The technical tools provided in this paper should also be of use for these extended models.

7 Appendix A: Order Statistics and Stochastic Dominance Relations

We set here the framework for stochastic dominance arguments involving order statistics and functions thereof. The results given without proofs are taken from the excellent textbook by Shaked and Shanthikumar (1994).

Definition 1 For any two random variables, Y and Z with distributions G and H respectively denote their hazard rates by $r_y = \frac{G'(s)}{1-G(s)}$ and $r_z = \frac{H'(s)}{1-H(s)}$. Y is said to be smaller than Z in the hazard rate order (denoted by $Y \leq_{hr} Z$) if $\forall s$, $r_y(s) \geq r_z(s)$. Y is said to be smaller than Z in the usual stochastic order (denoted by $Y \leq_{st} Z$) if $\forall s$, $G(s) \geq H(s)$.

Theorem 5 The following relations hold:

- 1. If $Y \leq_{hr} Z$ then $Y \leq_{st} Z$;
- 2. If $Y \leq_{st} Z$ then $E[Y] \leq E[Z]$;
- 3. If $Y \leq_{st} Z$ and E[Y] = E[Z] then G = H;
- 4. If $Y \leq_{hr} Z$ and w is any increasing [decreasing] function then $w(Y) \leq_{hr} [\geq_{hr}] w(Z)$;
- 5. If $Y \leq_{st} Z$ and w is any increasing [decreasing] function then $w(Y) \leq_{st} [\geq_{st}] w(Z)$.

Definition 2 Let Y, Z be two random variables such that $E[g(Y)] \leq E[g(Z)]$ for all increasing convex [concave] functions g. Then Y is said to be smaller than Z in the increasing convex order, denoted by $Y \leq_{icx} Z$ [Y is said to be smaller than Z in the increasing concave order³³, denoted by $Y \leq_{icv} Z$]

³³In the economics literature this order is sometimes called "second-order stochastic dominance." But note that some authors use this term to obtain a variability ranking of random variables with the same mean. Here we need the more general definition.

Theorem 6 Let Y and Z be two random variables with distributions H and G respectively, such that $E[Y] \leq E[Z]$.

- 1. Assume that the distributions H and G are single-crossing such that $G \ge H$ for $x \le x^*$ and $G \le H$ and for $x \ge x^*$. Then $Y \le_{icx} Z$.
- 2. Assume that the distributions H and G are single-crossing such that $G \ge H$ for $x \ge x^*$ and $G \le H$ and for $x \le x^*$. Then $Y \le_{icv} Z$.

In order to apply the above results to our framework, let C denote the random variable governing a contestant's ability, and let F be the corresponding distribution function. We denote by $C_{(i,n)}$ the random variable corresponding to the i-th order statistic out of n independent variables, each identical to C (that is, $C_{(n,n)}$ is the highest order statistic, etc...), and we denote by $F_{(i,n)}$ the respective distributions. It is well known that:

$$F_{(i,n)}(s) = \sum_{j=i}^{n} {n \choose j} F(s)^{j} (1 - F(s))^{n-j}$$
(8)

$$dF_{(i,n)}(s) = \frac{n!}{(i-1)!(n-i)!}F(s)^{i-1}(1-F(s))^{n-i}F'(s)$$
(9)

Theorem 7 The following relations hold³⁴:

1.
$$C_{(i,n)} \leq_{hr} C_{(i+1,n)}$$
 for $i = 1, 2, ..., n-1$

2.
$$C_{(i-1,n-1)} \leq_{hr} C_{(i,n)}$$
 for $i = 2, 3, ..., n-1$

3.
$$C_{(i,n)} \leq_{hr} C_{(i,n-1)} \text{ for } i = 1, 2, ..., n-1$$

Fix agent j, and let $F_i^n(s), 1 \le i \le n$ denote the probability that agent j with type s meets n-1 competitors such that i-1 of them have lower types, and n-i have higher types. We then have

 $^{^{34}}$ For future references we reproduce here the strong results, involving the hazard rate order. In this paper we employ the weaker versions (implied by Theorem 5-1) for the usual stochastic order .

$$F_i^n(s) = \frac{(n-1)!}{(i-1)!(n-i)!} (F(s))^{i-1} (1 - F(s))^{n-i}$$
(10)

Lemma 3 The following relations hold:

1.
$$F_1^n(s) = 1 - F_{(1,n-1)}(s)$$

2.
$$F_i^n(s) = F_{(i-1,n-1)}(s) - F_{(i,n-1)}(s)$$
, for all $i = 2, ..., n-1$

3.
$$nF(s)dF_{(i,n-1)} = idF_{(i+1,n)}$$
, for all $i = 2, ..., n-1$

Proof: The relations follow immediately from the respective definitions given above. \Box

8 Appendix B: Proofs

8.1 Proof of Lemma 1

The proof consists of several steps:

1) We first show that:

$$b_{n,p}(m) > b_{n,r}(m); \tag{11}$$

$$b_{n,p}(c) < b_{n,r}(c)$$
 for c in a neighborhood $[1 - \varepsilon, 1)$. (12)

This proves that the continuous equilibrium effort functions $b_{n,r}(c)$ and $b_{n,p}(c)$ must cross at least once in the interval (m, 1).

By equation (3), we obtain $b_{n,i}(m) = \frac{1}{i}E[\frac{1}{C_{(i,n-1)}}]$. By Theorem 7-1, we know that $C_{(p,n-1)} \leq_{st} C_{(r,n-1)}$. Since the function $\frac{1}{x}$ is strictly decreasing, we obtain by Theorem 5-5 that $\frac{1}{C_{(p,n-1)}} \geq_{st} \frac{1}{C_{(r,n-1)}}$, and hence that $E[\frac{1}{C_{(p,n-1)}}] > E[\frac{1}{C_{(r,n-1)}}]$. Since r > p, we finally obtain $\frac{1}{p}E[\frac{1}{C_{(p,n-1)}}] > \frac{1}{r}E[\frac{1}{C_{(r,n-1)}}]$, which means that $b_{n,p}(m) > b_{n,r}(m)$.

In order to prove relation (12), note that $b_{n,p}(1) = b_{n,r}(1) = 0$. Moreover, we have $b_{n,p}^{(i)}(1) = 0$ for all derivatives of order $i, 1 \le i \le n - p - 1$, and $b_{n,r}^{(i)}(1) = 0$ for all derivatives of order $i, 1 \le i \le n - r - 1$. This yields

$$\lim_{c \to 1} \frac{b_{n,r}^{(n-r)}(c)}{b_{n,p}^{(n-r)}(c)} = \infty \tag{13}$$

The result for the neighborhood of 1 follows then by L'Hospital's rule.

2) We now show that the equation $b'_{n,p}(c) = b'_{n,r}(c)$ has a unique solution in the interval (m,1). We have:

$$b'_{n,p}(c) - b'_{n,r}(c)$$

$$= \frac{1}{p} \frac{1}{c} dF_{(p,n-1)}(c) F'(c) - \frac{1}{r} \frac{1}{c} dF_{(r,n-1)}(c) F'(c)$$

$$= \frac{1}{p} \frac{1}{c} \frac{(n-1)!}{(p-1)!(n-p-1)!} F(c)^{p-1} (1-F(c))^{n-p-1} F'(c) - \frac{1}{r} \frac{1}{c} \frac{(n-1)!}{(r-1)!(n-r-1)!} F(c)^{r-1} (1-F(c))^{n-r-1} F'(c)$$

$$= \frac{1}{c} (n-1)! F(c)^{p-1} (1-F(c))^{n-r-1} F'(c) \times \left[\frac{1}{p!(n-p-1)!} (1-F(c))^{r-p} - \frac{1}{r!(n-r-1)!} F(c)^{r-p} \right]$$
(14)

Hence, for $c \in (m, 1)$

$$b'_{n,p}(c) - b'_{n,r}(c) = 0 \Leftrightarrow \frac{r!(n-r-1)!}{p!(n-p-1)!} = \left(\frac{F(c)}{1 - F(c)}\right)^{r-p}$$
(15)

The function $H(c) = (\frac{F(c)}{1 - F(c)})^{r-p}$ is strictly monotonically increasing with H(m) = 0 and $H(1) = \infty$. Hence,

$$H(c) = \frac{r!(n-r-1)!}{p!(n-p-1)!} \Leftrightarrow b'_{n,p}(c) - b'_{n,r}(c) = 0$$
(16)

has a unique solution in (m, 1), as desired.

3) By step 1, we know that the equation $b_{n,p}(c) = b_{n,r}(c)$ must have at least one solution in the interval (m, 1). It remains to show that the solution is unique.

Assume, by contradiction, that on (m, 1) the equation $b_{n,p}(c) - b_{n,r}(c) = 0$ has at least two distinct solutions c_1, c_2 with $c_2 > c_1$. On the interval [m, 1] there are then at least three

distinct solutions (the additional one is of course c=1). Applying Rolle's Theorem, we obtain two points d_1 and d_2 such that $d_1 \in (c_1, c_2), d_2 \in (c_2, 1)$ and

$$b'_n(d_1) - b'_k(d_1) = b'_n(d_2) - b'_k(d_2) = 0. (17)$$

Since both $d_1, d_2 \in (m, 1)$ we obtain a contradiction to step 2. \Box

8.2 Proof of Proposition 2

Let $R_{n,p}$ denote the designer's payoff in a contest with n contestants and p equal prizes. Then we have:

$$R_{n,p} = n \int_{m}^{1} b_{n,p}(c) dF(c)$$

$$= \frac{n}{p} \int_{m}^{1} \left[\int_{c}^{1} \frac{1}{s} dF_{(p,n-1)}(s) \right] dF(c)$$

$$= \frac{n}{p} \left(F(c) \int_{c}^{1} \frac{1}{s} dF_{(p,n-1)}(s) \Big|_{m}^{1} + \int_{m}^{1} F(c) \frac{1}{c} dF_{(p,n-1)}(c) \right)$$

$$= \int_{m}^{1} \frac{1}{c} \frac{n}{p} F(c) dF_{(p,n-1)}(c)$$

$$= \int_{m}^{1} \frac{1}{c} dF_{(p+1,n)}(c) = E\left[\frac{1}{C_{(n+1,n)}}\right]$$
(18)

We integrated by parts in the second line, and we used Lemma 3-3 in fourth line. Note also that

$$F(c) \int_{c}^{1} \frac{1}{s} dF_{(p,n-1)}(s) \Big|_{m}^{1} = 0$$
 (19)

Theorems 7-1,3 and 5-1 in Appendix A imply that $C_{p,n} \leq_{st} C_{p+1,n}$ and that $C_{p,n} \leq_{st} C_{p,n-1}$. Since the function $\frac{1}{x}$ is decreasing, we obtain by Theorem 5-5 that $\frac{1}{C_{(p+1,n)}} \leq_{st} \frac{1}{C_{(p,n)}}$ and that $\frac{1}{C_{(p,n-1)}} \leq_{st} \frac{1}{C_{(p,n)}}$. The results follow then by Theorem 5-2 in Appendix A. \Box

8.3 Proof of Theorem 1 (concave cost functions)

Let g denote the inverse of the cost function. Then g is increasing and convex.

Let $b_{n,1}$ denote the random variable governing the equilibrium bid with n contestants, one prize worth 1 and linear cost functions. Similarly, let $b_{\frac{n}{t},1}$ denote the respective random variable with $\frac{n}{t}$ contestants, where t > 1.

By Lemma 2, there exists c^* such that $b_{n,1} \geq b_{\frac{n}{t},1}$ for $c \leq c^*$ and $b_{n,1} \leq b_{\frac{n}{t},1}$ for $c \geq c^*$. Because $\frac{1}{t} < 1$, the same property holds for the random variables $b_{n,1}$ and $\frac{1}{t}b_{\frac{n}{t},1}^{35}$. Since these last two functions are strictly decreasing, their distribution functions are, respectively, $1 - F(b_{n,1}^{-1})$ and $1 - F((\frac{1}{t}b_{\frac{n}{t},1})^{-1})$, where F is the distribution of abilities.

By the single-crossing property of $b_{n,1}$ and $\frac{1}{t}b_{\frac{n}{t},1}$, their distribution functions are also single-crossing in the sense of Theorem 6-1 in Appendix A. By Proposition 2-1, we know that

$$\frac{n}{t}E[b_{\frac{n}{t},1}] \le nE[b_{n,1}] \Leftrightarrow E[\frac{1}{t}b_{\frac{n}{t},1}] \le E[b_{n,1}] \tag{20}$$

Single-crossing and inequality 20 imply that $\frac{1}{t}b_{\frac{n}{t},1} \leq_{icx} b_{n,1}$ (for the increasing convex stochastic order, see Definition 2 and Theorem 6 in Appendix A). Hence, we obtain that for any increasing convex function g,

$$E[g(\frac{1}{t}b_{\frac{n}{t},1})] \le E[g(b_{n,1})] \tag{21}$$

In GA the designer has a payoff $nE[g(b_{n,1})]$. In t-PA the designer has a payoff of $t \cdot \frac{n}{t}E[g(\frac{1}{t}b_{\frac{n}{t},1})] = nE[g(\frac{1}{t}b_{\frac{n}{t},1})]$, and the desired result follows from inequality 21.

8.4 Proof of Proposition 3

The equilibrium effort function is given by

$$b_{n,1}(c) = \int_{c}^{1} \frac{1}{s} dF_{(1,n-1)}(s)$$
 (22)

The designer's payoff is given by:

$$P_{n,1} = E[b_{n,1}(C_{(1,n)})] (23)$$

We have:

$$P_{n,1} = \int_{m}^{1} b_{n,1}(c) dF_{(1,n)}(c) = \int_{m}^{1} \left[\int_{c}^{1} \frac{1}{s} dF_{(1,n-1)}(s) \right] dF_{(1,n)}(c)$$

The proof that $b_{\frac{n}{t},1}$ and $b_{n,1}$ are single-crossing is based on three facts: 1) $b_{n,1}(m) > b_{\frac{n}{t},1}(m)$; 2) $b_{n,1}(c) < b_{\frac{n}{t},1}(c)$ for c in a neighborhood $[1-\varepsilon,1)$; 3) $(b_{n,1})'(c) - (b_{\frac{n}{t},1})'(c)$ is monotonic. It can be easily verified that these facts continue to hold if we replace $b_{\frac{n}{t},1}$ by $\beta b_{\frac{n}{t},1}$, where $0 < \beta < 1$.

$$= \int_{m}^{1} F_{(1,n)}(s) \frac{1}{s} dF_{(1,n-1)}(s)$$

$$= \int_{m}^{1} \frac{1}{s} [1 - (1 - F(s))^{n}] dF_{(1,n-1)}(s)$$

$$= \int_{m}^{1} \frac{1}{s} dF_{(1,n-1)}(s) - \int_{m}^{1} \frac{1}{s} (1 - F(s))^{n} dF_{(1,n-1)}(s)$$

$$= \int_{m}^{1} \frac{1}{s} dF_{(1,n-1)}(s) - (n-1) \int_{m}^{1} \frac{1}{s} (1 - F(s))^{2n-2}$$

$$= \int_{m}^{1} \frac{1}{s} dF_{(1,n-1)}(s) - \frac{n-1}{2n-1} \int_{m}^{1} \frac{1}{s} F_{(1,2n-1)}(s)$$

$$= b_{n,1}(m) - \frac{n-1}{2n-1} b_{2n,1}(m)$$

$$= E[\frac{1}{C_{(1,n-1)}}] - \frac{n-1}{2n-1} E[\frac{1}{C_{(1,2n-1)}}]$$
(24)

We used integration by parts in the second line, and the formula for $F_{(1,n)}$ and the binomial expansion formula in the third line. Taking the limit, we obtain:

$$\lim_{n \to \infty} P_{n,1} = \lim_{n \to \infty} \left[E\left[\frac{1}{C_{(1,n-1)}}\right] - \frac{n-1}{2n-1} E\left[\frac{1}{C_{(1,2n-1)}}\right] \right]$$
$$= \frac{1}{m} - \frac{1}{2m} = \frac{1}{2m}$$

8.5 Proof of Theorem 2

The designer's expected payoff in GA is

$$\int_{m}^{1} b_{n,1}(c)dF_{(1,n)}(c) \tag{25}$$

In t-PA there are n contestants, each exerting an effort of $\frac{1}{t}b_{\frac{n}{t},1}(c)$. The designer is interested in the highest realization, thus her payoff is given by

$$\int_{m}^{1} \frac{1}{t} b_{\frac{n}{t},1}(c) dF_{(1,n)}(c) \tag{26}$$

Denote $\Delta_t(c) = b_{n,1}(c) - \frac{1}{t}b_{\frac{n}{t},1}(c)$. By Lemma 2 and the argument in the proof of Theorem 1 (concave cost functions), the random variables $\frac{1}{t}b_{\frac{n}{t},1}$ and $b_{n,1}$ are single-crossing: there exists a unique point $c^* = c(n,t)$ such that $\Delta_t(c) > 0$ for all $c \in [m,c^*)$ and $\Delta_t(c) < 0$ for all $c \in (c^*,1]$. The difference between the designer's payoffs in GA and t-PA is

$$\Delta = \int_{m}^{1} \Delta_{t}(c) dF_{(1,n)}(c) = n \int_{m}^{1} \Delta_{t}(c) (1 - F(c))^{n-1} dF(c)$$
 (27)

The analog difference for the case where the designer maximizes total effort is positive by Theorem 1:

$$\widetilde{\Delta} = n \int_{m}^{1} \Delta_{t}(c) dF(c) > 0$$
(28)

Note that the expression for Δ is obtained by multiplying each term in expression $\widetilde{\Delta}$ by the decreasing function $H(c) = (1 - F(c))^{n-1}$. Hence all positive terms in $\widetilde{\Delta}$ are multiplied by relatively high values of H(c), while all negative terms are multiplied by relatively lower values. Therefore, if $\widetilde{\Delta}$ is positive, Δ must be positive too. \square

8.6 Proof of Proposition 4

Let $R(\alpha)$ be the designer's expected payoff where α is the first-stage prize. We want to show that the marginal effect of a first-stage prize on the designer's revenue is negative. Therefore the optimal prize in the first stage should be zero, or, equivalently, the entire-prize sum should be allocated in the second stage. The marginal effect of a prize awarded in the first stage is

$$R'(\alpha) = n \int_{m}^{1} \frac{d}{d\alpha} b_1(c) dF(c) + t \int_{m}^{1} \frac{d}{d\alpha} b_2(c) dG(c)$$
 (29)

Substituting (5) and (4) in (29) yields

$$R'(\alpha) = n \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dF_{(1,k-1)}(s) dF(c)$$

$$-nt \int_{m}^{1} \int_{c}^{1} \left[(1 - G(c))^{t-1} - c \int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) \right] \frac{1}{s} dF_{(1,k-1)}(s) dF(c)$$

$$-t^{2} \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) dG(c)$$

$$< n \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dF_{(1,k-1)}(s) dF(c) - t^{2} \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) dG(c)$$

$$= t \left[k \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dF_{(1,k-1)}(s) dF(c) - t \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) dG(c) \right]$$

The inequality holds since the expression $\left[(1-G(c))^{t-1}-c\int_c^1\frac{1}{s}dG_{(1,t-1)}(s)\right]$ is non-negative for all $c\in[m,1]$.

Thus, $R'(\alpha) < 0$ if

$$k \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dF_{(1,k-1)}(s) dF(c) \le t \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) dG(c)$$
(30)

For the left-hand side of equation (30) we obtain:

$$k \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dF_{(1,k-1)}(s) dF(c) = k \int_{m}^{1} F(c) \frac{1}{c} dF_{(1,k-1)}(c)$$

$$= k \int_{m}^{1} F(c) (k-1) \frac{1}{c} (1-F(c))^{k-2} dF(c)$$

$$= k \int_{m}^{1} \frac{1}{c} (k-1) (1-F(c))^{k-2} dF(c) - k \int_{m}^{1} (1-F(c)) (k-1) \frac{1}{c} (1-F(c))^{k-2} dF(c)$$

$$= k \int_{m}^{1} \frac{1}{c} (k-1) ((1-F(c))^{k-2} - (1-F(c))^{k-1}) dF(c)$$

$$= k \int_{m}^{1} \frac{1}{c} dF_{(1,k-1)}(c) - (k-1) \int_{m}^{1} \frac{1}{c} dF_{(1,k)}(c)$$

$$(31)$$

For the right hand side of equation (30) we obtain:

$$t \int_{m}^{1} \int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) dG(c) = t \int_{m}^{1} G(c) \frac{1}{c} dG_{(1,t-1)}(c)$$

$$= \int_{m}^{1} \frac{1}{c} t(t-1) G(c) (1 - G(c))^{t-2} dG(c)$$

$$= \int_{m}^{1} \frac{1}{c} dG_{(2,t)}(c) \ge \int_{m}^{1} \frac{1}{c} dG_{(2,2)}(c)$$

$$= \int_{m}^{1} \frac{1}{c} 2G(c) dG(c) = 2 \int_{m}^{1} \frac{1}{c} (1 - (1 - F(c))^{k}) k(1 - F(c))^{k-1} dF(c)$$

$$= 2 \int_{m}^{1} \frac{1}{c} dF_{(1,k)} - \int_{m}^{1} \frac{1}{c} dF_{(1,2k)}$$
(32)

The inequality in the third line of the above derivation follows by Theorems 5-4 and 7-3 in Appendix A. By inequality (30) and by derivations (31) and (32) we obtain that $R'(\alpha) < 0$ if

$$(k+1)\int_{m}^{1} \frac{1}{c} dF_{(1,k)}(c) \ge k \int_{m}^{1} \frac{1}{c} dF_{(1,k-1)}(c) + \int_{m}^{1} \frac{1}{c} dF_{(1,2k)}(c)$$

Define now:

$$v(c) := (k+1)dF_{(1,k)}(c); \ w(c) := kdF_{(1,k-1)}(c) + dF_{(1,2k)}(c)$$

and note that

$$\int_{m}^{1} (v(c) - w(c))$$

$$= \int_{0}^{1} (k+1)k(1-F(c))^{k-1} - k(k-1)(1-F(c))^{k-2} + 2k(1-F(c))^{2k-1}dF = 0. (33)$$

The Lemma below shows that the equation v(c) = w(c) has three intersection points: m, 1, and $x \in (m, 1)$. Moreover, v(c) > w(c) for all m < c < x and v(c) < w(c) for all x < c < 1. Since $\frac{1}{c}$ is a decreasing function, we obtain that

$$(k+1) \int_{m}^{1} \frac{1}{c} dF_{(1,k)}(c) - k \int_{m}^{1} \frac{1}{c} dF_{(1,k-1)}(c) - \int_{m}^{1} \frac{1}{c} dF_{(1,2k)}(c)$$

$$= \int_{m}^{1} \frac{1}{c} (v(c) - w(c)) > \int_{m}^{1} (v(c) - w(c)) = 0$$

Hence, the inequality (30) holds, and this implies that $R'(\alpha) < 0.\square$

Lemma 4 The equation v(c) = w(c) has three intersection points: m, 1, and $x \in (m, 1)$.

Proof: The proof is similar to the proof of Lemma 1. Note that $v(m) = w(m) = k(k+1)dF^1(m)$ and v(1) = w(1) = 0. We now show that v(x) = w(x) for a unique $x \in (m, 1)$.

We have:

$$v'(c) = -k(k-1)(k-1)(1-F(c))^{k-2}(dF^{1}(c))^{2}$$

$$w'(c) = -k(k-1)(k-2)(1-F(c))^{k-3}(dF(c))^{2} - 2k(2k-1)(1-F(c))^{2k-2}(dF(c))^{2}$$

For c = m, this yields

$$v'(m) = -k(k-1)^2 (dF(m))^2$$

$$w'(m) = (-k(k-1)(k-2) - 2k(2k-1))(dF(m))^2$$

A simple calculation shows that, for all $k \geq 2$,

$$|w'(m)| - |v'(m)| = (dF(m))^2 3k(k-1) > 0$$

Thus, v(c) > w(c) for small values of c.

For c=1, we get that $v^{(i)}(1)=0$ for all derivatives of order $i,1\leq i\leq k-2$, and $w^{(i)}(1)=0$ for all derivatives of order $i,1\leq i\leq k-3$. L'Hospital's rule yields

$$\lim_{c \to 1} \frac{v^{(k-2)}(c)}{w^{(k-2)}(c)} = 0$$

Thus v(c) < w(c) for large values of c.

The above discussion shows that v(c) = w(c) has a solution in the interval (m, 1). It remains to show that the solution is unique.

$$v'(c) - w'(c) = (1 - F(c))^{k-3} (dF(c))^2 (-k(k-1)(k-1)(1 - F(c)))$$
$$+k(k-1)(k-2) + 2k(2k-1)(1 - F(c))^{k+1})$$

Thus, for $c \in (m, 1)$,

$$v'(c) - w'(c) = 0 \Leftrightarrow$$

$$k(k-1)(k-1) - 2k(2k-1)(1 - F(c)) = k(k-1)(k-2)$$

The function H(c) = k(k-1)(k-1) - 2k(2k-1)(1-F(c)) is strictly monotonically increasing with H(1) > k(k-1)(k-2) and H(m) < k(k-1)(k-2). Thus, the equation v(c) = w(c) has a unique solution in the interval (m, 1) as desired.

8.7 Proof of Theorem 4

Consider a two-stage tournament where, in the first stage, there are t > 1 sub-contests. We compute the expected highest effort in the second stage³⁶ where the t winners from the first stage compete against each other for a unique prize worth 1. The symmetric equilibrium effort function in the second stage is given by

$$b_2(c) = \int_c^1 \frac{1}{s} dG_{(1,t-1)}(s) \tag{34}$$

³⁶It can be verified that the highest bid in the first stage converges to zero when the numbers of competitors approaches infinity.

where $G = F_{(1,\frac{n}{t})} = 1 - (1 - F)^{\frac{n}{t}}$. The expected value of the highest effort in the second stage is given by (see also the derivation in the proof of Proposition 3):

$$P_{t,1}^{G} = \int_{m}^{1} b_{2}(c)dG_{(1,t)}(c) = \int_{m}^{1} \left[\int_{c}^{1} \frac{1}{s} dG_{(1,t-1)}(s) \right] dG_{(1,t)}(c)$$

$$= \int_{m}^{1} \frac{1}{s} G_{(1,t)}(s) dG_{(1,t-1)}(s)$$

$$= \int_{m}^{1} \frac{1}{s} [1 - (1 - G(s))^{t}] (t - 1) (1 - G(s))^{t-2} dG(s)$$

$$= \int_{m}^{1} \frac{1}{s} [1 - (1 - F(s))^{n}] (t - 1) \frac{n}{t} (1 - F(s))^{n-1-\frac{n}{t}} dF(s)$$

$$= \frac{n(t-1)}{t} \int_{m}^{1} \frac{1}{s} (1 - F(s))^{n-1-\frac{n}{t}} dF(s) - \frac{n(t-1)}{t} \int_{m}^{1} \frac{1}{s} (1 - F(s))^{2n-1-\frac{n}{t}} dF(s)$$

$$= \frac{n(t-1)}{t} \frac{t}{n(t-1)} b_{n+1-\frac{n}{t}}(m) - \frac{n(t-1)}{t} \frac{t}{n(2t-1)} b_{2n+1-\frac{n}{t}}(m)$$

$$= b_{n+1-\frac{n}{t}}(m) - \frac{(t-1)}{(2t-1)} b_{2n+1-\frac{n}{t}}(m) = E\left[\frac{1}{C_{(1,n-\frac{n}{t})}}\right] - \frac{t-1}{2t-1} E\left[\frac{1}{C_{(1,2n-\frac{n}{t})}}\right]$$
(35)

Taking the limit, we obtain:

$$\lim_{n \to \infty} P_{t,1}^G = \lim_{n \to \infty} \left[E\left[\frac{1}{C_{(1,n-1)}}\right] - \frac{(t-1)}{(2t-1)} E\left[\frac{1}{C_{(1,2n-1)}}\right]$$

$$= \frac{1}{m} - \frac{(t-1)}{m(2t-1)} = \frac{1}{m} \left(1 - \frac{(t-1)}{(2t-1)}\right) = \frac{1}{m} \frac{t}{2t-1}$$

The claim follows by noting that the function $\frac{t}{2t-1}$ is monotonically decreasing for t > 1 and that $\lim_{t\to\infty} \frac{1}{m} \frac{t}{2t-1} = \frac{1}{2m}$ which is the expected value of the highest effort in the grand architecture (see Proposition 3).

8.8 Proof of Proposition 5

Let g denote the inverse of γ and d denote the inverse of δ . By assumption, there exists a strictly increasing and concave function μ such that $d = \mu \circ g$.

We compare GA with t-PA. For contestants with cost function γ , the designer's payoff is $nE[g(b_{n,1})]$ in GA and $nE[g(\frac{1}{t}b_{\frac{n}{t},1})]$ in t-PA. By our assumption we know that

$$E[g(\frac{1}{t}b_{\frac{n}{t},1})] \ge E[g(b_{n,1})] \tag{36}$$

Recall that the decreasing random variables $\frac{1}{t}b_{\frac{n}{t},1}$ and $b_{n,1}$ are single-crossing. Since g is increasing, we obtain that $g(\frac{1}{t}b_{\frac{n}{t},1})$ and $g(b_{n,1})$ are single-crossing in the sense of Theorem 6-2. Together with inequality (36), this yields $g(\frac{1}{t}b_{\frac{n}{t},1}) \geq_{icv} g(b_{n,1})$ (see Definition 2 in Appendix A for the *increasing concave stochastic order*). Hence, for the concave function μ we obtain that

$$E[\mu(g(\frac{1}{t}b_{\frac{n}{t},1}))] \ge E[\mu(g(b_{n,1}))]$$
(37)

By definition, this is equivalent to

$$E[d(\frac{1}{t}b_{\frac{n}{t},1})) \ge E[d(b_{n,1})]$$

as desired. \Box

8.9 Proof of Proposition 6

For a maximizer of the expected highest effort, the payoffs in GA and in t-PA are, respectively:

$$\int_{m}^{1} g(b_{n,1}(c))dF_{(1,n)}(c) \tag{38}$$

$$\int_{m}^{1} g(\frac{1}{t} b_{\frac{n}{t},1}(c)) dF_{(1,n)}(c) \tag{39}$$

In the case of a preferred t-PA we must have

$$\int_{m}^{1} g(\frac{1}{t} b_{\frac{n}{t},1}(c)) dF_{(1,n)}(c) \ge \int_{m}^{1} g(b_{n,1}(c)) dF_{(1,n)}(c) \tag{40}$$

For the maximizer of expected total effort, the payoffs in GA and in t-PA are, respectively:

$$n \int_{m}^{1} g(b_{n,1}(c))dF(c) \tag{41}$$

$$t \cdot \frac{n}{t} \int g(\frac{1}{t}b_{\frac{n}{t},1}(c))dF(c) = n \int_{m}^{1} g(\frac{1}{t}b_{\frac{n}{t},1}(c))dF(c)$$
 (42)

By the same method as in the proof of Theorem 2, we obtain that inequality (40) implies that

$$n \int_{m}^{1} g(\frac{1}{t} b_{\frac{n}{t},1}(c)) dF(c) \ge n \int_{m}^{1} g(b_{n,1}(c)) dF(c)$$
(43)

 \Box

9 References

- 1. Amegashie, J.A. (1999), "The Design of Rent-Seeking Competitions: Committees, Preliminary and Final Contests", *Public Choice*, 99: 63-76
- 2. Amman, E., and Leininger, W. (1996), "Asymmetric All-pay Auctions with Incomplete Information: The Two-Player Case", Games and Economic Behavior, 14, 1-18
- 3. Athey, S. (2000), "Monotone Comparative Statics under Uncertainty," forthcoming in Quarterly Journal of Economics.
- 4. Baye, M., Kovenock, D., and de Vries, C. (1993), "Rigging the Lobbying Process," American Economic Review, 83, 289-294.
- 5. Baye, M., Kovenock, D., and de Vries, C. (1996), "The All-Pay Auction with Complete Information," *Economic Theory*, 8, 291-305.
- 6. Barut, Y., and Kovenock, D. (1998), "The Symmetric Multiple Prize All-Pay Auction with Complete Information," *European Journal of Political Economy*, 14, 627-644.
- 7. Che, Y., and Gale, I. (2003), "Optimal Design of Research Contests," American Economic Review, 93, 646-671.
- 8. Bulow, J., and Klemperer, P. (1999), "The Generalized War of Attrition," *American Economic Review* 89, 175-189.
- 9. Clark, D., and Riis, C. (1998), "Competition over More than One Prize," *American Economic Review*, 88, 276-289.
- Dasgupta, P. (1986), "The Theory of Technological Competition," in J. Stiglitz and G. Mathewson (eds.), New Developments in the Analysis of Market Structure, Cambridge: MIT Press.
- 11. Dixit, A. (1987), "Strategic Behavior in Contests," American Economic Review, 77(5), 891-898.

- 12. Frank, R. and Cook. P. (1995), *The Winner-Take-All Society*, New York: The Free Press.
- 13. Fullerton, R. and McAfee, P. (1999), "Auctioning Entry into Tournaments," *Journal of Political Economy* 107, 573-605.
- 14. Gavious, A., Moldovanu, B., and Sela, A. (2003) "Bid Costs and Endogenous Bid Caps," Rand Journal of Economics, 33(4), 709-722.
- 15. Glazer, A., and Hassin, R. (1988), "Optimal Contests," Economic Inquiry 26, 133-143.
- 16. Green, J., and Stokey, N. (1983), "A Comparison of Tournaments and Contracts," Journal of Political Economy, 91(3), 349-364.
- 17. Hilman, A., and Riley, J. G. (1989) "Politically Contestable Rents and Transfers," *Economics and Politics*, 17-39.
- 18. Inderst, R., Mueller, H., and Waerneryd, K. (2001), "Distributional Conflict in Organizations," Discussion Paper, University of Mannheim.
- 19. Krishna, V., and Morgan, J. (1997), "An Analysis of the War of Attrition and the All-Pay Auction", *Journal of Economic Theory*, 343-362.
- 20. Krishna, V., and Morgan, J. (1998), "The Winner-Take-All Principle in Small Tournaments," Advances in Applied Microeconomics 7, 61-74.
- 21. Lazear, E., and Rosen, S. (1981), "Rank Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841-864.
- 22. Moldovanu, B., and Sela, A. (2001), "The Optimal Allocation of Prizes in Contests," American Economic Review, 91, 542-558.
- 23. Moulin, H. (1986), Game Theory for the Social Sciences, New York, New York University Press.
- 24. Nalebuff, B., and Stiglitz, J. (1983), "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics*, 14, 21-43.

- 25. Rosen, S. (1986), "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, **76**, 701-715.
- Shaked, M., and Shanthikumar, J. (1994), Stochastic Orders and Their Applications,
 Probability and Mathematical Statistics, San Diego: Academic Press.
- 27. Taylor, C. (1995), "Digging for Golden Carrots: An Analysis of Research Tournaments," *American Economic Review*, 85, 873-890.
- 28. Tullock, G. (1980), "Efficient Rent-Seeking", in J. Buchanan et al., eds, *Towards a Theory of the Rent-Seeking Society*, College Station: Texas A&M University Press.
- 29. Varian, H. (1980), "A Model of Sales", American Economic Review, 70(4), 651-659.
- 30. Weber, R. (1985), "Auctions and Competitive Bidding," in H.P. Young, ed., Fair Allocation, Providence RI: American Mathematical Society.
- 31. Windham, P.(1999), "A Taxonomy of Technology Prizes and Contests," in *Concerning Federally Sponsored Inducement Prizes in Engineering and Science*, National Academy Press.