

Head Starts in Contests

Ron Siegel*

Department of Economics

Northwestern University

September 2011

(first version: December 2010)

Abstract

This paper studies equilibrium behavior in a class of games that models asymmetric multiprize competitions in which players' costs are not necessarily strictly decreasing. Such costs accommodate head starts, which capture incumbency advantages, prior investments, and technological differences. I provide an algorithm that constructs the unique equilibrium in which players do not choose weakly-dominated strategies, and apply it to study multiprize all-pay auctions with head starts. A comparison to the standard all-pay auction shows that the strategic effects of head starts differ substantially from those of differing valuations.

*Department of Economics, Northwestern University, Evanston, IL 60208 (e-mail: r-siegel@northwestern.edu). I thank Ed Lazear, Ilya Segal, Jeremy Bulow, Robert Wilson, Jonathan Levin, and Yuval Salant for helpful discussions in the early stages of this project. Some of the work was conducted at the Institute for Advance Studies at Hebrew University; I thank the Institute for its hospitality. I thank Henrique Roscoe for very useful comments and excellent research assistantship. I thank Jeremy Bulow, Simone Galperti, Philipp Kircher, Wolfgang Leininger, Gilat Levy, Marco Ottaviani, Andrea Prat, Canice Prendergast, Ronny Razin, Yuval Salant, Michael Whinston, and seminar participants at Ben Gurion, Bonn, Chicago Booth, École Polytechnique, Hong Kong, LSE, PUC, Tel-Aviv, and the Second Brazilian Game Theory Workshop for helpful comments and suggestions.

1 Introduction

A head start is an advantage a competitor has at the outset of a competition, which must be overcome by other competitors if they are to have any chance of winning. For example, in a competition for promotions in which a worker's tenure plays a role, workers who have worked longer for the firm enjoy a head start. In a political campaign, the incumbent may have a head start in the form of a constituency advantage over an entrant. In a lobbying scenario, some lobbyists may have a reputational advantage. In a research and development (R&D) setting, some firms may start out with more advanced technologies than other firms. In a sports competition, some competitors may enjoy a head start in the form of a "handicap." But despite the prevalence of head starts in real-world competitions, their effects on the outcome of a competition are not well understood, because the relevant existing models typically do not allow for head starts.

This paper investigates equilibrium behavior in a single-prize and multiprize *contest* model that allows for a wide range of asymmetries among players, including head starts. In a contest, each player chooses a non-negative "score" and pays the associated cost, and the players with the highest scores obtain one prize each. Each player is characterized by his valuation for a prize and a continuous cost function, which specifies the cost associated with each score. Valuations and cost functions may differ across players, which allows for a high degree of heterogeneity among them. The primitives of the model are commonly known, capturing players' knowledge of the asymmetries among them.¹ Contests are defined in Section 2

In contrast to most existing models, players' costs need not be strictly, or even weakly, increasing in score. This accommodates head starts, subsidies that are contingent on a minimal level of performance, and non-monotonic mappings between a player's performance and the cost he incurs. A head start is modeled as an interval with lower bound 0 on which the player's costs are 0. A subsidy is modeled as an interval on which the player's costs are constant, where the lower bound of the interval corresponds to the minimal level of performance required to obtain the subsidy, and the length of the interval corresponds to the increase in performance resulting from the subsidy. Non-monotonic costs model situations in which better performance is not necessarily associated with higher costs. For example, competing researchers often enjoy

¹Models of competition with incomplete information were studied by Erwin Amann and Wolfgang Leininger (1996), Benny Moldovanu and Aner Sela (2001, 2006), Todd Kaplan et al. (2003), Sergio Parreiras and Anna Rubinchik (2009), and Siegel (2011) among others. When not restricted to two or ex-ante symmetric players, these models do not fully characterize equilibrium behavior. In contrast, the contest model presented here postulates complete information and provides a full characterization of equilibrium. Other models of competition postulate a probabilistic relation between competitors' efforts and prize allocation. See Gordon Tullock (1980) and Edward Lazear and Sherwin Rosen (1981). For a comprehensive treatment of the literature on competitions with sunk investments, see Shmuel Nitzan (1994) and Kai A. Konrad (2007).

certain parts of the research process (such as proving a conjecture), but dislike other parts of the process (such as writing a transparent proof). In the context of a competition for promotions, many professionals are intrinsically motivated and derive utility from working reasonably hard (but not too hard) even in the absence of compensation.

Costs that are not strictly increasing introduce a modeling subtlety, which is easily seen by considering head starts. Suppose, for example, that in a competition for promotions a worker's overall performance is determined by adding his tenure t to some non-negative function of the quality of his work. This implies that the worker has a head start of t , which is a lower bound on his performance in the competition. But in the contest in which the player's costs are 0 up to t , the player can choose scores lower than t . These scores are, of course, weakly dominated by t . To avoid this modeling mismatch, I restrict attention to equilibria in which players do not choose weakly dominated scores. To characterize such equilibria, it suffices to consider a restricted class of contests, in which players' costs are weakly increasing. This is because, for any contest, a simple mapping identifies a contest with weakly increasing costs whose equilibria coincide with those of the original contest when players do not choose weakly dominated scores. Section 2.1 describes this mapping.

Section 3 presents, for a large class of contests with weakly increasing costs, a constructive characterization of the unique mixed-strategy equilibrium in which players do not choose weakly dominated scores. This class includes contests with $m + 1$ players and m prizes, as well as many contests with more than $m + 1$ players. The construction relies on knowledge of players' equilibrium payoffs, which is provided by Ron Siegel's (2009) payoff result, and generalizes the equilibrium uniqueness results of Siegel (2010), which apply to contests with strictly increasing costs.²

Section 4 applies the equilibrium construction results to study multiprize all-pay *auctions* with head starts, in which all players have the same commonly known valuation for a prize.³ Each player is characterized by a non-negative head start, all players pay their bids, and a player's score equals the sum of his bid and his head start. There is a unique equilibrium, which depends qualitatively on the relative sizes of players' head starts. A Matlab procedure that constructs the equilibrium is available on my website.⁴ Figure 1 depicts the equilibrium when there are three players and two prizes.⁵ Player i 's head start is a_i , with $a_1 > a_2 > a_3 = 0$, and the value of each of the two prizes is $V > a_1$. In equilibrium, each player has an "atom" at his

²The beginning of Section 3 explains why Siegel's (2010) results cannot in general be applied to contests with weakly increasing costs, and also why applying these results to approximating contests with strictly increasing costs is not very useful.

³One setting in which players have the same valuations is when prizes are monetary.

⁴<http://faculty.wcas.northwestern.edu/~rsi665/>.

⁵Players' strategies are described in Section 4.1.

head start (that is, chooses his head start with positive probability), and with the remaining probability chooses scores from one or two intervals.

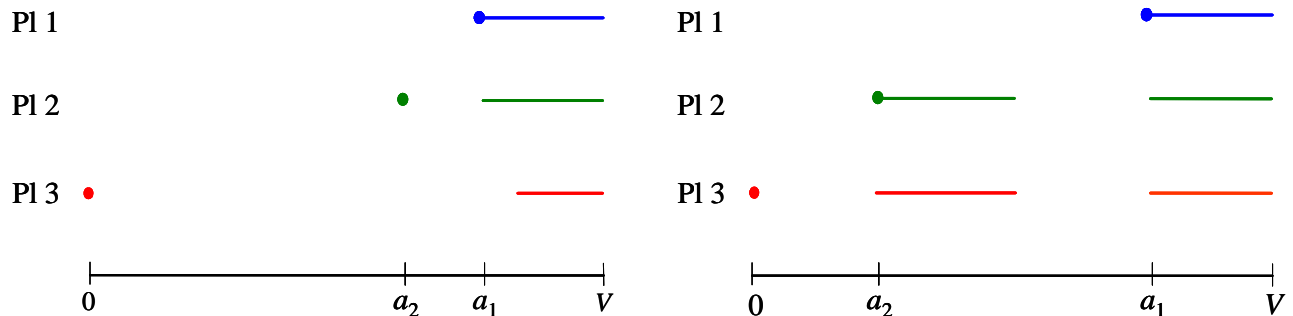


Figure 1: The supports of players’ strategies (dots represent atoms) in the unique equilibrium of a three-player, two-prize all-pay auction with head starts when a_2 is relatively large (left) and when a_2 is relatively small (right)

When a_2 is relatively small (the right-hand side of Figure 1), the equilibrium can be thought of as including two “competition zones:” a “minor league” and a “major league.” Players 2 and 3 choose whether to invest and, conditional on investing, in which league to compete. Player 1 competes only in the major league when he invests. In the major league, players compete for both prizes. In the minor league, players compete for only one prize (player 1 always beats any player who competes in the minor league). Conditional on competing within a league, all players are equally aggressive.⁶ When a_2 is relatively large (the left-hand side of Figure 1), the equilibrium includes only one competition zone, the major league. And conditional on competing, player 3 is more aggressive than the other, stronger players. Thus, when facing two strong opponents, player 3 displays an “all-or-nothing” behavior. The two different equilibrium configurations generated by varying players’ head starts contrast with the single equilibrium configuration, depicted in Figure 5 below, that arises when there are two players and one prize. Two-player all-pay auctions with head starts have been studied by Kai A. Konrad (2002, 2004) and Todd Kaplan et al. (2003).⁷ All-pay auctions with two prizes have also been studied, independently from this paper, by Casas-Arce and Martínez-Jerez (2010).⁸

⁶Aggressive is formalized in Sections 4.1 and 4.2 as conditional first-order stochastic dominance of players’ strategies.

⁷They considered potentially different valuations for the different players, which the equilibrium construction results in this paper also accommodate. René Kirkegaard (2009) analyzed head starts in an incomplete-information all-pay auction with two-players and one prize.

⁸In contrast to this paper, Casas-Arce and Martínez-Jerez (2010) did not solve the all-pay auction with more than two prizes, and their analysis and results do not extend to contests with non-linear costs.

Section 4.2 characterizes, for any number of prizes, all the equilibrium configurations that can be generated by varying players' head starts. The number of such configurations is exponential in the number of prizes, and coincides with a well-known combinatorial object, the Catalan number.⁹ This contrasts with the single equilibrium configuration that arises in standard all-pay auctions, in which every player's score equals his bid and players differ only in their valuations for a prize. Standard all-pay auctions have been used to model rent-seeking and lobbying activities (Arye L. Hillman and Dov Samet (1987), Hillman and John G. Riley (1989), Michael R. Baye, Dan Kovenock, and Casper de Vries (1993, 1996)), competitions for a monopoly position (Tore Ellingsen (1991)), competitions for multiple prizes (Derek J. Clark and Christian Riis (1998)), sales (Hal Varian (1980)), and R&D races (Partha Dasgupta (1986)).¹⁰ Figure 2 depicts the unique equilibrium of a standard all-pay auction with three players and two prizes, in which $V_1 > V_2 > V_3$.

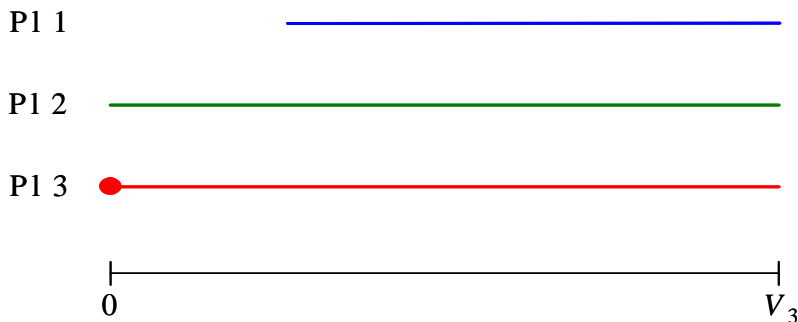


Figure 2: The supports of players' strategies (dots represent atoms) in the unique equilibrium of a three-player, two-prize standard all-pay auction with $V_1 > V_2 > V_3$

Section 4.3 compares the equilibrium predictions of standard all-pay auctions to those of all-pay auctions with head starts. In both contests, stronger players (those with higher valuations or larger head starts) win a prize with higher probability. But whereas in a standard all-pay auction stronger players expend more, the ranking of expenditures is reversed in an all-pay auction with head starts, regardless of the equilibrium configuration. In addition, the two contests differ in their predictions of a natural measure of players' aggressiveness. Section 4.4 concludes the analysis of all-pay auctions with head starts by considering some contest design issues.

⁹The Catalan number arises in many combinatorial contexts. To the best of my knowledge, however, this is the first time it appears in a game-theoretic setting.

¹⁰Variants of the all-pay auction have been used to investigate the effect of lobbying caps (Che & Gale (1998, 2006) and Kaplan & Wettstein (2006)), R&D races with endogenous prizes (Che & Gale (2003)), non-linear, ordered costs with one prize (Julio González-Díaz (2009)), and the effects of conditional investments (Siegel (2010)).

Appendix A contains the proofs of Proposition 1 and of results from Section 3. Appendix B applies the equilibrium construction results to derive the equilibrium of a three-player, two-prize all-pay auction with head starts. Appendix C contains proofs of results from Section 4. The Online Appendix contains two examples.

2 Model and Existing Results

In a *contest*, n players compete for m homogeneous prizes, $0 < m < n$. The set of players $\{1, \dots, n\}$ is denoted by \mathcal{N} . Players compete by each choosing a score from $[0, \infty)$, simultaneously and independently. Each of the m players with the highest scores wins one prize. In case of a relevant tie, any procedure may be used to allocate the tie-related prizes among the tied players.

Each player i is characterized by his *valuation for a prize*, $V_i > 0$, and his *cost function*, $c_i : [0, \infty) \rightarrow \mathbb{R}$. Given $\mathbf{s} = (s_1, \dots, s_n)$, where s_i is player i 's chosen score, *player i 's payoff* is

$$u_i(\mathbf{s}) = P_i(\mathbf{s}) V_i - c_i(s_i),$$

where $P_i : [0, \infty)^n \rightarrow [0, 1]$ is *player i 's probability of winning*, which satisfies

$$P_i(\mathbf{s}) = \begin{cases} 0 & \text{if } s_j > s_i \text{ for } m \text{ or more players } j \neq i, \\ 1 & \text{if } s_j < s_i \text{ for } n - m \text{ or more players } j \neq i, \end{cases}$$

and $\sum_{j \in \mathcal{N}} P_j(\mathbf{s}) = m$. Differences among players are captured by their different valuations for a prize and by differences in their cost functions. The primitives of the contest are commonly known.

I make the following assumptions.

C1 c_i is continuous and piecewise analytic.¹¹

Assumption C1 allows for non-monotonic costs. Such costs are useful in modeling competitions in which better performance is not necessarily more costly, such as competitions for promotions between intrinsically motivated workers. Head starts are accommodated by having costs equal to 0 on some interval of scores with lower bound 0.

Let $l_i = \inf_{s_i \geq 0} c_i(s_i)$.

¹¹A function is piecewise analytic on $[0, T]$ if $[0, T]$ can be partitioned into a finite number of closed intervals such that the restriction of the function to each interval is analytic. Analytic functions include polynomials, the exponent function, trigonometric functions, and power functions. Sums, products, compositions, reciprocals, and derivatives of analytic functions are analytic (see, for example, Charles P. Chapman (2002)).

C2 $c_i(s_i) = l_i$ for some $s_i \geq 0$, and $\lim_{s_i \rightarrow \infty} c_i(s_i) > V_i + l_i$.

The second part of Assumption C2 means that sufficiently high scores are prohibitively costly. Assumption C3 below is a genericity condition, which completes the description of a contest. It uses the following definition.

Definition 1 (i) *Player i 's reach r_i is the highest score whose cost does not exceed his highest possible payoff from losing by more than his valuation for a prize. That is, $r_i = \max \{s_i : c_i(s_i) = V_i + l_i\}$. Re-index players in (any) decreasing order of their reach, so that $r_1 \geq r_2 \geq \dots \geq r_n$.*

(ii) *Player $m + 1$ is the marginal player.*

(iii) *The threshold T of the contest is the reach of the marginal player: $T = r_{m+1}$.*

(iv) *Let $c_i^L(s_i) = \min_{x \geq s_i} c_i(x)$. The power w_i of player i is his payoff if he chooses the least costly score no lower than the threshold and wins: $w_i = V_i - c_i^L(T)$. By definition, $w_{m+1} = -l_{m+1}$.*

Note that $c_i^L \equiv c_i$ when costs are weakly increasing. For example, in a standard all-pay auction (Hillman and Samet (1987), Hillman and Riley (1989), Clark and Riis (1998)) all costs are linear, a player's reach is equal to his valuation for a prize (because $l_i = 0$), and a player's power is equal to his valuation for a prize minus the marginal player's valuation for a prize.

C3 (i) The marginal player is the only player i with power $-l_i$. (ii) The marginal player's cost function is increasing immediately below the threshold, i.e., for some $\varepsilon > 0$ and every x in $(T - \varepsilon, T)$ we have $c_{m+1}(x) < c_{m+1}(T)$.

For example, a standard all-pay auction satisfies part (ii) of Assumptions C3 because players' costs are strictly increasing. If the valuation of the marginal player is different from those of the other players, then part (i) of Assumption C3 is satisfied as well.

2.1 Equilibrium and Weakly Dominated Scores

A player's (mixed) strategy is a probability distribution over $[0, \infty)$, typically described as a cumulative distribution function (CDF), which for every score x specifies the probability that the player chooses a score lower than or equal to x . I say that *player i does not choose weakly dominated scores* if his CDF assigns probability 0 to the set of scores

$$\{x : c_i(x) \geq c_i(y) \text{ for a score } y > x\}.$$

Otherwise, I say that *player i chooses weakly dominated scores*. An equilibrium is a profile of strategies, one for each player, such that each player's strategy assigns probability 1 to the player's set of best responses given the other players' strategies.

As discussed in the introduction, this paper focuses on equilibria in which players do not choose weakly dominated scores. To investigate such equilibria it suffices to study a restricted class of contests, in which players' cost functions are non-decreasing and equal 0 at 0. To see why, given a contest C consider the modified contest \hat{C} in which player i 's costs are $\hat{c}_i(\cdot) = c_i^L(\cdot) - l_i$ and his valuation is V_i .

Proposition 1 *The set of equilibria in C in which players do not choose weakly dominated scores is the same as the set of equilibria in \hat{C} in which players do not choose weakly dominated scores. Moreover, in any such equilibrium player i 's payoff in C is precisely l_i lower than his payoff in \hat{C} .*

The proof of Proposition 1 is in Appendix A. It is straightforward to verify that \hat{C} is a contest in which players' costs are weakly increasing and equal 0 at 0 (so $\hat{l}_i = 0$ for every player i). For such contests, Assumptions C1-C3 are equivalent to the following Assumptions M1-M3.

M1 c_i is continuous, piecewise analytic, and weakly increasing.

M2 $c_i(0) = 0$, and $\lim_{s_i \rightarrow \infty} c_i(s_i) > V_i$.

M3 (i) The marginal player is the only player with power 0. (ii) The marginal player's cost function is strictly increasing immediately below the threshold.

By Proposition 1, to characterize equilibria in which players do not use weakly dominated scores it suffices to do so for contests that satisfy Assumptions M1-M3. Therefore, for the rest of the paper I restrict attention to contests that satisfy Assumptions M1-M3, and the terms "contest" and "contests" refer to such contests. Definition 1 simplifies to $r_i = \max\{s_i : c_i(s_i) = V_i\}$, $T = r_{m+1}$, and $w_i = V_i - c_i(T)$.

2.2 Existing Results

The contests considered here are a special case of Siegel's (2009) all-pay contest model.¹² The following two results, which I use in solving for equilibrium, are immediate corollaries of results in Siegel (2009).¹³ The first result characterizes players' equilibrium payoffs (without solving for equilibrium).

¹²In particular, a contest here is a generic separable contest of Siegel (2009). But it is not a special case of Siegel's (2010) contest model, since Siegel (2010) requires strictly increasing costs (which, in particular, preclude head starts).

¹³The results follow, respectively, from Theorem 1 and Theorem 2 in Siegel (2009).

Theorem 1 *In any equilibrium of a contest, the expected payoff of every player equals the maximum of his power and 0.*

In addition to giving a closed-form formula for players' equilibrium payoffs, Theorem 1 shows that players $1, \dots, m$ have positive expected payoffs (because of part (i) of Assumption M3), and players $m + 1, \dots, n$ have expected payoffs of 0.

A player *participates* in an equilibrium of a contest if with positive probability he chooses scores whose cost is positive. The second result provides a sufficient condition for players $m + 2, \dots, n$ not to participate in any equilibrium.

Theorem 2 *If the normalized cost function of the marginal player is strictly lower than that of player $i > m + 1$, that is,*

$$\frac{c_{m+1}(x)}{V_{m+1}} < \frac{c_i(x)}{V_i} \text{ for all } x \text{ such that } c_i(x) > 0,$$

then player i does not participate in any equilibrium. In particular, if this condition holds for players $m + 2, \dots, n$, then in any equilibrium only players $1, \dots, m + 1$ may participate.

3 Equilibrium Construction

I provide an algorithm that solves for the unique equilibrium of an $(m + 1)$ -player contest in which players do not choose weakly dominated scores. The algorithm generalizes the one described in Siegel (2010), which applies to contests with strictly increasing costs (in which no scores are weakly dominated). Siegel's (2010) algorithm cannot in general be applied to contests with weakly increasing costs, because it relies on certain equilibrium properties that may not hold when costs are weakly increasing.¹⁴

An alternative approach to dealing with weakly increasing costs would be to apply Siegel's (2010) algorithm to contests with strictly increasing costs that "approximate" the original contest, and consider the limit of the corresponding equilibria. This is problematic for three reasons. First, it is not clear how to derive analytical results by taking this limit, because the process entails "taking the limit" of an algorithm. Second, even if an equilibrium can be derived in this way, it may have the undesirable property that players choose weakly dominated scores. This can be seen by examining the relatively simple case of two players, for which Siegel's (2010) Theorem 3 provides a closed-form formula when costs are strictly increasing (instead of an algorithm). In this case, the "limiting equilibrium" corresponds to applying the formula to the

¹⁴These properties are that all players' CDFs reach 1 precisely at the threshold, every positive score up to the threshold is a best response for at least two players, and players' equilibrium CDFs are continuous on $(0, T)$. For example, the latter two properties do not hold in the equilibria depicted in Figure 1, and the first property does not hold in the equilibrium depicted in Figure 4 in the Online Appendix.

original contest with weakly increasing costs. In the context of Figure 4 below, which depicts a two-player all-pay auction in which player 1 has a head start of $a_1 > 0$, the formula would specify player 1’s equilibrium strategy to be $G_1(x) = x$ for every x in $[0, 1]$. In particular, player 1 would choose scores lower than his head start with probability a_1 . Figure 5 describes another equilibrium, in which players do not choose weakly dominated scores. This illustrates the third difficulty with the “limit” approach, namely, that it leaves unaddressed the issue of equilibrium uniqueness, even within the class of equilibria in which players do not choose weakly dominated scores.

Section 3.1 describes an equilibrium construction algorithm that applies to contests with weakly increasing costs, and shows that the resulting profile of CDFs is indeed an equilibrium in which players do not choose weakly dominated scores.¹⁵ Section 3.3 shows that this is the unique equilibrium in which players do not choose weakly dominated scores, and explains to what extent the construction applies to n -player contests.

3.1 The Algorithm

The algorithm applies to an $(m + 1)$ -player contest and has five stages. Stages 1, 2, and 4 parallel the algorithm of Siegel (2010). Stages 3 and 5 are novel additions that deal with the possibility that players’ cost functions may be constant on some intervals. I denote by $\mathbf{G} = (G_1, \dots, G_{m+1})$ the profile of CDFs constructed by the algorithm, where $G_i(x)$ is the probability that player i chooses a score lower than or equal to x .

The construction of \mathbf{G} proceeds from 0 to T by identifying *active* players and *checkpoints*. A player is active on his best response set, and a checkpoint is a score above which the set of active players may change. Stage 1 identifies a score $x_0 \geq 0$ and the value $\mathbf{G}(x_0)$ such that $\mathbf{G}(x) = 0$ for any $x < x_0$ and the value $\mathbf{G}(x_0)$ suffices to continue the construction above x_0 .¹⁶ If $G_{m+1}(x_0) = 1$, then x_0 is the last checkpoint and the algorithm proceeds to Stage 5, which completes the definition of \mathbf{G} . Otherwise, x_0 is the first checkpoint and the algorithm proceeds to Stage 2. Given a checkpoint x that has been reached, Stage 2 determines a set of candidate players at x , $\mathcal{CP}(x)$, which is a superset of the set $\mathcal{A}^+(x)$ of players active immediately above x . If $\mathcal{CP}(x)$ contains a player whose cost function is “constant at x ,” i.e., constant on some interval of scores with lower endpoint x , then the algorithm proceeds to Stage 3. Otherwise, the cost functions of all players in $\mathcal{CP}(x)$ are “increasing at x ,” i.e. strictly increasing on some interval with lower bound x , and the algorithm proceeds to Stage 4. Stage 3 sets $\mathcal{A}^+(x)$ to be the set

¹⁵For a two-player contest, the equilibrium constructed by the algorithm can also be derived from the equilibrium specified by Siegel’s (2010) Theorem 3 by, for any maximal interval on which a player’s cost function is constant, transferring all the mass the player places on the interval to an atom at the top of the interval.

¹⁶This generalizes Step 1 of Siegel’s (2010) algorithm. In his algorithm, $x_0 = 0$ because all costs are strictly increasing.

of all players in $\mathcal{CP}(x)$ whose costs are constant at x , and identifies the next checkpoint \bar{x} as the upper endpoint of the longest interval with lower endpoint x on which the cost function of a player in $\mathcal{A}^+(x)$ is constant. \mathbf{G} is then set to equal $\mathbf{G}(x)$ on $[x, \bar{x})$, and at most one player has an atom at \bar{x} . The player and the size of the atom are identified. Stage 4 uses $\mathcal{CP}(x)$ to define a function $S_{x,y}$ with a unique fixed point $H(x,y)$ for scores y immediately above x , and uses $H(x,y)$ to determine the set $\mathcal{A}^+(x)$. \mathbf{G} is then extended continuously to $[x, \bar{x}]$ using a closed-form formula, where \bar{x} is the next checkpoint. The checkpoint \bar{x} is also identified. If the CDF of at least one player reaches 1 at \bar{x} in Stages 3 or 4, then \bar{x} is the last checkpoint and the algorithm proceeds to Stage 5. Otherwise, the algorithm returns to Stage 2 with \bar{x} as the current checkpoint. Stage 5 completes the definition of \mathbf{G} by identifying, for each player whose CDF has not reached 1 at the last checkpoint, a score at which he has an atom that brings his CDF to 1. I show that the algorithm terminates after reaching a finite number of checkpoints, and that the resulting \mathbf{G} is indeed an equilibrium. Whenever possible, I use notation consistent with that of Siegel (2010).¹⁷ The algorithm is illustrated in Figure 3.

¹⁷An exception is my usage of checkpoints in contrast to his usage of *switching* points. A switching point of Siegel (2010) is a score above which the set of active players *does* change. Every switching point of a contest with strictly increasing costs is a checkpoint, but some checkpoints may not be switching points.

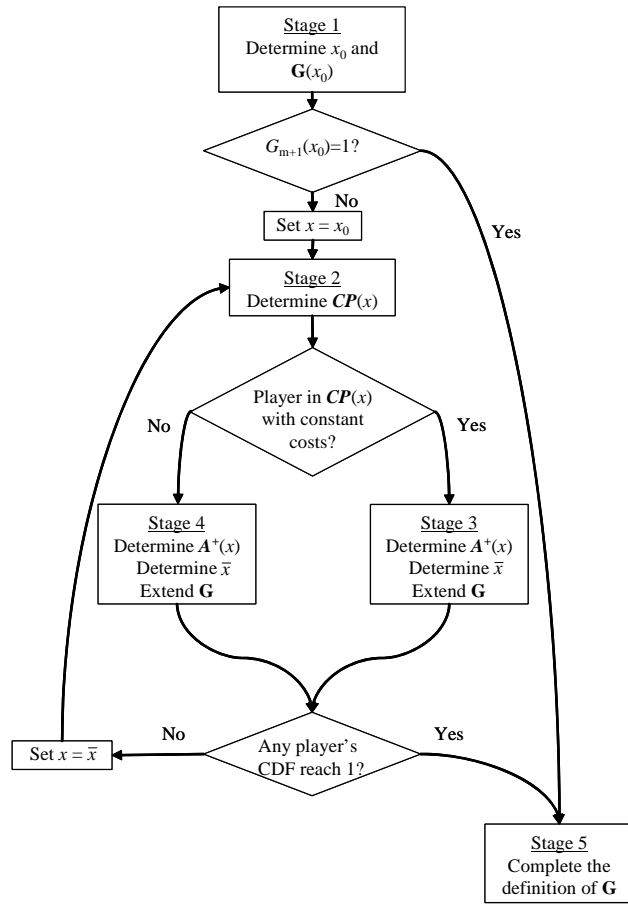


Figure 3: The equilibrium construction algorithm

Stage 1: Let $x_0 = \max \{x : c_{m+1}(x) = 0\}$. Because scores below x_0 are weakly dominated by x_0 for the marginal player, set $G_{m+1}(x) = 0$ for $x < x_0$. The value $G_{m+1}(x_0)$ is set such that at least one player other than the marginal player can have best responses immediately above x_0 , and no player can obtain more than his power by choosing scores immediately above x_0 . The following lemma determines $G_{m+1}(x_0)$ and $G_i(x)$ for every player $i < m + 1$ and $x \leq x_0$. Its proof and those of other results in this section are in Appendix A.

Lemma 1 *In any equilibrium of an $(m + 1)$ -player contest in which players do not choose weakly dominated scores, the marginal player has an atom of size $\min_{i < m+1} \left(\frac{c_i(x_0) + w_i}{V_i} \right)$ at x_0 , and the value of every other player's CDF at x_0 is 0.*

In accordance with Lemma 1, set $G_{m+1}(x_0) = \min_{i < m+1} \left(\frac{c_i(x_0) + w_i}{V_i} \right) \leq 1$ and $G_i(x) = 0$ for every player $i < m + 1$ and $x \leq x_0$. If $G_{m+1}(x_0) = 1$, then set $x^L = x_0$ as the last checkpoint and proceed to Stage 5. Otherwise, set $x = x_0$ as the first checkpoint and proceed to Stage 2.

Stage 2: Suppose \mathbf{G} has been defined up to $x < T$, and $G_i(x) < 1$ for every player i . We would like to identify a set of players, the “candidate players at x ,” who may be active immediately above x . Consider a player i who chooses scores immediately above x when other

players choose scores according to \mathbf{G} . Because CDFs are right-continuous and the player loses only when his score is lower than those of all other players, $1 - \prod_{j \neq i} (1 - G_j(x)) + \varepsilon$ is the probability that the player wins a prize, for some small $\varepsilon \geq 0$. As player i chooses scores closer to x , ε approaches 0. Therefore, if

$$(1 - \prod_{j \neq i} (1 - G_j(x))) V_i - c_i(x) < w_i, \quad (1)$$

then player i cannot be active immediately above x , because he cannot obtain his equilibrium payoff there. The candidate players at x are the other players, and for each such player i we have

$$(1 - \prod_{j \neq i} (1 - G_j(x))) V_i - c_i(x) = w_i. \quad (2)$$

Let

$$\mathcal{CP}(x) = \{\text{players } i \text{ for which (2) holds}\}$$

denote the set of candidate players at x . The set $\mathcal{CP}(x)$ contains at least two players.¹⁹ If $\mathcal{CP}(x)$ contains at least one player whose cost function is constant at x , proceed to Stage 3. Otherwise, proceed to Stage 4.

Stage 3: Suppose that $\mathcal{CP}(x)$ contains a player i whose cost function is constant at x . Constant costs imply that the *other* players' CDFs must remain constant immediately above x , otherwise player i would win a prize with too high a probability and therefore obtain more than his power immediately above x . And player i 's CDF must also remain constant immediately above x , because scores immediately above x are weakly dominated by slightly higher scores. Therefore, all players' CDFs remain constant immediately above x , so the only players for whom scores immediately above x are best responses are those players in $\mathcal{CP}(x)$ whose cost functions are constant at x . Therefore, let

$$\mathcal{A}^+(x) = \{\text{players in } \mathcal{CP}(x) \text{ whose cost functions are constant at } x\}.$$

Let $\bar{x}_i = \max\{y : c_i(y) = c_i(x)\}$ and $\bar{x} = \max\{\bar{x}_i : i \in \mathcal{A}^+(x)\}$, and choose a player i in $\mathcal{A}^+(x)$ for whom $\bar{x}_i = \bar{x}$. Based on the previous paragraph, extend \mathbf{G} to $[x, \bar{x}]$ as follows. For every player $j \neq i$ and every score y in $[x, \bar{x}]$ set $G_j(y) = G_j(x) < 1$, and for every score y in $[x, \bar{x}]$ set $G_i(y) = G_i(x) < 1$. Note that $\mathcal{CP}(\bar{x})$ contains player i ; set $G_i(\bar{x})$ to the lowest value such that $\mathcal{CP}(\bar{x})$ contains at least one other player. From (2) with l instead of i this means that

$$G_i(\bar{x}) = \min_{l \neq i} \left\{ 1 - \frac{1 - \frac{w_l + c_l(\bar{x})}{V_l}}{\prod_{j \neq l, i} (1 - G_j(x))} \right\} \text{ if } m > 1 \quad (3)$$

and

$$G_i(\bar{x}) = \frac{w_l + c_l(\bar{x})}{V_l} \text{ for } l \neq i \text{ if } m = 1. \quad (4)$$

¹⁸Lemma 11 in Appendix A shows that $>$ does not hold for any player.

¹⁹Lemma 11 in Appendix A shows this.

This value of $G_i(\bar{x})$ is the only value consistent with equilibrium. Indeed, setting $G_i(\bar{x})$ to a higher value implies that some player other than i can obtain more than his power immediately above \bar{x} . Setting $G_i(\bar{x})$ to a lower value implies that players other than i do not have best responses immediately above \bar{x} and therefore do not choose scores immediately above \bar{x} with positive probability. Because player i 's costs are increasing at \bar{x} , this means that player i too does not have best responses immediately above \bar{x} , so no player has best responses immediately above \bar{x} . This violates the following lemma.

Lemma 2 *In any equilibrium of an $(m + 1)$ -player contest, any score at which not all players' CDFs equal 1 is a best response for at least one player.*

Setting $G_i(\bar{x})$ as specified above implies that (i) $G_i(\bar{x}) = G_i(x)$ if there is another player $j \neq i$ in $\mathcal{A}^+(x)$ for whom $\bar{x}_j = \bar{x}$, and (ii) $G_i(\bar{x}) > G_i(x)$ (so player i has an atom at \bar{x}) if i is the only player in $\mathcal{A}^+(x)$ for whom $\bar{x}_i = \bar{x}$. If $G_i(\bar{x}) = 1$, then set $x^L = \bar{x}$ as the last checkpoint and proceed to Stage 5. Otherwise, set $x = \bar{x}$ as the next checkpoint and proceed to Stage 2.

Stage 4: Suppose that the cost functions of all players in $\mathcal{CP}(x)$ are increasing at x . Then Step 2 part two, Step 3, and Step 4 of Siegel's (2010) algorithm can be used to uniquely identify the set $\mathcal{A}^+(x)$ of players active immediately above x , extend \mathbf{G} up to the first checkpoint \bar{x} above x , and identify \bar{x} . I describe these steps briefly; a detailed description is given in Section 2.1 of Siegel (2010). For scores y immediately above x , let $q_i(y) = 1 - \frac{w_i + c_i(y)}{V_i} > 0$ and $\varepsilon_i(y) = -\frac{q'_i(y)}{q_i(y)} = \frac{c'_i(y)}{V_i - w_i - c_i(y)} > 0$ for every player i in $\mathcal{CP}(x)$,²⁰ and let

$$S_{x,y}(H) = \sum_{i \in \mathcal{CP}(x)} \max \{H - \varepsilon_i(y), 0\}.$$

The function $S_{x,y}(\cdot)$ is piecewise linear and convex. Since $S_{x,y}(0) = 0$, $S'_{x,y}(0) = 0$, and $\varepsilon_i(y) > 0$ for every player i in $\mathcal{CP}(x)$, which contains at least two players, $S_{x,y}$ starts below the diagonal and reaches a slope of at least 2. It therefore has a unique positive fixed point, denoted $H(x, y)$. Let $H(x, x) = \lim_{y \downarrow x} H(x, y)$. Siegel (2010) shows that $H(x, \cdot)$ and $\varepsilon_i(\cdot)$ for every player i in $\mathcal{CP}(x)$ are analytic in a right-neighborhood of x , so have right-derivatives of any order at x . The lowest-order right-derivatives at x of $\varepsilon_i(\cdot)$ and $H(x, \cdot)$ that differ (beginning with the zeroth-order derivative, that is, $\varepsilon_i(x)$ and $H(x, x)$) determine whether player i in $\mathcal{CP}(x)$ is in $\mathcal{A}^+(x)$ ($<$ means the player is in $\mathcal{A}^+(x)$, $>$ means the player is not in $\mathcal{A}^+(x)$). (This will "generically" stop at the first derivatives.) If all derivatives are equal, then player i is in $\mathcal{A}^+(x)$. Because $H(x, y)$ is higher than $\varepsilon_i(y)$ for at least two players i in $\mathcal{CP}(x)$ and scores y immediately above x , $\mathcal{A}^+(x)$ contains at least two players.

Once $\mathcal{A}^+(x)$ is identified, (2) with y instead of x for every player i in $\mathcal{A}^+(x)$ and the fact that the CDFs of players not in $\mathcal{A}^+(x)$ do not increase immediately above x can be used as follows

²⁰The inequalities follow from Lemma 12 in Appendix A.

to uniquely extend \mathbf{G} to scores y in $[x, \bar{x}]$. For every player i not in $\mathcal{A}^+(x)$, set $G_i(y) = G_i(x)$. For every player i in $\mathcal{A}^+(x)$, set

$$G_i(y) = 1 - \frac{\prod_{j \in \mathcal{A}^+(x)} q_j(y)^{\frac{1}{|\mathcal{A}^+(x)|-1}}}{q_i(y) D^{\frac{1}{|\mathcal{A}^+(x)|-1}}}, \quad (5)$$

where $D = \prod_{j \notin \mathcal{A}^+(x)} (1 - G_j(x)) > 0$ (if $\mathcal{A}^+(x) = \mathcal{N}$, then $D = 1$) and $|\mathcal{S}|$ is the cardinality of a set \mathcal{S} .²¹ This continuous extension of \mathbf{G} guarantees that immediately above x players in $\mathcal{A}^+(x)$ obtain precisely their power and players not in $\mathcal{A}^+(x)$ do not obtain more than their power. The first checkpoint \bar{x} above x is the first score above x at which G_i reaches 1 for some player i in $\mathcal{A}^+(x)$, or the cost function of a player in $\mathcal{A}^+(x)$ is constant, or the cost functions of all players in $\mathcal{A}^+(x)$ are increasing and one of the following happens: the hazard rate $(1 - G_i(y))' / G_i(y)$ of a player i in $\mathcal{A}^+(x)$ drops to 0,²² a player not in $\mathcal{A}^+(x)$ obtains his power,²³ or the score is a concatenation point of the cost function of a player in $\mathcal{A}^+(x)$ (recall that costs are piecewise-defined functions). If $G_i(\bar{x}) = 1$ for some player i in $\mathcal{A}^+(x)$, then set $x^L = \bar{x}$ as the last checkpoint and proceed to Stage 5. Otherwise, set $x = \bar{x}$ as the next checkpoint and proceed to Stage 2.

Stage 5: Because some player's CDF reaches 1 at x^L , no player i has best responses at scores y for which $c_i(y) > c_i(x^L)$. Therefore, for every player i let $x_i^{\max} = \max\{y : c_i(y) = c_i(x^L)\}$, set $G_i(y) = G_i(x^L)$ for every score y in (x^L, x_i^{\max}) , and set $G_i(y) = 1$ for every score $y \geq x_i^{\max}$. This completes the construction of \mathbf{G} .

The following lemma shows that the algorithm always stops.

Lemma 3 *The algorithm reaches Stage 5 at a checkpoint $x^L \leq T$ via a finite number of checkpoints.*

Theorem 3 summarizes the construction and shows that the resulting \mathbf{G} is an equilibrium.

Theorem 3 *For any $(m + 1)$ -player contest the algorithm constructs an equilibrium \mathbf{G} in which players do not choose weakly dominated scores. The equilibrium is characterized by a partition into a finite number of intervals of positive length, on the interior of which the set of active players remains constant.*

3.2 A Simple Example

To gain some intuition for how the algorithm works, consider the contest with two players whose cost functions are depicted in Figure 4. The common value of the prize is 1, player 1 has zero

²¹If $q_i(\bar{x}) = 0$, which happens if $c_i(\bar{x}) = c_i(T)$, then set $G_i(\bar{x}) = \lim_{y \uparrow \bar{x}} G_i(y)$.

²²Ignoring any player whose hazard rate is identically 0 starting at x .

²³That is, (2) holds for some player i not in $\mathcal{A}^+(x)$.

costs up to $a_1 < 1$ and marginal costs 1 starting from a_1 , and player 2 has marginal costs 1 starting from 0.²⁴ This is a special case of the all-pay auctions with head starts considered in Section 4.

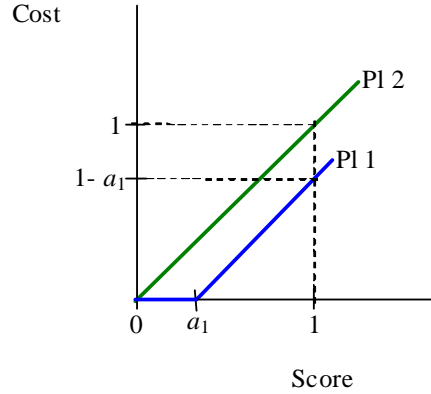


Figure 4: Players’ costs in a two-player contest with valuations and constant marginal costs of 1 and a head start for player 1

The contest’s threshold is 1 (equal to player 2’s reach), player 1’s power is a_1 , and player 2’s power is 0. Because player 2’s costs are increasing at 0, Stage 1 specifies that $x_0 = 0$. By Lemma 1, $G_2(0) = a_1 < 1$, and Stage 2 shows that $\mathcal{CP}(0) = \{1, 2\}$. Because player 1’s costs are constant at 0, the algorithm proceeds to Stage 3, which specifies that $\mathcal{A}^+(0) = \{1\}$. Therefore, the first checkpoint above 0 is $\bar{x} = a_1$, and no player’s CDF increases in $(0, a_1)$. From (4) we have that $G_1(a_1) = a_1$. The algorithm proceeds to Stage 2, and we have that $\mathcal{CP}(a_1) = \{1, 2\}$. Because both players’ costs are increasing at a_1 , the algorithm proceeds to Stage 4, and $\mathcal{A}^+(a_1) = \{1, 2\}$. From (5) we have that $G_1(y) = 1 - q_1(y)q_2(y)/q_1(y) = 1 - q_2(y) = y$ and $G_2(y) = a_1 + y - a_1 = y$ for scores y in $[a_1, \bar{x}]$, where $\bar{x} = 1$ is the score at which players’ CDFs reach 1. We then have $x^L = 1$, and Stage 5 specifies that both players’ CDFs equal 1 starting from 1. Players’ atoms and densities in the equilibrium are depicted in Figure 5.

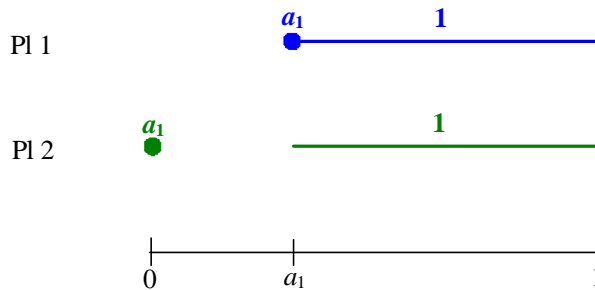


Figure 5: Players’ atoms and densities in the equilibrium of the contest depicted in Figure 4

²⁴If $a_1 \geq 1$, then player 2 has an atom of size 1 at 0, and the algorithm proceeds to Stage 5, which specifies that player 1 has an atom of size 1 at a_1 .

The Online Appendix describes the execution of the algorithm for two additional two-player contests.

3.3 Equilibrium Uniqueness and n -player Contests

The following result shows that the equilibrium \mathbf{G} constructed by the algorithm is the unique equilibrium in which players do not choose weakly dominated scores. The proof uses the property of \mathbf{G} that every player's best response set is a finite union of disjoint intervals to show that every equilibrium in which players do not choose weakly dominated scores coincides with \mathbf{G} .

Theorem 4 *For any $(m + 1)$ -player contest the algorithm constructs the unique equilibrium of the contest in which players do not choose weakly dominated scores.*

A contest with more than $m + 1$ players may have multiple equilibria in which players do not choose weakly dominated scores.²⁵ Uniqueness is guaranteed, however, if the costs of the marginal player are lower than those of player $m + 2, \dots, n$. This is the content of the following result.

Theorem 5 *If for every player $m + 2, \dots, n$ the conditions of Theorem 2 hold and the highest score whose cost is 0 is no higher than x_0 (the highest score whose cost is 0 for player $m + 1$), then the contest has a unique equilibrium in which players do not choose weakly dominated scores. In this equilibrium, players $m + 2, \dots, n$ choose their respective highest score whose cost is 0 and players $1, \dots, m + 1$ behave as in the equilibrium constructed by the algorithm for the reduced contest with players $1, \dots, m + 1$.*

Note that the conditions of Theorem 2 place no restrictions on how the cost functions of players $\mathcal{N} \setminus \{m + 1\}$ relate to each other.

4 All-Pay Auctions with Head Starts

In an *all-pay auction with head starts*, $n \geq 2$ risk-neutral players compete for $m < n$ identical prizes of common and commonly known value $V > 0$. Each player i makes a non-negative, irreversible investment, and this investment is added to his head start a_i , with $a_1 \geq \dots \geq a_n \geq 0$. Each of the m players with the highest sum wins one prize. Relevant ties are decided using any tie-breaking rule (which is specified in advance).

To model this game as a contest, let $V_i = V$ and

$$c_i(x) = \begin{cases} 0 & \text{if } x \leq a_i \\ x - a_i & \text{if } x > a_i \end{cases}.$$

²⁵See Example 3 in Siegel (2009).

For the contest to meet condition M3, assume that a_{m+1} is distinct from the head start of every other player. Player i 's reach is $a_i + V$, so the contest's threshold is $a_{m+1} + V$. Theorem 1 shows that in any equilibrium of the contest the payoff of every player $i < m + 1$ is equal to his power, i.e., $\min\{a_i - a_{m+1}, V\}$, and the payoff of the other players is 0. Not every equilibrium of the contest, however, is an equilibrium of the original game. This is because in the contest player i can choose scores lower than a_i , which are weakly dominated by a_i , whereas in the original game the lowest investment of 0 leads to the score a_i . Nevertheless, it is easy to see that the set of equilibria of the original game coincides with the set of equilibria of the contest in which players do not choose weakly dominated scores. We therefore have the following corollary of Theorem 5.

Corollary 1 *An all-pay auction with head starts in which $a_m > a_{m+1} > a_{m+2}$ has a unique equilibrium. In the equilibrium, players $m + 2, \dots, n$ and every player i for whom $a_i \geq a_{m+1} + V$ choose their respective head starts.*

Proof. The conditions of Theorem 2 hold for players $m + 2, \dots, n$ in the contest corresponding to the original game. Therefore, by Theorem 5 the contest has a unique equilibrium in which players do not choose weakly dominated scores, and in this equilibrium players $m + 2, \dots, n$ choose their respective head starts. The power of a player i for whom $a_i \geq a_{m+1} + V$ equals V , so he cannot choose scores higher than a_i in the equilibrium, and scores lower than a_i are weakly dominated by a_i . ■

Before deriving additional properties of the equilibrium in Section 4.2, let us consider the case of two prizes and three players in some detail (Corollary 1 shows that additional players do not participate).

4.1 Three Players and Two Prizes

If $a_2 \geq a_3 + V$, then every player chooses his head start, and players 1 and 2 each win a prize with certainty.²⁶ If $a_2 < a_3 + V \leq a_1$, then player 1 chooses a_1 and wins a prize with certainty, and players 2 and 3 compete as in the two-player contest of Figure 4.²⁷

If $a_1 < a_3 + V$, then players' behavior is more complicated, and depends on how large a_1 is relative to a_2 and a_3 . Appendix B applies the algorithm to derive the equilibrium. Specifically,

²⁶Stage 1 of the algorithm specifies that $x_0 = a_3$ and $G_3(a_3) = 1$, and Stage 5 of the algorithm specifies that $G_2(a_2) = G_1(a_1) = 1$.

²⁷Stage 1 specifies that $x_0 = a_3$ and $G_3(a_3) = (a_2 - a_3)/V$, Stage 3 specifies that both players' CDFs remain constant on (a_3, a_2) , Stage 4 specifies that $G_2(y) = G_3(y) = (y - a_3)/V$ for every y in $[a_2, a_3 + V]$, and both players' CDFs reach 1 at $a_3 + V$.

if $a_1 > 2a_2 - a_3 - (a_2 - a_3)^2/V$, then players' equilibrium strategies are

$$G_1(y) = \begin{cases} 0 & \text{if } y < a_1 \\ 1 - \sqrt{1 - \frac{y-a_3}{V}} & \text{if } a_1 \leq y \leq a_3 + V \\ 1 & \text{if } a_3 + V < y \end{cases},$$

$$G_2(y) = \begin{cases} 0 & \text{if } y < a_2 \\ \frac{y-a_3}{V} & \text{if } a_2 \leq y \leq a_3 + V \left(1 - \sqrt{1 - \frac{a_1-a_3}{V}}\right) \\ 1 - \sqrt{1 - \frac{a_1-a_3}{V}} & \text{if } a_3 + V \left(1 - \sqrt{1 - \frac{a_1-a_3}{V}}\right) < y < a_1 \\ 1 - \sqrt{1 - \frac{y-a_3}{V}} & \text{if } a_1 \leq y \leq a_3 + V \\ 1 & \text{if } a_3 + V < y \end{cases},$$

and

$$G_3(y) = \begin{cases} 0 & \text{if } y < a_3 \\ \frac{a_2-a_3}{V} & \text{if } a_3 \leq y < a_2 \\ \frac{y-a_3}{V} & \text{if } a_2 \leq y \leq a_3 + V \left(1 - \sqrt{1 - \frac{a_1-a_3}{V}}\right) \\ 1 - \sqrt{1 - \frac{a_1-a_3}{V}} & \text{if } a_3 + V \left(1 - \sqrt{1 - \frac{a_1-a_3}{V}}\right) < y < a_1 \\ 1 - \sqrt{1 - \frac{y-a_3}{V}} & \text{if } a_1 \leq y \leq a_3 + V \\ 1 & \text{if } a_3 + V < y \end{cases}.$$

The equilibrium is depicted in Figure 6.

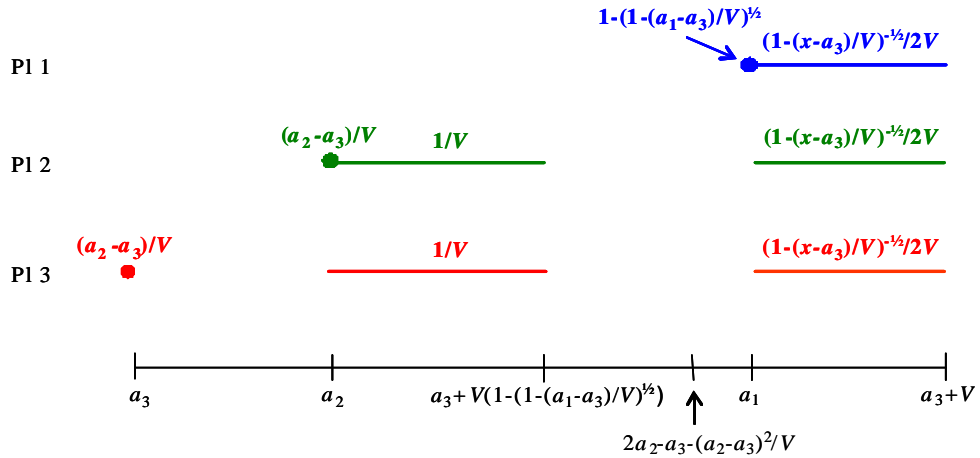


Figure 6: Players' equilibrium atoms and densities when $a_1 > 2a_2 - a_3 - (a_2 - a_3)^2/V$

If $a_1 \leq 2a_2 - a_3 - (a_2 - a_3)^2/V$, then players' equilibrium strategies are

$$G_1(y) = \begin{cases} 0 & \text{if } y < a_1 \\ \frac{y-a_2}{V-a_2+a_3} & \text{if } a_1 \leq y < 2a_2 - a_3 - \frac{(a_2-a_3)^2}{V} \\ 1 - \sqrt{1 - \frac{y-a_3}{V}} & \text{if } 2a_2 - a_3 - \frac{(a_2-a_3)^2}{V} \leq y \leq V + a_3 \\ 1 & \text{if } V + a_3 < y \end{cases},$$

$$G_2(y) = \begin{cases} 0 & \text{if } y < a_2 \\ \frac{a_1 - a_2}{V - a_2 + a_3} & \text{if } a_2 \leq y < a_1 \\ \frac{y - a_2}{V - a_2 + a_3} & \text{if } a_1 \leq y < 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \\ 1 - \sqrt{1 - \frac{y - a_3}{V}} & \text{if } 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \leq y \leq V + a_3 \\ 1 & \text{if } V + a_3 < y \end{cases},$$

and

$$G_3(y) = \begin{cases} 0 & \text{if } y < a_3 \\ \frac{a_2 - a_3}{V} & \text{if } a_3 \leq y < 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \\ 1 - \sqrt{1 - \frac{y - a_3}{V}} & \text{if } 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \leq y \leq V + a_3 \\ 1 & \text{if } V + a_3 < y \end{cases}.$$

The equilibrium is depicted in Figure 7.

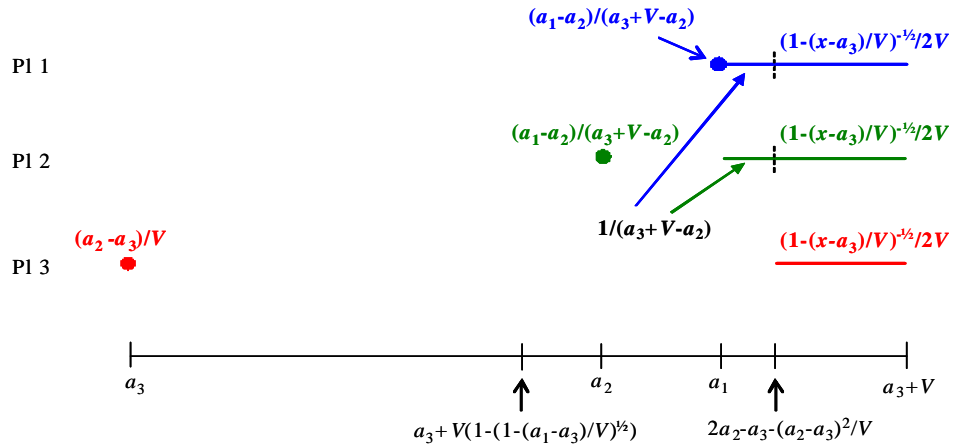


Figure 7: Players' equilibrium atoms and densities when $a_1 \leq 2a_2 - a_3 - (a_2 - a_3)^2 / V$

For example, if $V = 1$, $a_3 = 0$, and $a_2 = \frac{1}{2}$, then Figure 6 describes the equilibrium for $a_1 > \frac{3}{4}$, and Figure 7 describes the equilibrium for $a_1 \leq \frac{3}{4}$. The qualitative difference between the equilibria is due to the size of player 2's atom at a_2 , which (3) specifies to be the minimal size so that $\mathcal{CP}(a_2)$ contains a player in addition to player 2. In Figure 6, this additional player is player 3, whereas in Figure 7 this additional player is player 1.

Players' behavior can be interpreted as follows. When player 1's head start is relatively large (Figure 6), players compete in two disjoint "competition zones." In the lower zone, (a_2, a_1) , only players 2 and 3 compete. In the higher zone, $(a_1, a_3 + V)$, all three players compete. Within the lower zone, player 2 and 3 stop competing at $a_3 + V \left(1 - \sqrt{1 - (a_1 - a_3)/V}\right) < a_1$. Each player can be thought of as making a three-stage decision: whether to make an investment (choose a score whose cost is positive), in which competition zone to compete (player 1 competes only in the higher zone), and which score to choose within a zone. When player 1's head start is relatively small (Figure 7), players compete only in the higher competition zone. Moreover, conditional on competing in the higher zone, player 3 is more aggressive than the other two players, in

the sense that his conditional CDF (conditional on choosing a score in $(a_1, a_3 + V)$) first-order stochastically dominates (FOSD) the conditional CDFs of the other two players. Player 3's behavior may be interpreted as an “all-or-nothing” response to his facing two strong opponents who have relatively similar head starts. The next subsection generalizes these observations to all-pay auctions with head starts and any number of prizes.

Two additional equilibrium properties can be deduced from examining players' CDFs, regardless of whether the equilibrium is described by Figure 6 or Figure 7. The first property is that player 1's CDF FOSD that of player 2, which in turn FOSD that of player 3. This implies that a player with a higher head start wins a prize with a higher probability than a player with a lower head start. The second property is that player 1's (expected) expenditures are lower than those of player 2, which are in turn lower than those of player 3. In Figure 6 this can easily be seen by comparing player 1's strategy to player 2's strategy and then player 2's strategy to player 3's strategy. In Figure 7 this is less obvious, because players 1 and 2 incur costs when competing in the lower part of the higher competition zone, whereas with the corresponding probability player 3 does not invest. Nevertheless, the higher costs incurred by player 3 when choosing scores in the upper part of the higher competition zone more than offset this difference in expenditures. The next subsection shows that this ranking of players' probabilities of winning and their expenditures holds for any number of prizes.

4.2 Equilibrium Properties

In this subsection, I restrict attention to all-pay auctions with $m + 1$ players, whose head starts are all strictly less than $a_{m+1} + V$. This is without loss of generality, because Corollary 1 shows that without this restriction the “real” competition is between players $k + 1, \dots, m + 1$, who compete for $m - k$ prizes, where k is the number of players whose head starts are at least $a_{m+1} + V$.²⁸ All statements refer to the unique equilibrium $\mathbf{G} = (G_1, \dots, G_{m+1})$.

The first result shows that players' strategies can be ranked in terms of FOSD.

Proposition 2 *For any two players $i < j$, G_i FOSD G_j , i.e., $G_i(x) \leq G_j(x)$ for any score x . Moreover, if $G_i(x) = G_j(x)$ and $x \geq a_i$ then $G_i(y) = G_j(y)$ for all $y \geq x$.*

The proof of Proposition 2 is in Appendix C, as are the proofs of the other results in this section. A corollary of the ranking of players' strategies is that players' probabilities of winning a prize can also be ranked.

²⁸For any $n > m + 1$, players $m + 2, \dots, n$ choose their head starts and win a prize with probability 0 and players $1, \dots, k$ choose their head starts and win a prize with probability 1. All other players compete by choosing scores in $[a_{m+1}, a_{m+1} + V]$. It is easy to see that players $k + 1, \dots, m + 1$ behave as in the unique equilibrium of the reduced contest that includes only players $k + 1, \dots, m + 1$ and $m - k$ prizes.

Corollary 2 For every two players $i < j$, the probability that player i wins a prize is at least as high as that of player j .

To further characterize players' strategies, note that no player chooses scores in (a_{m+1}, a_m) .²⁹ And how players choose scores in $(a_m, a_{m+1} + V)$ depends qualitatively on all players' head starts, as shown in Section 4.1. Letting $a_0 = a_{m+1} + V$, each interval (a_i, a_{i-1}) , $i \leq m$, can be thought of as a "competition zone," in which a subset \mathcal{A}_i of the players $i, \dots, m+1$ compete, that is,

$$\mathcal{A}_i = \{j \geq i : \text{player } j \text{ has best responses in } (a_i, a_{i-1})\}.$$

I refer to the sequence $\mathcal{A}_1, \dots, \mathcal{A}_m$ as the "competitive characteristic." For example, the competitive characteristic of the equilibrium in Figure 6 is $\{1, 2, 3\}, \{2, 3\}$, and that of the equilibrium in Figure 7 is $\{1, 2, 3\}, \phi$. The following theorem enumerates properties of the competitive characteristic, shows that each player's set of best responses within a competition zone is an interval in which his CDF strictly increases, and shows that if players $k > j$ compete in the same competition zone, then player k is more aggressive than player j , conditional on choosing scores within the competition zone.

Theorem 6 The unique equilibrium \mathbf{G} of an all-pay auction with head starts has the following properties.

(a) $\mathcal{A}_1 = \{1, 2, \dots, m+1\}$. For every $i > 1$ (i) $\mathcal{A}_i = \phi$ or $\mathcal{A}_i = \{i, i+1, \dots, k_i\}$ for some $k_i > i$, and (ii) for every $i' < i$, $\mathcal{A}_i \cap \mathcal{A}_{i'} = \phi$ or $\mathcal{A}_i \subseteq \mathcal{A}_{i'}$.

(b) For every j in \mathcal{A}_i , player j 's set of best responses in (a_i, a_{i-1}) is an interval with lower bound $l_{i,j}$ and upper bound h_i , and $G_j(x) > G_j(y)$ for any two scores $x > y$ in $(l_{i,j}, h_i)$. Moreover, $h_i < a_{i-1}$ if $i > 1$, and $h_1 = a_{m+1} + V$. In particular, the upper bound h_i is the same for all the players in \mathcal{A}_i , so \mathbf{G} does not change in (h_i, a_{i-1}) . For players $k > j$ in \mathcal{A}_i we have (i) $l_{i,k} \geq l_{i,j}$ and (ii) $l_{i',j} = l_{i',k}$ for every $i' < i$ for which $\mathcal{A}_i \subseteq \mathcal{A}_{i'}$. Also, if $\mathcal{A}_i \neq \phi$ then $l_{i,i} = l_{i,i+1} = a_i$.

(c) If players $k > j$ are in \mathcal{A}_i , then $G_k(x) = G_j(x)$ for every $x \geq l_{i,k}$.

(d) For every player i , $\lim_{x \rightarrow a_{m+1} + V} G(x) = 1$ (no atoms at the threshold).

(e) If players $k > j$ are in \mathcal{A}_i , then player k 's conditional CDF FOSD that of player j . That is, for every x in (a_i, a_{i-1}) ,

$$\frac{G_k(x) - G_k(a_i)}{G_k(a_{i-1}) - G_k(a_i)} \leq \frac{G_j(x) - G_j(a_i)}{G_j(a_{i-1}) - G_j(a_i)}.$$

The competitive characteristics of the equilibria in Figures 6 and 7 satisfy part (a) of Theorem 6. The equilibrium in Figure 6 has

$$l_{1,1} = l_{1,2} = l_{1,3} = a_1, h_1 = a_3 + V, l_{2,2} = l_{2,3} = a_2, h_2 = a_3 + V \left(1 - \sqrt{1 - (a_1 - a_3)/V}\right) < a_1,$$

²⁹These scores are weakly dominated for players $1, \dots, m$. Consequently, player $m+1$ wins with probability 0 at these scores, which are costly for him, so he does not choose them.

the CDFs of players 2 and 3 coincide starting from a_2 , and the CDFs of all players coincide starting from a_1 . The inequality in part (e) of Theorem 6 is an equality for both competition zones. The equilibrium in Figure 7 has

$$l_{1,1} = l_{1,2} = a_1 < l_{1,3} = 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V}, h_1 = a_3 + V,$$

the CDFs of players 1 and 2 coincide starting from a_1 , and the CDFs of all players coincide starting from $l_{1,3}$. The inequality in part (e) of Theorem 6 is an equality for $k = 2$ and $j = 1$, and is a strict inequality for $k = 3$ and $j = 1, 2$.

While part (a) of Theorem 6 provides necessary conditions for a sequence of subsets of $\{1, \dots, m + 1\}$ to form the competitive characteristic, it leaves open the question of which competitive characteristics can be generated by changing players' head starts. To address this question, denote by \mathcal{A}^m the set of sequences $\mathcal{A}_1, \dots, \mathcal{A}_m$ that satisfy the conditions listed in part (a) of Theorem 6. For example, \mathcal{A}^2 consists of the sequences $\{1, 2, 3\}, \{2, 3\}$ and $\{1, 2, 3\}, \phi$. The following result shows that \mathcal{A}^m is related to a well-known combinatorial object.

Theorem 7 *The cardinality of \mathcal{A}^m is the Catalan number $C^m = \frac{(2m)!}{(m+1)!m!}$.*

The Catalan number C^m appears in many combinatorial contexts. For example, it is the number of different ways in which m open parentheses and m close parentheses can be sequenced so that they are correctly matched (for $m = 2$ there are two such sequences, “ $(())$ ” and “ $()()$ ”). The Catalan number C^m is also the number of complete binary trees with $m + 1$ leaves.³⁰ The idea underlying Theorem 7 is that each sequence in \mathcal{A}^m can be mapped to a sequence of $m + 1$ progressively finer partitions of $\{1, \dots, m + 1\}$, and each such sequence of partitions can be mapped to a sequence of correctly matched parentheses. The proof of Theorem 7 describes these mappings, and immediately after the proof in Appendix C I describe an explicit bijection between the set of sequences of correctly matched parentheses and \mathcal{A}^m . As discussed above, the two sequences $\{1, 2, 3\}, \{2, 3\}$ and $\{1, 2, 3\}, \phi$ that comprise \mathcal{A}^2 are the competitive characteristics of the equilibria in Figures 6 and 7. Theorem 8 shows that this equivalence holds for any number of prizes.

Theorem 8 *Every sequence in \mathcal{A}^m is a competitive characteristic for some head starts $a_{m+1} < \dots < a_1 < a_{m+1} + V$.*

Theorems 7 and 8 fully describe the number and type of qualitatively different equilibria that can be generated by varying players' head starts, as captured by the competitive characteristic. This is summarized by the following corollary.

³⁰See Thomas Koshy's (2009) book for additional applications and a derivation of the formula in Theorem 7.

Corollary 3 *The number of competitive characteristics that can be generated by varying players' head starts is the Catalan number $C^m = \frac{(2m)!}{(m+1)!m!}$. The set of these competitive characteristics is \mathcal{A}^m .*

Figure 8 depicts the five qualitatively different equilibria that correspond to the competitive characteristics for $m = 3$.

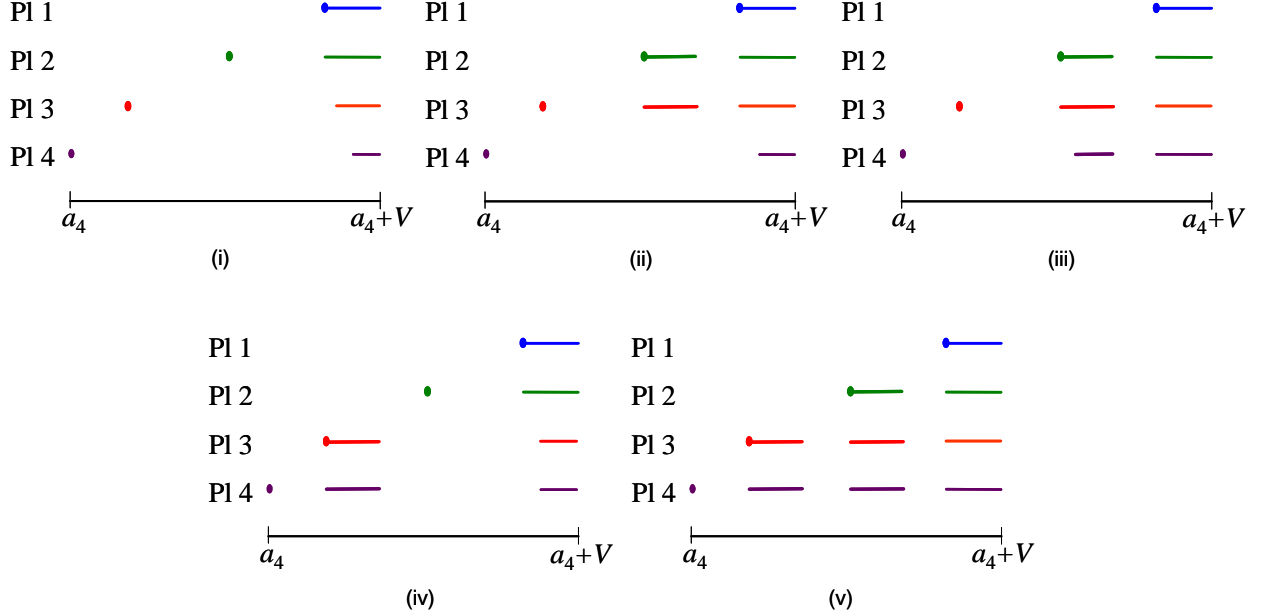


Figure 8: The supports of players' strategies (dots represent atoms) in the equilibria corresponding to the five possible competitive characteristics for $m = 3$: (i) $\{1, 2, 3, 4\}, \phi, \phi$, (ii) $\{1, 2, 3, 4\}, \{2, 3\}, \phi$, (iii) $\{1, 2, 3, 4\}, \{2, 3, 4\}, \phi$, (iv) $\{1, 2, 3, 4\}, \phi, \{3, 4\}$, and (v) $\{1, 2, 3, 4\}, \{2, 3, 4\}, \{3, 4\}$

Some intuition for the connection between the Catalan number and the competitive characteristic can be gained by considering the execution of the algorithm. For every i in $2, \dots, m + 1$, Stage 3 of the algorithm is executed once in $[a_i, a_{i-1})$, at the point at which player $i - 1$ obtains his payoff (so no player chooses scores between that point and a_{i-1}). Each of these m executions of Stage 3 can be thought of as an “open parentheses.” And for every i in $2, \dots, m + 1$, there is a first execution of Stage 4 of the algorithm in $[a_{i-1}, a_{m+1} + V)$ in which player i becomes active (and these are all the executions of Stage 4). Each of these m executions of Stage 4 can be thought of as a “close parentheses.” Moreover, the open parentheses that corresponds to i is correctly matched with the close parentheses that corresponds to i , so the sequence of m open parentheses and m close parentheses obtained by the execution of the algorithm is balanced. Figure 9 illustrates this for plate (ii) of Figure 8. Varying players' head starts changes the order in which the $2m$ instances of Stages 3 and 4 are executed: decreasing a_{i-1} causes Stage 3 in $[a_i, a_{i-1})$ to be executed closer to a_i , which reduces the number of instances of Stage 4 executed

in $[a_i, a_{i-1})$, whereas increasing a_{i-1} causes Stage 3 in $[a_i, a_{i-1})$ to be executed farther from a_i and allows for more instances of Stage 4 to be executed in $[a_i, a_{i-1})$.

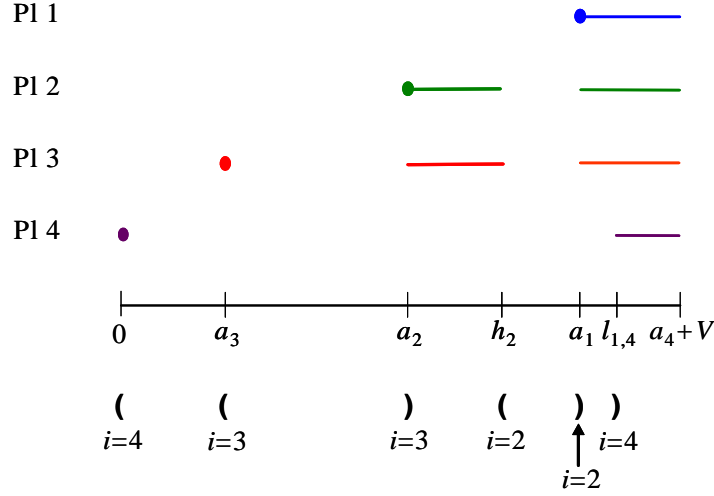


Figure 9: The sequence of parentheses corresponding to the execution of the algorithm

In Figure 9, the notation h_2 and $l_{1,4}$ is as introduced in Theorem 6. In $[a_4, a_3)$, Stage 3 is executed at a_4 (open parentheses, $i = 4$), and Stage 4 is not executed. In $[a_3, a_2)$, Stage 3 is executed at a_3 (open parentheses, $i = 3$), and Stage 4 is not executed. In $[a_2, a_1)$, Stage 4 is executed once, at a_2 , and this is the first execution of Stage 4 in $[a_2, a_4 + V)$ in which player 3 becomes active (close parentheses, $i = 3$). Stage 3 is executed at h_2 (open parentheses, $i = 2$). Finally, in $[a_1, V)$ Stage 4 is executed twice. The first time is at a_1 , and this is the first execution of Stage 4 in $[a_1, V)$ in which player 2 becomes active (close parentheses, $i = 2$). The second time is at $l_{1,4}$, and this is the first execution of Stage 4 in $[a_3, V)$ in which player 4 becomes active (close parentheses, $i = 4$).

I now turn to players' equilibrium (expected) expenditures. In the equilibria depicted in Figure 6 and plate (v) of Figure 8, the best response set of every player $i > 1$ nests that of player $i - 1$. In such cases, by using part (c) of Theorem 6, it is straightforward to show that the expenditures of player i are higher than those of player $i - 1$. But when players' best response sets are not nested, which occurs in other competitive characteristics, some player $i - 1$ chooses scores from regions to which player i assigns probability 0. This is what happens in Figure 7 in the interval $\left(a_1, 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V}\right)$ for $i = 3$. It is therefore not immediately obvious whether players' expenditures can be ranked in general. The following corollary of the previous results shows that an unambiguous ranking exists regardless of the competitive characteristic.

Proposition 3 *For any two players $i < j$, the expenditures of player j are at least as high as those of player i .*

4.3 Comparison to Standard All-pay Auctions

In a *standard all-pay auction*, $n \geq 2$ risk-neutral players compete for $m < n$ identical prizes. Player i 's commonly known valuation for a prize is V_i , with $V_1 \geq \dots \geq V_n > 0$. Each player makes a non-negative investment, and the m players who make the highest investments win one prize each. Relevant ties are decided using any tie-breaking rule (which is specified in advance).

This game is a contest in which $c_i(x) = x$, so no scores are weakly dominated. For the contest to meet condition M3, assume that V_{m+1} is distinct from the valuation of every other player. Player i 's reach is V_i , so the contest's threshold is V_{m+1} . Theorem 1 shows that in any equilibrium of the contest the payoff of every player $i < m+1$ is equal to his power, $V_i - V_{m+1}$, and the payoff of the other players is 0. Theorem 5 shows that the contest has a unique equilibrium, in which players $m+2, \dots, n$ choose 0. Clark and Riis (1998) and Siegel (2010) solved for this equilibrium (which can also be solved by the algorithm presented in this paper) and described the CDFs of players $1, \dots, m+1$ in closed form. Theorem 8, Corollary 5, and (11) in Siegel (2010) provide the following properties of the equilibrium.

Theorem 9 *The unique equilibrium \mathbf{G} of a standard all-pay auction has the following properties.*

(a) *For any two players $i < j \leq m+1$, G_i FOSD G_j , so player i 's expenditures and probability of winning a prize are at least as high as those of player j .*

(b) *The best response set of player $i \leq m+1$ is an interval with lower bound l_i and upper bound V_{m+1} , and $G_i(x) > G_i(y)$ for any two scores $x > y$ in (l_i, V_{m+1}) . For any two players $i < j \leq m+1$ we have $l_i \geq l_j$. Also, $l_m = l_{m+1} = 0$.*

Players' atoms and best-response sets in the equilibrium are depicted in Figure 10.

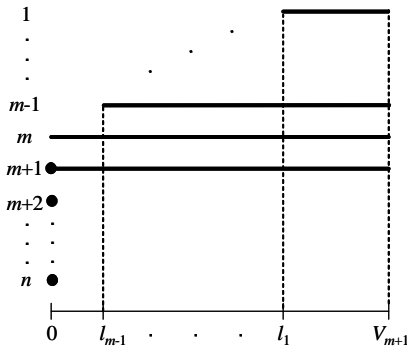


Figure 10: The supports of players' strategies (dots represent atoms) in the unique equilibrium of a standard all-pay auction

Certain real-world asymmetries among competitors are naturally modeled by head starts, whereas other asymmetries are better modeled by differences in valuations for a prize. A comparison between the results of Section 4.2 and Theorem 9 elucidates the similarities and differences between the equilibrium predictions of the two models. Both models have a unique

equilibrium, in which players $m + 2, \dots, n$ do not participate. Stronger players choose higher scores than weaker players, in terms of FOSD and therefore on average, and consequently win a prize with higher probability. But whereas the expenditures of stronger players are higher than those of weaker players when players differ in their valuations, the opposite is true with head starts.³¹ Another difference is that with head starts players $1, \dots, m + 1$ compete in one or more competition zones, depending on the relative values of the head starts. Within each competition zone, all competing players choose scores from an interval, and these intervals all end at the same score. The intervals are shorter for weaker players, who are more aggressive than stronger players in terms of conditional FOSD. Varying players' head starts changes the competition zones in which different players compete, so that player i may choose scores from up to $\min\{i, m\}$ disjoint intervals of positive length. The possible qualitatively different equilibrium configurations are captured by the different competitive characteristics, of which there are an exponential number in m . In contrast, when players differ in their valuations each player chooses scores from one interval, and these intervals all end at the threshold and are shorter for stronger players. The equilibrium can be thought of as consisting of one competition zone, $(0, V_{m+1})$, within which players $1, \dots, m + 1$ compete and stronger players are more aggressive than weaker players. Changing players' valuations does not change the qualitative form of the equilibrium. This difference between the equilibrium predictions of the two models is illustrated by the difference between Figures 8 and 10.

To conclude the comparison between the models, consider a symmetric standard all-pay auction with n players and m prizes in which all players' valuations equal $V > 0$. This game, which is a special case of the all-pay auction with heterogeneous prizes analyzed by Yasar Barut and Kovenock (1998), has a continuum of equilibria. It is easy to verify that one equilibrium is for $m + 1$ players to employ the CDF $G(x) = 1 - (1 - x/V)^{1/m}$ on $[0, V]$ and for the other players to choose 0. This equilibrium is the limit of the equilibria that correspond to a sequence of standard all-pay auctions in which players have different valuations that approach V (provided that $V_m > V_{m+1} > V_{m+2}$ along the sequence).³² The following result shows that this equilibrium is also the limit of the equilibria that correspond to a sequence of all-pay auctions with head starts that approach 0.

Proposition 4 *Consider a sequence indexed by k of all-pay auctions with head starts $a_1^k \geq \dots > a_{m+1}^k > \dots \geq a_n^k \geq 0$ for which $a_1^k \xrightarrow[k \rightarrow \infty]{} 0$. In the limit of the corresponding equilibria \mathbf{G}^k players $1, \dots, m + 1$ employ the CDF $G(x) = 1 - (1 - x/V)^{1/m}$ on $[0, V]$ and players $m + 2, \dots, n$ choose 0.*

³¹Moreover, in a standard all-pay auction the expenditures of players $1, \dots, m + 1$ are positive and no player wins a prize with probability 1. In contrast, a player with a sufficiently large head start expends 0 and wins a prize with probability 1.

³²This follows from (2), (3), and (4) in Clark and Riis (1998), and also from (11) and (12) in Siegel (2010).

4.4 Contest Design

The payoff and equilibrium characterizations can also be used for contest design. As an example, consider an $(m + 1)$ -player all-pay auction with heads starts and m prizes of value $V > a_1$, and two types of intervention by the contest administrator: handicaps and subsidies.³³ A handicap increases the handicapped player's score without affecting his output and at no direct cost to the administrator. A subsidy increases the player's output (and therefore his score) by allocating some of the prize money to defray the player's expenditures. Because of the linearity of players' costs (above their head starts), the payoff result is sufficient to determine the optimal handicaps and subsidies when the goal is to maximize aggregate expected expenditures or output.³⁴

Because every player i 's power is $a_i - a_{m+1}$, the aggregate expected expenditures are

$$mV - \sum_{i=1}^m a_i + ma_{m+1}.$$

If handicaps h_1, \dots, h_{m+1} are administered, the resulting game is strategically equivalent to an all-pay auction in which every player i 's head start is $a_i + h_i$. This implies that handicaps $h_i = a_1 - a_i$ lead to aggregate expected expenditures of

$$mV - \sum_{i=1}^m (a_i + h_i) + m(a_{m+1} + h_{m+1}) = mV - \sum_{i=1}^m (a_i + a_1 - a_i) + m(a_{m+1} + a_1 - a_{m+1}) = mV.³⁵$$

These are the highest possible aggregate expected expenditures, because aggregate expected expenditures are bounded above by the aggregate value of the prizes. Note that a player's output equals his head start plus his expenditures (the handicap is not added to the output), so maximizing aggregate output is equivalent to maximizing aggregate expected expenditures. Therefore, the maximal aggregate expected output of $mV + \sum_{i=1}^{m+1} a_i$ is achieved by setting $h_i = a_1 - a_i$. This analysis also implies that adding additional players cannot increase the aggregate expenditures or output.

If subsidies are administered, then a subsidy of b_i increases player i 's output by b_i and reduces the amount of prize money by b_i . To maximize aggregate expected output, set $b_i = a_1 - a_i$ (just

³³Similar interventions have been investigated by Qiang Fu, Jingfeng Lu, and Yuanzhu Lu (2009) in a two-player stochastic contest.

³⁴The output is the "real" score, excluding handicaps but including subsidies. This could correspond, for example, to the quality or quantity of what is produced in the course of the competition.

³⁵Although Condition M3 is violated when all players are identical, Corollary 3 in Siegel (2009) shows that in this case every player's equilibrium payoff is 0.

like the handicaps above). This gives the maximal aggregate expected output of

$$\begin{aligned}
& \underbrace{mV - \sum_{i=1}^{m+1} b_i}_{\text{prize value}} - \underbrace{\left(\sum_{m=1}^m (a_i + b_i) - m(a_{m+1} + b_{m+1}) \right)}_{\text{expected payoffs}} + \underbrace{\sum_{i=1}^{m+1} (a_i + b_i)}_{\text{subsidies and head starts}} \\
& \underbrace{\hspace{10em}}_{\text{expected expenditures}} \\
& = mV - \sum_{i=2}^{m+1} (a_1 - a_i) + (m+1)a_1 = mV + \sum_{i=1}^{m+1} a_i,
\end{aligned}$$

just like with optimal handicaps.

A tradeoff is introduced when subsidies are chosen to maximize aggregate expected expenditures. This is because increasing the marginal player's output enhances competition, which increases expenditures, but also decreases the aggregate prize value, which decreases competition and lowers expenditures. As long as $a_i + b_i \geq a_j + b_j$ for $i < j$, aggregate expenditures are

$$mV - \sum_{i=1}^{m+1} b_i - \left(\sum_{m=1}^m (a_i + b_i) - m(a_{m+1} + b_{m+1}) \right),$$

so subsidizing a player who is not the marginal player lowers expenditures both because it increases the player's payoff and because it lowers the prize value. Subsidizing the marginal player increases expenditures by $(m-1)b_{m+1}$, so it is optimal to give the marginal player a subsidy of $a_m - a_{m+1}$. At this point, any additional subsidy to player $m+1$ requires that the same additional subsidy be given to player m (otherwise player m would become the marginal player and expenditures would decrease, as discussed above). Therefore, an additional subsidy of z increases expenditures by $(m-3)z$. If $m \geq 3$ then it is optimal to set an additional subsidy of $z = a_{m-1} - a_m$. At this point, any additional subsidy must be given to player $m-1$ as well. Continuing in this way, we see that the optimal subsidy equates the head starts of players $k, k+1, \dots, m+1$ with that of player $k-1$, where $k \geq 2$ is the lowest integer that satisfies

$$m \geq (m+1-k) + 1 + m - k + 1 \iff m \geq 2m - 2k + 3 \iff k \geq \frac{m+3}{2}.^{36}$$

Design questions for which knowledge of players' equilibrium payoffs is insufficient may be addressed by applying the equilibrium construction algorithm. I illustrate this by briefly considering some design questions in a three-player all-pay auction with head starts and two prizes of value V . Players' equilibrium strategies are described in Section 4.1. Suppose that we are interested in increasing player 1's expenditures (and output) by subsidizing or handicapping player 2. If $a_1 > 2a_2 - a_3 - (a_2 - a_3)^2/V$, then Figure 6 describes the equilibrium as long as

³⁶It can be shown that in an $(m+1)$ -player all-pay auction with head starts players' payoffs equal their power and there is a unique equilibrium, given by the algorithm, even when Condition M3 does not hold. This continues to be true if additional players with head starts lower than a_{m+1} are added. These additional players do not participate.

the subsidy/handicap is not too large. Considering player 1's equilibrium CDF in this case, it is immediate that a handicap for player 2 has no effect, and a subsidy for player 2 reduces player 1's expenditures because it lowers the prizes' value. If Figure 7 describes the equilibrium, then it can be shown that a handicap for player 2 of up to $a_1 - a_2$ monotonically increases player 1's expenditures. This implies that a handicap of $a_1 - a_2$ is optimal. It can be shown, however, that if a_1 is small, then any subsidy to player 2 reduces player 1's expenditures.³⁷ Another object of possible interest is the first-order statistic of players' scores (this would correspond to the quality of the best innovation in an R&D setting). If Figure 6 describes the equilibrium, it can be shown that the first-order statistic is convex in a_1 . This implies that the optimal subsidy to player 1 is either 0 or the maximal subsidy (so that $a_1 + b_1 = V - \frac{b_1}{2}$). The former is optimal if a_1 is small, the latter is optimal if a_1 is large.

5 Conclusion

This paper has examined equilibrium behavior in a single-prize and multiprize contest model that accommodates various asymmetries, including head starts. When $m + 1$ players compete for m prizes, there is a unique equilibrium in which players do not choose weakly dominated scores. As long as no additional players participate, which is the case when weaker players' costs are everywhere higher than those of the marginal player, uniqueness is maintained. The equilibrium is constructed by an algorithm, and the construction shows that players may choose scores from several intervals, have atoms at positive scores, and have aggregate gaps in their best-response sets.

To study the effects of head starts, I applied the algorithm to all-pay auctions with head starts and identical valuations. These contests have a unique equilibrium, in which players choose scores from one or more competition zones separated by gaps. Within each competition zone weaker players are more aggressive than stronger players. Weaker players' expected expenditures are also higher than those of stronger players. The qualitative form of the equilibrium changes with the relative values of players' head starts, and is closely related to sequences that consist of m open parentheses and m close parentheses that are correctly matched, whose cardinality is the Catalan number. These equilibrium properties contrast with the equilibrium predictions of the standard all-pay auction, in which players differ in their valuations for a prize.

One direction for future research is to study all-pay auctions that combine head starts and differing valuations. The algorithm can be applied to these contests as long as the participation result holds (for example, when $a_i < a_{m+1}$ and $V_i < V_{m+1}$ for every player $i > m + 1$). The

³⁷It can also be shown that if Figure 6 describes the equilibrium, then a subsidy to player 1 increases his output: the decrease in expenditures due to the lower prize value is smaller than the increase in score due to the subsidy.

resulting equilibrium would likely combine some features of the standard all-pay auction with those of the all-pay auction with head starts and identical valuations. Classes of contests with non-linear costs can also be studied using the tools developed here. Contest design, which has been touched upon only briefly in this paper, is another direction for future research and application of the results. Because the equilibrium can be explicitly constructed for a large class of contests, many target functions are, at least in principle, amenable to analysis. Such an analysis could help a contest designer optimally employ the tools at his disposal. Finally, the contest framework can be used to study environments in which competitors' costs are non-monotonic, or in which competitors receive bonuses upon reaching certain milestones.

A Proofs of Proposition 1 and Results from Section 3

I denote by $\tilde{\mathbf{G}} = (\tilde{G}_1, \dots, \tilde{G}_{m+1})$ an equilibrium of an $(m+1)$ -player contest, where $\tilde{G}_i(x)$ is the probability that player i chooses a score lower than or equal to x . I say that player i is *stronger* (*weaker*) than player j if $i < j$ ($i > j$).

A.1 Proof of Proposition 1

Denote by \tilde{C} the contest in which player i 's costs are c_i^L and his valuation is V_i . I begin by proving that the first part of the proposition holds with \tilde{C} instead of \hat{C} . By definition of c_i^L , a score is weakly dominated for player i in \tilde{C} if and only if it is weakly (or strongly) dominated for player i in C . And if x is not weakly dominated for player i , then $c_i(x) = c_i^L(x)$. This proves the following claim.

Claim 1 *In any profile of strategies in which players do not choose weakly dominated scores, every player's payoff is the same in C and \hat{C} .*

Consider an equilibrium in C in which players do not choose weakly dominated scores. If this is not an equilibrium in \tilde{C} , then some player i has a profitable deviation x in \tilde{C} , which implies that he has a profitable deviation $y \geq x$ that is not weakly dominated. By Claim 1, y is a profitable deviation for for player i in C . The same argument proves the other direction. Claim 1 also shows that in any such equilibrium every player's payoff is the same in C and \tilde{C} . To conclude, note that \hat{C} is derived from \tilde{C} by adding the constant l_i to player i 's Bernoulli utility (adding $-l_i$ to his costs). This implies that \hat{C} and \tilde{C} are strategically equivalent, and also that for every strategy profile the payoff in \tilde{C} of player i is precisely l_i lower than his payoff in \hat{C} .

A.2 Proof of Lemma 1

The proof relies on the following lemma.

Lemma 4 *Suppose that in $\tilde{\mathbf{G}}$ no player chooses weakly dominated scores, and denote by s_{inf} the infimum of the union of the best response sets of players $1, \dots, m$. Then $\tilde{G}_{m+1}(s_{\text{inf}}) > 0$, $\tilde{G}_i(s_{\text{inf}}) = 0$ for every player $i < m+1$, and $s_{\text{inf}} = x_0$.*

Proof. By definition of s_{inf} , there exists a player $i < m+1$ and a set of scores $\{x_k\}_{k=1}^{\infty}$ that are best responses for player i such that $x_k \rightarrow s_{\text{inf}}$. Because player i 's power is positive, his probability of winning when choosing any score x_k in the set is bounded away from 0, so by right-continuity of players' CDFs there is a player $j \neq i$ with $\tilde{G}_j(s_{\text{inf}}) > 0$. Suppose in contradiction that $\tilde{G}_{m+1}(s_{\text{inf}}) = 0$. Then $j \neq m+1$, and because $\tilde{G}_j(x) = 0$ for every $x < s_{\text{inf}}$, player j has an atom at s_{inf} . The same argument for player j instead of player i shows that

there is another player, in addition to player j , with an atom at s_{inf} . Because $\tilde{G}_{m+1}(s_{\text{inf}}) = 0$, not all players with an atom at s_{inf} win a prize with probability 1 when choosing s_{inf} , so by the Tie Lemma of Siegel (2009) (henceforth: Tie Lemma) they win a prize with probability 0 and have a payoff of 0 when choosing s_{inf} . But players $1, \dots, m$ have a positive payoff, a contradiction. Thus, $\tilde{G}_{m+1}(s_{\text{inf}}) > 0$. Because $\tilde{G}_{m+1}(x) = 0$ for $x < x_0$, we have that $x_0 \leq s_{\text{inf}}$. If $x_0 = s_{\text{inf}}$, then the marginal player has an atom at s_{inf} , so by the Tie Lemma $\tilde{G}_i(s_{\text{inf}}) = 0$ for $i < m + 1$. Therefore, to complete the proof it suffices to rule out $x_0 < s_{\text{inf}}$. Suppose in contradiction that $x_0 < s_{\text{inf}}$. By choosing a score x in (x_0, s_{inf}) the marginal player wins a prize with probability 0 and incurs a positive cost, so no score in (x_0, s_{inf}) is a best response for him. Therefore, if $\tilde{G}_{m+1}(s_{\text{inf}}) > \tilde{G}_{m+1}(x_0)$ then the marginal player has an atom at s_{inf} , which implies that s_{inf} is a best reply for him. This means that another player has an atom at s_{inf} (because $c_{m+1}(s_{\text{inf}}) > c_{m+1}(x_0) = 0$), which contradicts the Tie Lemma. If $\tilde{G}_{m+1}(s_{\text{inf}}) = \tilde{G}_{m+1}(x_0) < 1$, consider a player $i < m + 1$ such that (i) the player has best responses that approach s_{inf} and (ii) no other player $j < m + 1$ has an atom at s_{inf} (the Tie Lemma guarantees that such a player i exists). Then s_{inf} is a best response for player i . But the CDFs of players other than i do not increase on $(x_0, s_{\text{inf}}]$, so all scores in (x_0, s_{inf}) are at least as good as s_{inf} for player i , contradicting the definition of s_{inf} . Finally, if $\tilde{G}_{m+1}(s_{\text{inf}}) = \tilde{G}_{m+1}(x_0) = 1$, then all scores in $(x_0, s_{\text{inf}}]$ are at least as good as s_{inf} for any player i who has best responses that approach s_{inf} , contradicting the definition of s_{inf} . ■

To complete the proof, it remains to show that $\tilde{G}_{m+1}(x_0) = \min_{i < m+1} \left(\frac{c_i(x_0) + w_i}{V_i} \right)$. If $\tilde{G}_{m+1}(x_0)$ were larger, some player $i < m + 1$ would obtain more than his power at scores immediately above x_0 . If $\tilde{G}_{m+1}(x_0)$ were smaller, then by Lemma 4 and right-continuity of $\tilde{\mathbf{G}}$ no player $i < m + 1$ would have best responses in $[x_0, x_0 + \varepsilon]$ for some $\varepsilon > 0$. This would contradict $s_{\text{inf}} = x_0$.

A.3 Proof of Lemma 2

Suppose in contradiction that $\tilde{G}_j(y) < 1$ for some player j at a score y that is not a best response for any player. This implies that no player has an atom at y , so by continuity of players' cost functions and CDFs at y , no player has best responses in some neighborhood of y . Denote by BR_y^k the set of player k 's best responses above y , and by $s_{\text{inf},y}$ the infimum of $\cup_{k=1}^{m+1} BR_y^k$. Because $\tilde{G}_j(y) < 1$, BR_y^j is not empty, so $s_{\text{inf},y} > y$ is well defined and no player has best responses in $[y, s_{\text{inf},y})$. If $s_{\text{inf},y}$ were not a best response for at least one player, then by the argument above no player would have best responses in a neighborhood of $s_{\text{inf},y}$, contradicting the definition of $s_{\text{inf},y}$. Consider a player k for whom $s_{\text{inf},y}$ is a best response. Because $c_k(s_{\text{inf},y}) \geq c_k(y)$ and y is not a best response for player k , he must win a prize with higher probability when choosing $s_{\text{inf},y}$ than when choosing y . This means that some player other than k has an atom at $s_{\text{inf},y}$, because no player has best responses in $[y, s_{\text{inf},y})$. This implies that $s_{\text{inf},y}$ is a best response for that other

player, which requires another player to have an atom at $s_{\text{inf},y}$. Therefore, there are at least two players with an atom at $s_{\text{inf},y}$. By the Tie Lemma of Siegel (2009), either all the players with an atom at $s_{\text{inf},y}$ win a prize with probability 0 or all win a prize with probability 1 when choosing $s_{\text{inf},y}$. If the former is true, then y is a best response for the players with an atom at $s_{\text{inf},y}$, a contradiction. If the latter is true, then denote by \mathcal{N}' the set of players l for whom $\tilde{G}_l(s_{\text{inf},y}) = 1$. This set is non-empty, otherwise no player could win a prize with probability 1 by choosing $s_{\text{inf},y}$. Moreover, every player in \mathcal{N}' has an atom at $s_{\text{inf},y}$, otherwise $\tilde{G}_l(y) = 1$ for a player l in \mathcal{N}' , so scores slightly below $s_{\text{inf},y}$ would be best-responses for every player with an atom at $s_{\text{inf},y}$, a contradiction. Now consider the positive-probability event that the $(m+1) - |\mathcal{N}'|$ players not in \mathcal{N}' choose scores strictly higher than $s_{\text{inf},y}$ and the $|\mathcal{N}'|$ players in \mathcal{N}' choose $s_{\text{inf},y}$.³⁸ Conditional on this event, $m - ((m+1) - |\mathcal{N}'|) = |\mathcal{N}'| - 1$ prizes are divided among the $|\mathcal{N}'|$ players in \mathcal{N}' , so the players in \mathcal{N}' cannot all win a prize with probability 1 when choosing $s_{\text{inf},y}$, a contradiction.

A.4 Proof of Lemma 3

The proof requires two lemmas.

Lemma 5 *Suppose that a checkpoint $x < T$ has been reached with $G_i(x) < 1$ for every player i , so Stage 2 is executed. If the algorithm proceeds to Stage 3, then the next checkpoint \bar{x} satisfies (i) $\bar{x} < T$ and (ii) $G_i(x) \leq G_i(\bar{x}) \leq 1$ for every player i .*

Proof. By definition of the checkpoint \bar{x} identified in Stage 3, $\mathcal{A}^+(x)$ contains at least one player i for whom $c_i(x) = c_i(\bar{x})$. Because player i is in $\mathcal{CP}(x)$, (2) holds. Because $G_j(x) < 1$ for every player j , (2) implies that $V_i - c_i(x) > w_i$. By definition, $w_i = V_i - c_i(T)$, so $c_i(x) < c_i(T)$ and consequently $c_i(\bar{x}) < c_i(T)$, which implies that (i) holds. For (ii), recall that Stage 3 specifies to set $G_j(\bar{x}) = G_j(x)$ for every player $j \neq i$. To see that $G_i(x) \leq G_i(\bar{x}) \leq 1$, note that $\bar{x} < T$ implies $V_j - c_j(\bar{x}) \geq w_j$ for every player j . Therefore, the lowest value of $G_i(\bar{x})$ such that a player other than i is in $\mathcal{CP}(\bar{x})$ is at most 1. Moreover, because for every player \tilde{i} either (1) or (2) holds with \tilde{i} instead of i , $c_{\tilde{i}}(\bar{x}) \geq c_{\tilde{i}}(x)$, and $G_j(\bar{x}) = G_j(x)$ for every player $j \neq i$, the lowest value of $G_i(\bar{x})$ such that a player other than i is in $\mathcal{CP}(\bar{x})$ is at least $G_i(x)$. ■

Lemma 6 *Suppose that a checkpoint $x < T$ has been reached with $G_i(x) < 1$ for every player i , so Stage 2 is executed. If the algorithm proceeds to Stage 4, then the next checkpoint \bar{x} satisfies (i) $\bar{x} \leq T$, (ii) G_i is non-decreasing on $[x, \bar{x}]$ and $G_i(\bar{x}) \leq 1$ for every player i , and (iii) if $\bar{x} = T$ then $G_i(\bar{x}) = 1$ for some player i .*

Proof. The extension of \mathbf{G} to $[x, \bar{x}]$ in Stage 4 is such that the hazard rate $-(1 - G_i(y))' / G_i(y)$ (where the derivative is a right-derivative) of every player i in $\mathcal{A}^+(x)$ at scores y immediately

³⁸If $|\mathcal{N}'| = m + 1$, then this event is simply the event that all players choose $s_{\text{inf},y}$.

above x equals $H(x, y) - \varepsilon_i(y) \geq 0$. This shows that \mathbf{G} is non-decreasing immediately above x . The checkpoint \bar{x} is chosen so that if player i 's hazard rate is non-negative immediately above x , then his hazard rate remains non-negative up to \bar{x} . Therefore, \mathbf{G} is non-decreasing in $[x, \bar{x}]$. Also, if $\bar{x} \leq T$ then $G_i(\bar{x}) \leq 1$ for every player i , because \bar{x} is chosen so that when \mathbf{G} is extended to $[x, \bar{x}]$ no player can obtain more than his power by choosing a score in $[x, \bar{x}]$. If $\bar{x} \geq T$, then the CDF of at least one player is 1 at T when \mathbf{G} is extended as specified in Stage 4, because for a player to obtain his power by choosing T he must win a prize with probability 1, and every player in $\mathcal{CP}(x)$ obtains his power in $[x, \bar{x}]$. Thus $\bar{x} \geq T$ implies $\bar{x} = T$. ■

To prove Lemma 3, first suppose that $G_{m+1}(x_0) = 1$. Then Stage 1 is followed by Stage 5. By part (ii) of Condition M3, $c_{m+1}(T) > c_{m+1}(0) = 0$, so $x^L = x_0 < T$. For the remainder of the proof, suppose that $G_{m+1}(x_0) < 1$, and recall that by definition $G_i(x_0) = 0$ for every player $i \neq m + 1$. Lemmas 5 and 6 suggest two possibilities. The first possibility is that $x^L \leq T$ with $G_i(x^L) = 1$ for some player i is reached after Stages 3 and 4 are executed a finite number of times. In this case, the statement of the lemma holds. The second possibility is that Stage 3 or 4 (or both) are executed infinitely many times. To complete the proof, it suffices to rule out this possibility. That Stage 3 is executed at most a finite number of times follows because the checkpoint identified in Stage 3 is the endpoint of a maximal interval on which a player's cost function is constant, and every player's cost function is constant on at most a finite number of such intervals because cost functions are piecewise analytic. If Stage 4 is executed an infinite number of times, then there exists a sequence of consecutive checkpoints $x_1 < x_2 < \dots$, all identified via iterations of Stage 4. By Lemmas 5 and 6, and because Stage 5 is not reached, the value of every player's CDF at every checkpoint in the sequence is less than 1, and $\lim_{k \rightarrow \infty} x_k \leq T$. The proof of Lemma 7 in Siegel (2010) shows that because cost functions are piecewise analytic there is only a finite number of checkpoints x_k in the sequence for which $\mathcal{A}^+(x_k) \neq \mathcal{A}^+(x_{k+1})$. This means that there is an infinite subsequence of consecutive checkpoints $x_{i_1} < x_{i_2} < \dots$ in some subinterval $[a, b]$ of $[0, T]$ for which $\mathcal{A}^+(x_{i_1}) = \mathcal{A}^+(x_{i_2}) = \dots$. This, together with (5) and the fact that the CDFs of players not in $\mathcal{A}^+(\cdot)$ do not increase, means that \mathbf{G} is piecewise analytic on $[a, b]$.³⁹ And because players' cost functions are piecewise analytic, the criteria specified in Stage 4 for identifying the next checkpoint mean that there can only be a finite number of checkpoints in $[a, b]$, a contradiction.

A.5 Proof of Theorem 3

To prove that \mathbf{G} is an equilibrium, it suffices to show three things: (i) \mathbf{G} is a profile of CDFs (non-decreasing, right-continuous, and $\lim_{x \rightarrow \infty} \mathbf{G}(x) = 1$), (ii) no player can obtain a payoff higher than his power by choosing any score when the other players choose scores according to

³⁹A proof similar to that of Lemma 4 in Siegel (2010) shows that \mathbf{G} is analytic at b .

\mathbf{G} , and (iii) every player i assigns G_i -measure 1 to scores that give him a payoff equal to his power when the other players choose scores according to \mathbf{G} . (i) follows from the construction of \mathbf{G} , Lemma 3, and part (ii) of Lemmas 5 and 6. For (ii), by construction of \mathbf{G} no player can obtain more than his power by choosing a score in $[0, x^L]$, where x^L is the last checkpoint identified by the algorithm. Also, by definition of power no player can obtain more than his power above T . It therefore suffices to show that no player can obtain more than his power by choosing a score in (x^L, T) . The following lemma shows that if $x^L < T$, then $G_{m+1}(x^L) = 1$ and $c_j(x^L) = c_j(T)$ for every player $j < m + 1$. This is enough: by Stage 5, the marginal player cannot increase his probability of winning a prize by increasing his chosen score from x^L to a score in (x^L, T) , and every other player obtains his power by choosing a score in (x^L, T) .

Lemma 7 *If $x^L < T$, then $G_{m+1}(x^L) = 1$ and $c_j(x^L) = c_j(T)$ for every player $j < m + 1$.*

Proof. By definition, $G_i(x^L) = 1$ for some player i , and x^L is identified in Stage 1, 3, or 4. If x^L is identified in Stage 1, then $x^L = x_0$ and

$$1 = G_{m+1}(x_0) = \min_{i < m+1} \left(\frac{c_i(x_0) + w_i}{V_i} \right) = \min_{i < m+1} \left(1 - \frac{(c_i(T) - c_i(x_0))}{V_i} \right),$$

which implies that $c_j(x^L) = c_j(T)$ for all $j < m + 1$. If x^L is identified in Stage 3 or 4, then consider a player i for which $G_i(x^L) = 1$. Player i is in $\mathcal{A}^+(x)$, where x is the checkpoint preceding x^L , and is therefore in $\mathcal{CP}(x^L)$. Suppose $i < m + 1$. In Stage 3 this means that the marginal player obtains strictly more than his power immediately above x^L (Assumption M3 part (ii)), so $G_i(x^L)$ can be set lower so that the marginal player is in $\mathcal{CP}(x^L)$. This contradicts the definition of $G_i(x^L)$ in Stage 3 as the lowest value such that $\mathcal{CP}(x^L)$ contains at least two players. In Stage 4, $i < m + 1$ means that the marginal player can obtain strictly more than his power immediately below x^L (by continuity of the extension of \mathbf{G} in Stage 4), contradicting the property of the extension of \mathbf{G} in Stage 4 that no player can obtain more than his payoff in $[x, x^L]$. Given that $G_{m+1}(x^L) = 1$, the same contradictions showed for the marginal player hold for any player $j < m + 1$ with $c_j(x^L) < c_j(T)$. ■

For (iii), note that a player's CDF increases continuously only in Stage 4, and \mathbf{G} is extended in Stage 4 in such a way that all players whose CDFs increase obtain their power (because they are all in $\mathcal{A}^+(x)$). It remains to show that a player obtains precisely his power whenever he has an atom. By construction, a player can be assigned an atom in Stage 1, 3, or 5. Stage 1 assigns an atom only to the marginal player, at x_0 , and the marginal player obtains a payoff of 0 (equal to his power) when choosing x_0 . Each execution of Stage 3 identifies a checkpoint \bar{x} and assigns an atom to at most one player, at \bar{x} . That player obtain his power at \bar{x} , because he is in $\mathcal{CP}(\bar{x})$ and there is no tie at \bar{x} (because no other player has an atom there). For Stage 5, consider two cases. If $x^L < T$, then Lemma 7 shows that Stage 5 assigns atoms to a subset of players $1, \dots, m$ at scores at which they win with probability 1 and whose cost equals their

cost of choosing the threshold, so they obtain their power by choosing these scores. If $x^L = T$, then x^L is identified in Stage 4 (part (i) of Lemma 5). Because the extension of \mathbf{G} in Stage 4 is continuous, the players whose CDF reaches 1 in Stage 4 at x^L do not have an atom at $x^L = T$. Thus, any player j with an atom at $x_j^{\max} \geq T$ wins a prize with probability 1 by choosing x_j^{\max} . And by Stage 5, $c_j(x_j^{\max}) = c_j(x^L) = c_j(T)$. Therefore player j obtains his power by choosing x_j^{\max} .

It remains to show that in \mathbf{G} players do not choose weakly dominated scores. Recall that a score x is weakly dominated for player i if $c_i(x) = c_i(y)$ for a score $y > x$. Consider the stages of the algorithm in which a player's CDF may increase. If a player's CDF increases in Stage 4, then his cost function is increasing (the cost functions of all candidate players in Stage 4 are increasing). If a player's CDF increases in Stages 1, 3, or 5, then he has an atom, and by construction that atom is at the upper endpoint of a maximal interval in which the player's cost function is constant. In both cases the player does not choose weakly dominated scores.

A.6 Proof of Theorem 4

Suppose that in $\tilde{\mathbf{G}}$ no player chooses weakly dominated scores. I will show that $\mathbf{G} = \tilde{\mathbf{G}}$. The following lemma shows that if \mathbf{G} and $\tilde{\mathbf{G}}$ differ, they differ at scores lower than x^L , the last checkpoint in \mathbf{G} .

Lemma 8 *If $\mathbf{G}(x^L) = \tilde{\mathbf{G}}(x^L)$, then $\mathbf{G}(y) = \tilde{\mathbf{G}}(y)$ for all $y > x^L$.*

Proof. Because $G_i(x^L) = \tilde{G}_i(x^L) = 1$ for some player i (which also implies $G_i(y) = \tilde{G}_i(y) = 1$ for all $y > x^L$), no other player j has best responses in \mathbf{G} or $\tilde{\mathbf{G}}$ at scores y for which $c_j(y) > c_j(x^L)$. Therefore, $\tilde{G}_j(y) = 1$ for every score $y \geq x_j^{\max}$, where x_j^{\max} is defined as in Stage 5. And because in $\tilde{\mathbf{G}}$ no player chooses weakly dominated scores, $\tilde{G}_j(y) = \tilde{G}_j(x^L)$ for every score y in (x^L, x_j^{\max}) . Consequently, $\mathbf{G}(y) = \tilde{\mathbf{G}}(y)$ for all $y > x^L$. ■

For the next step, divide the checkpoints in \mathbf{G} , excluding x^L , into “constant checkpoints” and “increasing checkpoints.” A constant checkpoint leads to Stage 3, i.e., it is a checkpoint at which at least one active player's cost function is constant. An increasing checkpoint leads to Stage 4, i.e., it is a checkpoint at which all active players' cost functions are increasing. The next lemma shows that if \mathbf{G} and $\tilde{\mathbf{G}}$ coincide at a constant checkpoint, then they coincide at all scores up to the following checkpoint. Denote by x_k and x_{k+1} two consecutive checkpoints.

Lemma 9 *If x_k is a constant checkpoint and $\mathbf{G}(x_k) = \tilde{\mathbf{G}}(x_k)$, then $\mathbf{G}(y) = \tilde{\mathbf{G}}(y)$ for all y in $[x_k, x_{k+1}]$.*

Proof. By construction, because x_k is a constant checkpoint, there is a player i in $\mathcal{CP}(x_k)$ for which $c_i(x_k) = c_i(x_{k+1})$. This means that $\tilde{G}_j(y) = \tilde{G}_j(x_k)$ for every other player j and

y in $[x_k, x_{k+1}]$, otherwise player i would obtain more than his power in $\tilde{\mathbf{G}}$. Now, because no player chooses weakly dominated scores in $\tilde{\mathbf{G}}$, $\tilde{G}_i(y) = \tilde{G}_i(x_k) = G_i(x_k) = G_i(y)$ for every y in $[x_k, x_{k+1}]$. It remains to show that $G_i(x_{k+1}) = \tilde{G}_i(x_{k+1})$. This is true because $G_i(x_{k+1}) < \tilde{G}_i(x_{k+1})$ implies that some player other than i obtains more than his power immediately above x_{k+1} , and $G_i(x_{k+1}) > \tilde{G}_i(x_{k+1})$ implies that Lemma 2 is violated immediately above x_{k+1} , as explained in the paragraph preceding the statement of Lemma 2. ■

Denote by $\tilde{\mathcal{CP}}(x)$ the set of players i for which (2) holds with \tilde{G}_i instead of G_i (the players for whom x is a best response in $\tilde{\mathbf{G}}$).

Lemma 10 *Suppose that $x_k < x^L$ is an increasing checkpoint. If $\tilde{\mathbf{G}}(x_k) = \mathbf{G}(x_k)$ and $\tilde{\mathcal{CP}}(x) = \mathcal{CP}(x)$ for every x in (x_k, x_{k+1}) , then $\tilde{\mathbf{G}}(x) = \mathbf{G}(x)$ for every $x \in [x_k, x_{k+1}]$.*

Proof. By construction, $\mathcal{CP}(x) = \mathcal{A}^+(x_k)$ for every x in (x_k, x_{k+1}) . This implies that for every player i not in $\mathcal{A}^+(x_k)$ and every x in (x_k, x_{k+1}) we have $\tilde{G}_i(x) = \tilde{G}_i(x_k) = G_i(x_k) = G_i(x)$, because player i does not have best responses in (x_k, x_{k+1}) so his CDF does not increase (in both equilibria). For every player i in $\mathcal{A}^+(x_k)$ and every x in (x_k, x_{k+1}) (2) holds with \tilde{G}_i instead of G_i , so \tilde{G}_i is given by (5). This shows that $\tilde{\mathbf{G}}(x) = \mathbf{G}(x)$ for every $x \in [x_k, x_{k+1}]$. To see that $\tilde{\mathbf{G}}(x_{k+1}) = \mathbf{G}(x_{k+1})$, suppose first that $x_{k+1} < x^L$. In this case, by construction \mathbf{G} is continuous at x_{k+1} , so the same must be true for $\tilde{\mathbf{G}}$, otherwise a player in $\mathcal{A}^+(x_k)$ would get in $\tilde{\mathbf{G}}$ more than his power. Therefore, $\tilde{\mathbf{G}}(x_{k+1}) = \mathbf{G}(x_{k+1})$. Now suppose that $x_{k+1} = x^L$. By construction, $G_i(x^L) = 1$ and G_i is continuous at x^L for some player i , so the same is true for $\tilde{\mathbf{G}}$. Consider another player j . Because in both $\tilde{\mathbf{G}}$ and \mathbf{G} players do not choose weakly dominated scores, if player j 's cost function is constant at x^L , then he does not have an atom at x^L , so $\tilde{G}_j(x^L) = \lim_{x \rightarrow x^L} \tilde{G}_j(x) = \lim_{x \rightarrow x^L} G_j(x) = G_j(x^L)$. If player j 's cost function is increasing at x^L , then in both $\tilde{\mathbf{G}}$ and \mathbf{G} he does not have best responses above x^L (he wins with probability 1 by choosing x^L), so $\tilde{G}_j(x^L) = G_j(x^L) = 1$. ■

Suppose that $\mathbf{G} \neq \tilde{\mathbf{G}}$, and denote by x_k the highest checkpoint such that $\mathbf{G}(y) = \tilde{\mathbf{G}}(y)$ for every y in $[0, x_k]$ (Lemmas 1 and 8 show that x_k is well defined and $x_k < x^L$). Lemma 9 shows that x_k is an increasing checkpoint. Lemma 10 shows that $\tilde{\mathcal{CP}}(x) \neq \mathcal{CP}(x)$ for some x in (x_k, x_{k+1}) . I conclude the proof by showing that $\tilde{\mathcal{CP}}(x) \neq \mathcal{CP}(x)$ is impossible. That $\tilde{\mathcal{CP}}(x) \subseteq \mathcal{CP}(x)$ follows from Lemma 9 in Siegel (2010), which applies here.⁴⁰ This inclusion means that in both equilibria only the CDFs of players in $\mathcal{A}^+(x_k) = \mathcal{CP}(x)$ increase on (x_k, x_{k+1}) . Because x_k is an increasing checkpoint, the costs of all players in $\mathcal{A}^+(x_k)$ are increasing at every x in (x_k, x_{k+1}) . Therefore, in both equilibria the CDFs of these players are continuous on (x_k, x_{k+1}) .⁴¹ Lemmas 10 and 11 in Siegel (2010) now imply that $\mathcal{CP}(x) \subseteq \tilde{\mathcal{CP}}(x)$. Therefore, $\mathbf{G} = \tilde{\mathbf{G}}$.

⁴⁰In the proof of the lemma, replace $v_i(x)$ with $V_i - c_i(x)$ for every player i , and note that here we have $\bar{G}(x_k) = \mathbf{G}(x_k)$ by definition of x_k .

⁴¹If a player had an atom in (x_k, x_{k+1}) , then because no player's CDF reaches 1 before x_{k+1} , no other player would choose scores immediately below x_k . But then the player would be better off by shifting the mass of his

A.7 Proof of Theorem 5

Suppose that all players behave as specified in the statement of the theorem. Then, players $1, \dots, m+1$ do not have profitable deviations, because their probability of winning at any score higher than x_0 coincides with their probability of winning at that score in the reduced contest with players $1, \dots, m+1$, and their probability of winning at any score lower than or equal to x_0 is at most their probability of winning at that score in the reduced contest. This is because, by assumption, players $m+2, \dots, n$ do not choose scores higher than x_0 . Players $m+2, \dots, n$ do not have profitable deviations, because the proof of Theorem 2 in Siegel (2009) shows that if a player $i > m+1$ could obtain a positive payoff when players $1, \dots, m+1$ play according to the equilibrium constructed by the algorithm, then the marginal player could do the same, and this deviation in the reduced contest would contradict Theorem 1 applied to the reduced contest. This shows that the behavior specified in the statement of the theorem is an equilibrium. For uniqueness, consider an equilibrium in which players do not choose weakly dominated scores. Because players $m+2, \dots, n$ do not participate and do not choose weakly dominated scores, they behave as in the statement of the theorem. In particular, they do not choose scores above x_0 . But now the same arguments used to describe the algorithm show that the only behavior of players $1, \dots, m+1$ consistent with equilibrium is the one specified by the algorithm (Lemma 1 holds, and above x_0 a player in $\{1, \dots, m+1\}$ wins a prize if and only if his score is higher than at least one other player in $\{1, \dots, m+1\}$).

A.8 Lemmas 11 and 12

Lemma 11 *At every checkpoint x that has been reached, $\mathcal{CP}(x)$ contains at least two players, and (1) holds for every player i not in $\mathcal{CP}(x)$.*

Proof. The proof is by induction on the number of checkpoints that have been reached. The first checkpoint is x_0 , and $\mathbf{G}(x_0)$ is set to the lowest value such that $\mathcal{CP}(x_0)$ contains player $m+1$ and at least one other player, so for every player i not in $\mathcal{CP}(x_0)$ (1) holds. For the induction step, consider two consecutive checkpoints $x < \bar{x}$. By the induction hypothesis, $\mathcal{CP}(x)$ contains at least two players, and for every player i not in $\mathcal{CP}(x)$ (1) holds. The first possibility is that $\mathcal{CP}(x)$ contains at least one player i whose cost function is constant at x . Then \bar{x} is identified in Stage 3, and $\mathbf{G}(\bar{x})$ is set to the lowest value such that $\mathcal{CP}(\bar{x})$ contains player i and at least one other player, so for every player i not in $\mathcal{CP}(\bar{x})$ (1) holds with \bar{x} instead of x . The second possibility is that the cost functions of all players in $\mathcal{CP}(x)$ are increasing at x . Then \bar{x} is identified in Stage 4, and \mathbf{G} is defined on $[x, \bar{x}]$ so that $\mathcal{A}^+(x) \subseteq \mathcal{CP}(\bar{x})$, and by definition of \bar{x} for every player i not in $\mathcal{CP}(\bar{x})$ (1) holds with \bar{x} instead of x . Because $\mathcal{A}^+(x)$ identified in Stage 4 contains at least two players, we are done. ■

atom to a lower score. The proof of Lemma 1 in Siegel (2009) provides a more detailed argument.

Lemma 12 *If c_i is increasing at $x < T$, then $q_i(y) > 0$ and $\varepsilon_i(y) > 0$ for every y in some open right-neighborhood of x .*

Proof. Because $q_i(y) = 1 - \frac{w_i + c_i(y)}{V_i}$, for the first inequality it suffices to show that $w_i + c_i(y) < V_i$, or $c_i(y) < c_i(T)$. Because c_i is increasing at x , $c_i(x) < c_i(T)$, so by continuity the same is true for scores y immediately above x . For the second inequality, because $\varepsilon_i(y) = \frac{c'_i(y)}{V_i - w_i - c_i(y)}$ and we have seen that $w_i + c_i(y) < V_i$, it suffices to show that $c'_i(y) > 0$. If this were not true in some right-neighborhood of x , then by analyticity c_i would be constant on some right-neighborhood of x , contradicting the fact that c_i is increasing at x . ■

B Derivation of the Equilibrium for the Three-Player, Two-Prize All-Pay Auction with Head Starts

Players' powers are $w_1 = a_1 - a_3$, $w_2 = a_2 - a_3$, and $w_3 = 0$. Stage 1 specifies that $x_0 = a_3$ and $G_3(a_3) = (a_2 - a_3)/V$. Stage 2 shows that $\mathcal{CP}(a_3) = \{2, 3\}$. Because player 2's costs are constant at a_3 , the algorithm proceeds to Stage 3, which gives $\mathcal{A}^+(a_3) = \{2\}$. Therefore, the first checkpoint above a_3 is a_2 , and no player's CDF increases in (a_3, a_2) . From (3) we have

$$\begin{aligned} G_2(a_2) &= \min \left\{ 1 - \frac{1 - \frac{w_1 + c_1(a_2)}{V}}{1 - G_3(a_3)}, 1 - \frac{1 - \frac{w_3 + c_3(a_2)}{V}}{1 - G_1(a_3)} \right\} \\ &= \min \left\{ \frac{a_1 - a_2}{a_3 + V - a_2}, \frac{a_2 - a_3}{V} \right\} = \begin{cases} \frac{a_2 - a_3}{V} & \text{if } a_1 \geq 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \\ \frac{a_1 - a_2}{a_3 + V - a_2} & \text{if } a_1 < 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \end{cases}. \end{aligned}$$

This implies that

$$\mathcal{CP}(a_2) = \begin{cases} \{2, 3\} & \text{if } a_1 > 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \\ \{1, 2, 3\} & \text{if } a_1 = 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \\ \{1, 2\} & \text{if } a_1 < 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} \end{cases}.$$

Case 1, $a_1 > 2a_2 - a_3 - (a_2 - a_3)^2/V$: Because the cost functions of players 2 and 3 are increasing at a_2 , the algorithm proceeds to Stage 4. Because $\mathcal{A}^+(a_2) = \mathcal{CP}(a_2) = \{2, 3\}$, from (5) we have that $G_2(y) = 1 - q_3(y) = (y - a_3)/V$ and $G_3(y) = 1 - q_2(y) = (y - a_3)/V$ for scores y in $[a_2, \bar{x}]$, where the checkpoint \bar{x} is the first score at which player 1 obtains his power:

$$\left(1 - \left(1 - \frac{\bar{x} - a_3}{V} \right)^2 \right) V = a_1 - a_3 \Rightarrow \bar{x} = a_3 + V \left(1 - \sqrt{1 - \frac{a_1 - a_3}{V}} \right) > a_2.^{42}$$

Proceeding to Stage 2, we have that $\mathcal{CP}(\bar{x}) = \{1, 2, 3\}$, and because the cost function of player 1 is constant at \bar{x} , the algorithm proceeds to Stage 3, which gives $\mathcal{A}^+(\bar{x}) = \{1\}$. Therefore, the

next checkpoint above $a_3 + V \left(1 - \sqrt{1 - (a_1 - a_3)/V}\right)$ is a_1 and no player's CDF increases in $(a_3 + V \left(1 - \sqrt{1 - (a_1 - a_3)/V}\right), a_1)$. From (3) we have

$$\begin{aligned} G_1(a_1) &= \min \left\{ 1 - \frac{1 - \frac{w_2 + c_2(a_1)}{V}}{1 - G_3\left(a_3 + V \left(1 - \sqrt{1 - \frac{a_1 - a_3}{V}}\right)\right)}, 1 - \frac{1 - \frac{w_3 + c_3(a_1)}{V}}{1 - G_2\left(a_3 + V \left(1 - \sqrt{1 - \frac{a_1 - a_3}{V}}\right)\right)} \right\} \\ &= \min \left\{ 1 - \sqrt{1 - \frac{a_1 - a_3}{V}}, 1 - \sqrt{1 - \frac{a_1 - a_3}{V}} \right\} = 1 - \sqrt{1 - \frac{a_1 - a_3}{V}}. \end{aligned}$$

Proceeding to Stage 2, we have that $\mathcal{CP}(a_1) = \{1, 2, 3\}$, and because all players' costs are increasing at a_1 the algorithm proceeds to Stage 4. Because $\mathcal{A}^+(a_1) = \mathcal{CP}(a_1) = \{1, 2, 3\}$, from (5) we have

$$G_i(y) = 1 - \frac{\prod_{j \in \{1, 2, 3\}} q_j(y)^{\frac{1}{|\{1, 2, 3\}| - 1}}}{q_i(y)} = 1 - \frac{\left(\frac{V + a_3 - y}{V}\right)^{\frac{3}{2}}}{\frac{V + a_3 - y}{V}} = 1 - \sqrt{1 - \frac{y - a_3}{V}} \quad (6)$$

for every player i and score y in $[a_1, \bar{x}]$, where the checkpoint $\bar{x} = a_3 + V$ is the score at which players' CDFs reach 1.

Case 2, $a_1 = 2a_2 - a_3 - (a_2 - a_3)^2/V$: Because the cost function of player 1 is constant at a_2 , the algorithm proceeds to stage 3, which gives $\mathcal{A}^+(a_2) = \{1\}$. Therefore, the next checkpoint above a_2 is a_1 and no player's CDF increases in (a_2, a_1) . The equilibrium starting from a_1 is as in Case 1: $G_i(y) = 1 - \sqrt{1 - (y - a_3)/V}$ for every player i and score y in $[a_1, a_3 + V]$. All players' CDF reach 1 at $a_3 + V$.

Case 3, $a_1 < 2a_2 - a_3 - (a_2 - a_3)^2/V$: Because the cost function of player 1 is constant at a_2 , the algorithm proceeds to stage 3, which gives $\mathcal{A}^+(a_2) = \{1\}$. Therefore, the next checkpoint above a_2 is a_1 , and no player's CDF increases in (a_2, a_1) . From (3) we have

$$G_1(a_1) = \min \left\{ 1 - \frac{1 - \frac{w_2 + c_2(a_1)}{V}}{1 - G_3(a_2)}, 1 - \frac{1 - \frac{w_3 + c_3(a_1)}{V}}{1 - G_2(a_2)} \right\} = \min \left\{ \frac{a_1 - a_2}{a_3 + V - a_2}, \frac{a_2 - a_3}{V} \right\} = \frac{a_1 - a_2}{a_3 + V - a_2}.$$

Proceeding to Stage 2, we have that $\mathcal{CP}(a_1) = \{1, 2\}$, and because both players' costs are increasing at a_1 the algorithm proceeds to Stage 4. Because $\mathcal{A}^+(a_1) = \mathcal{CP}(a_1) = \{1, 2\}$, from (5) we have

$$G_1(y) = 1 - \frac{q_1(y) q_2(y)}{q_1(y) (1 - G_3(a_1))} = \frac{y - a_2}{a_3 + V - a_2} = 1 - \frac{q_1(y) q_2(y)}{q_2(y) (1 - G_3(a_1))} = G_2(y)$$

for scores y in $[a_1, \bar{x}]$, where the checkpoint \bar{x} is the first score at which player 3 obtains his power:

$$\left(1 - \left(1 - \frac{\bar{x} - a_2}{a_3 + V - a_2}\right)^2\right) V - (\bar{x} - a_3) = 0 \Rightarrow \bar{x} = 2a_2 - a_3 - \frac{(a_2 - a_3)^2}{V} > a_1.$$

Proceeding to Stage 2, we have that $\mathcal{CP}(\bar{x}) = \{1, 2, 3\}$, and because all players' costs are increasing at \bar{x} the algorithm proceeds to Stage 4. Because $\mathcal{A}^+(\bar{x}) = \mathcal{CP}(\bar{x}) = \{1, 2, 3\}$, (6) tells us that $G_i(y) = 1 - \sqrt{1 - (y - a_3)/V}$ for every player i and score y in $[\bar{x}, a_3 + V]$. All players' CDFs reach 1 at $a_3 + V$.

C Proofs of Results from Section 4

To simplify the exposition, I assume that $a_{m+1} = 0$ and $V = 1$, and that all the a_i s are distinct. Relaxing these assumptions does not change the results or the logic of the proofs, but complicates the notation and exposition. One implication of these assumptions is that every player i 's power is a_i .

The following lemmas describe properties of \mathbf{G} used in the proofs below.

Lemma 13 *If i is in $\mathcal{CP}(x)$ for some x in $[a_i, 1)$, then for all k such that $x \geq a_k$ we have $G_i(x) \leq G_k(x)$, and if $G_i(x) = G_k(x)$, then k is in $\mathcal{CP}(x)$. Therefore, (i) if players $j > i$ are in $\mathcal{CP}(x)$ for some x in $[a_i, 1)$, then $G_i(x) = G_j(x)$, and (ii) if $G_i(x) = G_j(x)$ for players $j > i$ and x is in $[a_i, 1)$, then i is in $\mathcal{CP}(x)$ if and only if j is in $\mathcal{CP}(x)$.*

Proof. Because $x < 1$, $G_l(x) < 1$ for every player l (otherwise every other player would obtain more than his power by choosing scores slightly above x). From (2) we have

$$(1 - \prod_{l \neq i} (1 - G_l(x))) - (x - a_i) = a_i \Rightarrow (1 - \prod_{l \neq i} (1 - G_l(x))) = x.$$

If $G_i(x) > G_k(x)$, then

$$(1 - \prod_{l \neq k} (1 - G_l(x))) > x \text{ and } (1 - \prod_{l \neq k} (1 - G_l(x))) - (x - a_k) > a_k,$$

so if $x \geq a_k$ player k can obtain more than his power immediately above x . And if $G_i(x) = G_k(x)$, then $(1 - \prod_{l \neq k} (1 - G_l(x))) - (x - a_k) = a_k$ so k is in $\mathcal{CP}(x)$. ■

Lemma 14 *If the algorithm proceeds to Stage 4 at a checkpoint x , then $\mathcal{A}^+(x) = \mathcal{CP}(x)$.*

Proof. Proceeding to Stage 4 means that for every player i in $\mathcal{CP}(x)$ the costs are increasing at x so $x \geq a_i$. And for every player i and every score $y \geq a_i$, we have that

$$\varepsilon_i(y) = \frac{c'_i(y)}{V_i - w_i - c_i(y)} = \frac{1}{1 - a_i - (y - a_i)} = \frac{1}{1 - y}.$$

That ε_i for every player i in $\mathcal{CP}(x)$ does not depend on i implies that either all players in $\mathcal{CP}(x)$ are in $\mathcal{A}^+(x)$ or none of them are. Because $\mathcal{A}^+(x) \subseteq \mathcal{CP}(x)$ and $\mathcal{A}^+(x)$ contains at least two players, $\mathcal{A}^+(x) = \mathcal{CP}(x)$. ■

C.1 Proof of Proposition 2

I first show by reverse induction on $k = m, \dots, 1$ that for every two players $j > i$ and every two scores $y > x$ in $[a_{k+1}, a_k]$ (a) $G_i(x) \leq G_j(x)$, (b) if $G_i(x) = G_j(x)$ and $x \geq a_i$ then $G_i(y) = G_j(y)$, and (c) every player $k' \leq k$ does not have best responses in $[0, a_{k+1})$, and there is some h_{k+1} in $[a_{k+1}, a_k)$ such that every score in (h_{k+1}, a_k) is a best response only for player k .

By Stage 1 and Stage 3 (of the algorithm), for every x in $[0, a_m)$ we have $G_{m+1}(x) = a_m$ and $G_i(x) = 0$ for every player $i < m + 1$. Therefore, (a) and (b) hold for $k = m$. Every score in $(0, a_m)$ is a best response only for player m , so (c) holds.

Suppose that the claim holds for k, \dots, m . By (c), some x in (a_{k+1}, a_k) is a best response for player k , so k is in $\mathcal{CP}(x)$. Because player k 's costs are 0 up to a_k , at x Stage 3 is executed. Because no player $k' < k$ has best responses below a_k (by (c)) and the costs of every player $k' > k$ are increasing at every score above x , the checkpoint following x is a_k and k is in $\mathcal{CP}(a_k)$. By Stage 3, $G_i(a_k) = G_i(x)$ for every player $i \neq k$. Also by Stage 3, $G_k(a_k)$ is set so that at least one player other than k is in $\mathcal{CP}(a_k)$.

Suppose first that $\mathcal{CP}(a_k)$ contains a player $k' < k$. Then $k' = k - 1$, because for every $i < k - 1$ we have $w_{k-1} < w_i$ and $c_i(a_k) = c_{k-1}(a_k) = 0$. Now, because player $k - 1$ has constant costs at a_k , Stage 3 is executed at a_k and no player's CDF increases in (a_k, a_{k-1}) . This implies (b) (at a_{k-1} no player $l > k - 1$ has an atom) and (c) for $h_k = a_k$. To complete the induction step in this case, it remains to show that $G_k(a_k) \leq G_i(a_k)$ for every player $i > k$. But this is immediate from Lemma 13, because k is in $\mathcal{CP}(a_k)$.

Now suppose that $\mathcal{CP}(a_k)$ does not contain any player $k' < k$. Then, by Lemma 13 the CDFs at a_k of all players in $\mathcal{CP}(a_k)$ coincide, and are lower than the CDF of any player $k' > k$ not in $\mathcal{CP}(a_k)$. Because $\mathcal{CP}(a_k)$ does not contain players $k' < k$, Stage 4 is executed at a_k , and by Lemma 14 we have $\mathcal{A}^+(a_k) = \mathcal{CP}(a_k)$. Because $\mathcal{CP}(z) = \mathcal{A}^+(z) = \mathcal{A}^+(a_k)$ for every z in (a_k, y) , where y is the checkpoint following a_k , Lemma 13 shows that the CDFs of all players in $\mathcal{A}^+(a_k)$ coincide on $(a_k, y]$. It cannot be that $y \geq a_{k-1}$, otherwise an argument similar to the one used in the proof of Lemma 13 applied to scores slightly below a_{k-1} would show that player $k - 1$ can obtain more than his power by choosing a_{k-1} . Because $y < a_{k-1}$ and the hazard rates at y of all players in $\mathcal{A}^+(a_k)$ are identical, and therefore positive,⁴³ at y another player l becomes active (l is in $\mathcal{CP}(y) \setminus \mathcal{A}^+(a_k)$). If $l > k$, by Lemma 13 the CDFs of the players in $\mathcal{A}^+(a_k)$ at y coincide with $G_l(y)$, and y is the first score above a_k at which the CDFs of the players in $\mathcal{A}^+(a_k)$ coincide with the CDF of any player $k' > k$. Therefore, $G_l(y)$ is strictly lower than the CDF at y of every player $k' > k$ not in $\mathcal{CP}(y)$. Stage 4 is then executed at y , $\mathcal{A}^+(y) = \mathcal{CP}(y)$, and the CDFs of all players in $\mathcal{A}^+(y)$ coincide up to the checkpoint following y . Continuing in this way, it cannot be that only players $k' > k$ become active before a_{k-1} is reached, otherwise player $k - 1$ could obtain more than his power by choosing a_{k-1} . At the first checkpoint above

⁴³The hazard rate of player i at y is $H(a_k, y) - \varepsilon_i(y)$.

a_k at which a player $k' < k$ becomes active, player $k - 1$ becomes active (because for every $i < k - 1$ we have $w_{k-1} < w_i$ and $c_i(a_k) = c_{k-1}(a_k) = 0$). Denote this checkpoint by $h_k < a_{k-1}$. No score in $[a_k, h_k]$ is a best response for any player $k' < k - 1$, otherwise such a player would have become active earlier. At h_k player $k - 1$ has constant costs, so Stage 3 is executed and no player's CDF increases up to a_{k-1} (and at a_{k-1} no player $l > k - 1$ has an atom). Therefore, (a), (b), and (c) hold. This completes the induction step.

To complete the proof it suffices to show that for every two players $j > i$ and every two scores $y > x$ in $[a_1, 1)$ (a) and (b) hold. This can be done by following the analysis above of the case in which $\mathcal{CP}(a_k)$ does not contain any player $k' < k$, and noting that at every checkpoint up to 1 additional players become active.

Corollary 4 *If i is in $\mathcal{CP}(x)$ for some x in $[a_i, 1)$, then for all $k < i$ such that $x \geq a_k$ we have that (a) k is in $\mathcal{CP}(x)$ and (b) $G_i(y) = G_k(y)$ for every $y \geq x$.*

Proof. Lemma 13 shows that $G_i(x) \leq G_k(x)$, and Proposition 2 shows that $G_k(x) \leq G_i(x)$. Lemma 13 then shows (a), and Proposition 2 shows (b). ■

C.2 Proof of Corollary 2

Denote by $P_k(x)$ the probability that player k wins a prize when he chooses x and the other players choose scores according to \mathbf{G} , and by P_k the probability that player k wins a prize when all players choose scores according to \mathbf{G} . To show that $P_i \geq P_j$, note that if player i chooses $x \neq a_j$, then player j chooses a lower score with probability $G_j(x)$, and if j chooses $x \neq a_i$, then i chooses a lower score with probability $G_i(x)$. Therefore, because by Proposition 2 $G_i(x) \leq G_j(x)$, for any $x \neq a_i, a_j$ we have $P_i(x) \geq P_j(x)$. Moreover, because $G_i(a_i) \leq G_j(a_i)$ and players other than i do not have an atom at a_i , we have $P_i(a_i) \geq P_j(a_i)$ regardless of the tie-breaking rule. Together with the fact that i chooses a_j with probability 0, this implies that $P_i = \int P_i(x) dG_i \geq \int P_j(x) dG_i$, and because $P_j(\cdot)$ is non-decreasing, by FOSD we have $\int P_j(x) dG_i \geq \int P_j(x) dG_j = P_j$.

C.3 Proof of Theorem 6

Part (a): Part (a) of Corollary 4 shows that if $l > i$ is in \mathcal{A}_i , then so are $i, \dots, l - 1$. Therefore, if $m + 1$ is in \mathcal{A}_1 , then $\mathcal{A}_1 = \{1, 2, \dots, m + 1\}$. Suppose $m + 1$ is not in \mathcal{A}_1 , so G_{m+1} does not increase on $[a_1, 1)$. Because $G_{m+1}(a_1) < 1$ (otherwise player 1 would obtain more than his power by choosing scores slightly above a_1) and at every score up to 1 at least one player is active (Lemma 2), $\lim_{y \rightarrow 1} G_i(y) = 1$ for some player i (otherwise no player would obtain his power at scores immediately below 1). But this contradicts Proposition 2. To complete the proof of (i) it remains to show that $\mathcal{A}_i = \{i\}$ is impossible. This impossibility obtains because player i 's costs

are strictly increasing in (a_i, a_{i-1}) , so for a score x in (a_i, a_{i-1}) to be a best-response for player i , his probability of winning when choosing x must be higher than his probability of winning when choosing a lower score, for example $a_i + (x - a_i)/2$. This requires at least one other player to choose scores in $[a_i + (x - a_i)/2, x) \subseteq (a_i, a_{i-1})$ with positive probability. But if $\mathcal{A}_i = \{i\}$, then no player $i + 1, \dots, m + 1$ chooses scores in (a_i, a_{i-1}) with positive probability. And players $1, \dots, i - 1$ do not choose scores in (a_i, a_{i-1}) with positive probability, because such scores are weakly dominated by their respective head starts. For (ii), suppose that $\mathcal{A}_i \cap \mathcal{A}_{i'} \neq \emptyset$. To show that $\mathcal{A}_i \subseteq \mathcal{A}_{i'}$, by (i) it suffices to show that k_i is in $\mathcal{A}_{i'}$. Consider some j in $\mathcal{A}_i \cap \mathcal{A}_{i'}$. Because k_i is in \mathcal{A}_i and $j \leq k_i$, part (b) of Corollary 4 shows that there is a score x in (a_i, a_{i-1}) such that $G_j(y) = G_{k_i}(y)$ for every $y > x$. But then Lemma 13 shows that player k_i is active at any score y in (a_h, a_{h-1}) at which player j is active, so k_i is in $\mathcal{A}_{i'}$ because j is in $\mathcal{A}_{i'}$.

Part (b): Take a player j in \mathcal{A}_i , and consider the induction step for $k = i$ in the proof of Proposition 2. Because player $j \geq i$ has best responses in (a_i, a_{i-1}) (where $a_0 = 1$), the induction step shows that $\mathcal{CP}(a_i)$ does not contain any player $i' < i$ (otherwise only player $i - 1$ would have best responses in (a_i, a_{i-1})). And if $\mathcal{CP}(a_i)$ does not contain any player $i' < i$, then the induction step identifies a score $h_i \leq a_{i-1}$ such that if a player is active at a score in $[a_i, h_i)$ then the player remains active until h_i , and no player $i' \geq i$ has best responses in (h_i, a_{i-1}) . Moreover, $h_i < a_{i-1}$ if $i > 1$ (h_i is the score at which player $i - 1$ becomes active), and $h_1 = 1$. Therefore, the set of best responses in (a_i, a_{i-1}) of every player j in \mathcal{A}_i is an interval with upper bound h_i . That $G_j(x) > G_j(y)$ for any two scores $x > y$ in $(l_{i,j}, h_i)$ follows from the fact that G_j , given by (5), strictly increases in some right-neighborhood of y (where \mathcal{A}^+ does not change). For players $k > j$ in \mathcal{A}_i , part (a) of Corollary 4 implies that $l_{i,k} \geq l_{i,j}$, and part (a) of Theorem 6, part (b) of Corollary 4, and part (ii) of Lemma 13 together imply that $l_{i',j} = l_{i',k}$ for every $i' < i$ for which $\mathcal{A}_i \cap \mathcal{A}_{i'} \neq \emptyset$. To show that $l_{i,i} = l_{i,i+1} = a_i$ if $\mathcal{A}_i \neq \emptyset$, recall that Stage 3 at h_{i+1} (where $h_{m+1} = a_{m+1}$) shows that $\mathcal{CP}(a_i)$ contains a player j other than i , and the induction step in the proof of Proposition 2 shows that $j > i$. This implies, by Corollary 4, that i and $i + 1$ are in $\mathcal{CP}(a_i)$.

Part (c): Immediate from part (b) of Corollary 4 and right-continuity of CDFs at $l_{i,k}$.

Part (d): By part (a), $m + 1$ is in \mathcal{A}_1 . Therefore, part (b) of Corollary 4 shows that the CDFs of players $1, \dots, m + 1$ coincide on $[x, 1]$ for some x in $(a_1, 1)$. Because at least one player is active at every score up to 1 (Lemma 2), $\lim_{x \rightarrow 1} G_i(x) = 1$ for every player i , otherwise no player $1, \dots, m + 1$ could obtain his power at scores immediately below 1.

Part (e): For every x in $(a_i, l_{i,k})$ we have $G_k(x) = G_k(a_i)$, so it remains to show that the claimed inequality holds for every x in $(l_{i,k}, a_{i-1})$. For every x in $[l_{i,k}, a_{i-1}]$, part (c) shows that $G_k(x) = G_j(x)$, which means that the inequality is equivalent to

$$\frac{G_k(x) - G_k(a_i)}{G_k(a_{i-1}) - G_k(a_i)} \leq \frac{G_k(x) - G_j(a_i)}{G_k(a_{i-1}) - G_j(a_i)} = \frac{G_k(x) - G_k(a_i) + \varepsilon}{G_k(a_{i-1}) - G_k(a_i) + \varepsilon},$$

for some $\varepsilon \geq 0$ ($G_k(a_i) \geq G_j(a_i)$ by Proposition 2). But for any $y \leq z$ and $\varepsilon \geq 0$, $\frac{y}{z} \leq \frac{y+\varepsilon}{z+\varepsilon}$, so

the inequality holds.

C.4 Proof of Proposition 3

Choose a player $i \leq m$, and let $x^{\text{inf}} = \inf \{x \geq a_i : G_i(x) = G_{i+1}(x)\} \geq a_i$ and $P^{\text{inf}} = G_{i+1}(x^{\text{inf}})$. Part (a) of Theorem 6 shows that $i + 1$ is in \mathcal{A}_1 , so Proposition 2 and Corollary 4 show that $x^{\text{inf}} < 1$ and $G_i(y) = G_{i+1}(y)$ for every $y \geq x^{\text{inf}}$. Therefore, the difference between the expenditures of player $i + 1$ and player i at scores higher than x^{inf} is

$$\int_{x > x^{\text{inf}}} (x - a_{i+1}) dG_{i+1}(x) - \int_{x > x^{\text{inf}}} (x - a_i) dG_i(x) = (a_i - a_{i+1}) (1 - P^{\text{inf}}).$$

The expenditures of player i on $[0, a_i]$ are 0, and his expenditures on $[a_i, x^{\text{inf}}]$ are clearly at most $(1 - a_i) P^{\text{inf}}$. Therefore, to complete the proof it suffices to show that

$$(a_i - a_{i+1}) (1 - P^{\text{inf}}) - (1 - a_i) P^{\text{inf}} \geq 0.$$

For this, let us obtain a useful expression for P^{inf} . First, note that if player $i + 1$ is active at a score y in $[a_i, x^{\text{inf}})$, then by Corollary 4 we have $G_i(y) = G_{i+1}(y)$, contradicting the definition of x^{inf} . This means that $G_{i+1}(a_i) = G_{i+1}(x^{\text{inf}})$, because player $i + 1$ does not have atoms above a_{i+1} . Now, let $b = \prod_{k > i+1} (1 - G_k(a_i)) > 0$ ($b = 1$ for $i = m$). As shown in the proof of Proposition 2, i is in $\mathcal{CP}(a_i)$. Therefore, (2) and the fact that $G_k(a_i) = 0$ for $k < i$ show that

$$1 - b(1 - P^{\text{inf}}) = a_i, \text{ or } P^{\text{inf}} = 1 - \frac{1 - a_i}{b}.$$

Therefore,

$$(a_i - a_{i+1}) (1 - P^{\text{inf}}) - (1 - a_i) P^{\text{inf}} = (a_i - a_{i+1}) \frac{1 - a_i}{b} - (1 - a_i) \left(\frac{b - (1 - a_i)}{b} \right)$$

is non-negative if and only if

$$(a_i - a_{i+1}) (1 - a_i) - (1 - a_i) (b - (1 - a_i)) = (1 - a_i) (1 - b - a_{i+1}) \geq 0.$$

Because $a_i < 1$, it remains to show that $1 - b \geq a_{i+1}$. That $i + 1$ is in $\mathcal{CP}(a_{i+1})$ implies that $1 - \prod_{k > i+1} (1 - G_k(a_{i+1})) = a_{i+1}$, and because CDFs are non-decreasing we have $1 - b \geq a_{i+1}$.

C.5 Proof of Theorem 7

To show that the cardinality of \mathcal{A}^m is C^m , I first define \mathcal{B}^m , a set of sequences that correspond to all the balanced expressions with m open parentheses and m close parentheses. Each sequence in \mathcal{B}^m “mirrors” a different balanced expression.

Lemma 15 *Let*

$$\mathcal{B}^m = \left\{ m - \sum_{i=2}^m b_i, b_2, \dots, b_m : \begin{array}{l} \text{for every } i > 1, b_i \text{ is a non-negative integer and} \\ \sum_{i=k}^m b_i \leq m + 1 - k \text{ for every } k > 1 \end{array} \right\}.$$

The cardinality of \mathcal{B}^m is C^m .

Proof. Consider a sequence of parentheses consisting of m open parentheses and m close parentheses. It is immediate that the parentheses are correctly matched if and only if, when the sequence is read from left to right, at every point the number of close parentheses that have been read is no higher than the number of open parentheses that have been read. Now, interpret b_i as the number of close parentheses placed immediately after the $(m + 1 - i)^{\text{th}}$ open parentheses in the sequence. Then the vectors in \mathcal{B}^m correspond to precisely to all the ways in which m open parentheses and m close parentheses can be sequenced so they are correctly matched. For example, \mathcal{B}^2 consists of the two sequences 2, 0 and 1, 1, which correspond to the balanced expressions $(())$ and $() ()$, and \mathcal{B}^3 consists of the five sequences 3, 0, 0, 2, 1, 0, 1, 2, 0, 2, 0, 1, and 1, 1, 1, which correspond to the balanced expressions $((()))$, $(() ())$, $(()) ()$, $() (())$, and $() () ()$. ■

Because the cardinality of \mathcal{B}^m is C^m , the following lemma shows that the cardinality of \mathcal{A}^m is C^m .

Lemma 16 *\mathcal{A}^m and \mathcal{B}^m have the same cardinality.*

Proof. I will show that \mathcal{A}^m and \mathcal{B}^m each have the same cardinality as \mathcal{D}^m , where \mathcal{D}^m is the set of sequences $\mathcal{D}_1, \dots, \mathcal{D}_{m+1}$ of successively (weakly) finer partitions of $\{1, 2, \dots, m + 1\}$ that satisfy (i) $\mathcal{D}_1 = \{\{1, 2, \dots, m + 1\}\}$ and $\mathcal{D}_{m+1} = \{\{1\}, \{2\}, \dots, \{m + 1\}\}$, and (ii) for every $i > 1$, if $\mathcal{D}_{i-1} \neq \mathcal{D}_i$, then \mathcal{D}_{i-1} is obtained from \mathcal{D}_i by merging $\{i - 1\}$ with one or more “consecutive” partition elements in \mathcal{D}_i , so that the partition element in \mathcal{D}_{i-1} containing $i - 1$ is $\{i - 1, i, \dots, k\}$ for some $k > i - 1$. (This implies that if $j < i$ then $\{j\} \in \mathcal{D}_i$, and every partition element of every partition \mathcal{D}_i consists of consecutive integers.) For example, \mathcal{D}^2 consists of the two sequences of partitions $\{\{1, 2, 3\}\}, \{\{1\}, \{2\}, \{3\}\}, \{\{1\}, \{2\}, \{3\}\}$ and $\{\{1, 2, 3\}\}, \{\{1\}, \{2, 3\}\}, \{\{1\}, \{2\}, \{3\}\}$, and \mathcal{D}^3 consists of the five sequences of partitions

$$\{\{1, 2, 3, 4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \quad (7)$$

$$\{\{1, 2, 3, 4\}\}, \{\{1\}, \{2, 3\}, \{4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \quad (8)$$

$$\{\{1, 2, 3, 4\}\}, \{\{1\}, \{2, 3, 4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \quad (9)$$

$$\{\{1, 2, 3, 4\}\}, \{\{1\}, \{2\}, \{3, 4\}\}, \{\{1\}, \{2\}, \{3, 4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}, \quad (10)$$

and

$$\{\{1, 2, 3, 4\}\}, \{\{1\}, \{2, 3, 4\}\}, \{\{1\}, \{2\}, \{3, 4\}\}, \{\{1\}, \{2\}, \{3\}, \{4\}\}. \quad (11)$$

Now, for every sequence $\mathcal{D}_1, \dots, \mathcal{D}_{m+1}$ in \mathcal{D}^m , let $M^{\mathcal{D}\mathcal{A}}(\mathcal{D}_1, \dots, \mathcal{D}_{m+1})$ be the sequence $\mathcal{A}_1, \dots, \mathcal{A}_m$ defined by $\mathcal{A}_i = \mathcal{D}_i \setminus \mathcal{D}_{i+1}$. The function $M^{\mathcal{D}\mathcal{A}}$ maps the partition \mathcal{D}_i to the partition element in \mathcal{D}_i that contains i , unless that element is a singleton, in which case \mathcal{D}_i is mapped to the empty set. For example, $M^{\mathcal{D}\mathcal{A}}$ maps (7) to $\{1, 2, 3, 4\}, \phi, \phi$, (8) to $\{1, 2, 3, 4\}, \{2, 3\}, \phi$, (9) to $\{1, 2, 3, 4\}, \{2, 3, 4\}, \phi$, (10) to $\{1, 2, 3, 4\}, \phi, \{3, 4\}$, and (11) to $\{1, 2, 3, 4\}, \{2, 3, 4\}, \{3, 4\}$. Similarly, for every sequence $\mathcal{D}_1, \dots, \mathcal{D}_{m+1}$ in \mathcal{D}^m , let $M^{\mathcal{D}\mathcal{B}}(\mathcal{D}_1, \dots, \mathcal{D}_{m+1})$ be the sequence b_1, \dots, b_m defined by $b_i = |\mathcal{D}_{i+1}| - |\mathcal{D}_i|$. The function $M^{\mathcal{D}\mathcal{B}}$ maps partition \mathcal{D}_i to the number of partition elements in \mathcal{D}_{i+1} that were merged with $\{i\}$ to obtain \mathcal{D}_i . For example, $M^{\mathcal{D}\mathcal{B}}$ maps (7) to 3, 0, 0, (8) to 2, 1, 0, (9) to 1, 2, 0, (10) to 2, 0, 1, and (11) to 1, 1, 1. Using the definitions of $\mathcal{A}^m, \mathcal{B}^m$, and \mathcal{D}^m , it is straightforward to show that $M^{\mathcal{D}\mathcal{A}}$ and $M^{\mathcal{D}\mathcal{B}}$ are well-defined, one-to-one, and onto. ■

The one-to-one and onto function $M^{\mathcal{D}\mathcal{A}} \circ M^{\mathcal{D}\mathcal{B}^{-1}}$, which maps every sequence in \mathcal{B}^m to a sequence in \mathcal{A}^m , leads to the following procedure for mapping a balanced expression with m open parentheses and m close parentheses to a sequence in \mathcal{A}^m . Given such an expression, define the sequence $\mathcal{A}_1, \dots, \mathcal{A}_m$ recursively from \mathcal{A}_m to \mathcal{A}_1 as follows. Let $\mathcal{A}_{m+1} = \phi$. For $i \leq m$, denote by j_i the number of close parentheses placed immediately after the $(m+1-i)^{\text{th}}$ open parentheses. If $j_i = 0$, then $\mathcal{A}_i = \phi$. Otherwise, $\mathcal{A}_i = \cup_{k=0}^{j_i} \mathcal{F}_i^k$, where $\mathcal{F}_i^0 = \{i\}$ and $\mathcal{F}_i^k = \mathcal{A}_{1+\max \mathcal{F}_i^{k-1}} \cup \{1 + \max \mathcal{F}_i^{k-1}\}$ for $k > 0$.

For example, for $m = 2$, the expression $(())$ leads to the sequence $\{1, 2, 3\}, \phi$, and the expression $()()$ leads to the sequence $\{1, 2, 3\}, \{2, 3\}$. For $m = 3$, the expression $((()))$ leads to the sequence $\{1, 2, 3, 4\}, \phi, \phi$, the expression $(())()$ leads to the sequence $\{1, 2, 3, 4\}, \{2, 3\}, \phi$, the expression $(())()$ leads to the sequence $\{1, 2, 3, 4\}, \{2, 3, 4\}, \phi$, the expression $()()$ leads to the sequence $\{1, 2, 3, 4\}, \phi, \{3, 4\}$, and the expression $()()()$ leads to the sequence $\{1, 2, 3, 4\}, \{2, 3, 4\}, \{3, 4\}$.

C.6 Proof of Theorem 8

Figure 5 shows that the statement of the theorem holds for $m = 1$.

For every sequence $\mathcal{A}_1, \dots, \mathcal{A}_m$ in \mathcal{A}^m , let $\mathcal{D}_1, \dots, \mathcal{D}_{m+1}$ be the corresponding sequence of partitions given by $M^{\mathcal{D}\mathcal{A}}$ from Lemma 16.⁴⁴ I will prove by induction on $m \geq 2$ that for every sequence $\mathcal{A}_1, \dots, \mathcal{A}_m$ in \mathcal{A}^m there exists an $(m+1)$ -player all-pay auction for m prizes of common value 1 with head starts $0 < a_m < \dots < a_1 < 1$ such that in the unique equilibrium \mathbf{G} (i) among players $i, \dots, m+1$ precisely the players in \mathcal{A}_i compete (have best responses) in (a_i, a_{i-1}) (where $a_0 = 1$) and (ii) for every partition element in $\mathcal{D}_2 \setminus \{1\}$ there is a *distinct* score x in $[a_1, 1)$ such

⁴⁴ $\mathcal{D}_{m+1} = \{\{1\}, \dots, \{m+1\}\}$, and for $i < m+1$ $\mathcal{D}_i = (\mathcal{D}_{i+1} \cup \{\mathcal{A}_i\}) \setminus \mathcal{B}_i$, where \mathcal{B}_i is the set of partition elements in \mathcal{D}_{i+1} whose intersection with \mathcal{A}_i is non-empty (and are therefore subsets of \mathcal{A}_i), that is,

$$\mathcal{B}_i = \{\mathcal{F} \in \mathcal{D}_{i+1} : \mathcal{F} \cap \mathcal{A}_i \neq \phi\} = \{\mathcal{F} \in \mathcal{D}_{i+1} : \mathcal{F} \subseteq \mathcal{A}_i\}.$$

that for every player i in the partition element the set of scores in $[a_1, 1)$ at which (2) holds is $[x, 1)$.

For $m = 2$, \mathcal{A}^2 consists of the two sequences $\{1, 2, 3\}, \{2, 3\}$ and $\{1, 2, 3\}, \phi$. Figure 6 corresponds to the sequence $\{1, 2, 3\}, \{2, 3\}$: in $(a_1, 1)$ all three players compete, and in (a_2, a_1) only players 2 and 3 compete. Also, $\mathcal{D}_2 \setminus \{1\} = \{\{2, 3\}\}$, and the set of scores in $[a_1, 1)$ at which (2) holds for players $i = 2, 3$ is $[a_1, 1)$. Figure 7 corresponds to the sequence $\{1, 2, 3\}, \phi$: in $(a_1, 1)$ all three players compete, and in (a_2, a_1) no player competes. Also, $\mathcal{D}_2 \setminus \{1\} = \{\{2\}, \{3\}\}$, the set of scores in $[a_1, 1)$ at which (2) holds for player $i = 2$ is $[a_1, 1)$, the set of scores in $[a_1, 1)$ at which (2) holds for player $i = 3$ is $[2a_2 - a_2^2, 1)$, and $a_1 < 2a_2 - a_2^2$.

For the induction step, suppose that (i) and (ii) hold for m , and consider a sequence $\mathcal{A}_1, \dots, \mathcal{A}_{m+1}$ in \mathcal{A}^{m+1} and its corresponding sequence of partitions $\mathcal{D}_1, \dots, \mathcal{D}_{m+2}$. Define the sequence $\tilde{\mathcal{A}}_1, \dots, \tilde{\mathcal{A}}_m$ by $\tilde{\mathcal{A}}_1 = \{1, \dots, m\}$ and $\tilde{\mathcal{A}}_i = \{j > 0 : j + 1 \text{ is in } \mathcal{A}_{i+1}\}$ for $i > 1$. It is easy to verify that $\tilde{\mathcal{A}}_1, \dots, \tilde{\mathcal{A}}_m$ is in \mathcal{A}^m . Let $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_{m+1}$ be the sequence of partitions corresponding to $\tilde{\mathcal{A}}_1, \dots, \tilde{\mathcal{A}}_m$. By the induction hypothesis, there exists an $(m + 1)$ -player all-pay auction for m prizes of common value 1 with head starts $0 < \tilde{a}_m < \dots < \tilde{a}_1 < 1$ such that in its unique equilibrium $\tilde{\mathbf{G}} = (\tilde{G}_1, \dots, \tilde{G}_{m+1})$ both (i) and (ii) hold with $\tilde{\mathcal{A}}_i$ instead of \mathcal{A}_i and $\tilde{\mathcal{D}}_2 \setminus \{1\}$ instead of $\mathcal{D}_2 \setminus \{1\}$.

Now modify the contest by adding another prize of value 1 and another player with a head start a_1 in $(\tilde{a}_1, 1)$. The modified contest has $m + 2$ players, $m + 1$ prizes of value 1, and head starts $0 < a_{m+1} < \dots < a_1 < 1$, with $a_i = \tilde{a}_{i-1}$ for $i > 1$. Denote by $\mathbf{G} = (G_1, \dots, G_{m+2})$ the equilibrium of this contest. It is easy to see that the execution of the algorithm on $[0, \tilde{a}_1) = [0, a_2)$ is the same for both contests, so $G_i(x) = \tilde{G}_{i-1}(x)$ for every $i > 1$ and $x < a_2 = \tilde{a}_1$. This immediately implies that in \mathbf{G} , among players $i, \dots, m + 2$, precisely the players in \mathcal{A}_i compete in (a_i, a_{i-1}) , for $i > 2$. Moreover, part (a) of Theorem 6 shows that all players compete in $(a_1, 1)$. I will show that a_1 can be chosen so that in \mathbf{G} , among players $2, \dots, m + 2$, precisely the players in \mathcal{A}_2 compete in (a_2, a_1) and (ii) holds.

If $\mathcal{A}_2 = \phi$, set a_1 to any value in $(\tilde{a}_1, 1 - \Pi_{i=1}^{m+1}(1 - \tilde{G}_i(\tilde{a}_1)))$. This interval is non-empty, because $\tilde{G}_1(\tilde{a}_1) > 0$ and in $\tilde{\mathbf{G}}$ player 1 is active at \tilde{a}_1 so $1 - \Pi_{i=2}^{m+1}(1 - \tilde{G}_i(\tilde{a}_1)) = \tilde{a}_1$ (as shown in the proof of Proposition 2). Choosing a_1 in the specified interval guarantees that $G_2(a_2) < \tilde{G}_1(\tilde{a}_1)$ (otherwise in \mathbf{G} player 1 would obtain at least $1 - \Pi_{i=1}^{m+1}(1 - \tilde{G}_i(\tilde{a}_1)) > a_1$ by choosing scores immediately above a_2). And $G_2(a_2) < \tilde{G}_1(\tilde{a}_1)$ implies that $\mathcal{CP}(a_2) = \{1, 2\}$ in the modified contest, since $\tilde{G}_1(\tilde{a}_1)$ was minimal atom size at \tilde{a}_1 such that in $\tilde{\mathbf{G}}$ some player $i > 1$ would be active immediately above $\tilde{a}_1 = a_2$. Therefore, at a_2 the algorithm proceeds to Stage 3, \mathbf{G} does not change in (a_2, a_1) , and players $2, \dots, m + 2$ do not compete in (a_2, a_1) .

If $\mathcal{A}_2 \neq \phi$, then $\mathcal{A}_2 = \{2, 3, \dots, j\}$ for some $3 \leq j \leq m + 2$. Denote by $\tilde{x}_k \geq \tilde{a}_1$ the score identified in (ii) (applied to $\tilde{\mathbf{G}}$) that corresponds to the partition element in $\tilde{\mathcal{D}}_2 \setminus \{1\}$

that contains k . If $j = m + 2$, then set a_1 to any value in $\left(1 - \prod_{i=1}^{m+1} \left(1 - \tilde{G}_i(\tilde{x}_{m+1})\right), 1\right)$. (Note that $1 - \prod_{i=2}^{m+1} \left(1 - \tilde{G}_i(\tilde{x}_{m+1})\right) = \tilde{x}_{m+1}$, because every score in $(\tilde{a}_1, 1)$ is a best response in $\tilde{\mathbf{G}}$ for player 1, as shown in the proof of Proposition 2, so $a_1 > \tilde{x}_{m+1}$.) If $j < m + 2$, then, by part (b) of Theorem 6, $\tilde{x}_{j-1} \leq \tilde{x}_j$. To see that the inequality is strict, note that j and $j + 1$ belong to different partition elements in \mathcal{D}_2 (since $\mathcal{A}_2 = \mathcal{D}_2 \setminus \mathcal{D}_3$), and therefore in \mathcal{D}_3 (\mathcal{D}_2 is a coarsening of \mathcal{D}_3), so $j - 1$ and j belong to different partition elements in $\tilde{\mathcal{D}}_2 \setminus \{1\}$. So (ii) (applied to $\tilde{\mathbf{G}}$) shows that $\tilde{x}_{j-1} < \tilde{x}_j$. In this case, set a_1 to any value in $\left(1 - \prod_{i=1}^{m+1} \left(1 - \tilde{G}_i(\tilde{x}_{j-1})\right), 1 - \prod_{i=1}^{m+1} \left(1 - \tilde{G}_i(\tilde{x}_j)\right)\right)$ (part (b) of Theorem 6 shows that this interval is non-empty).

This value of a_1 guarantees that player 1 obtains his power for the first time at a score $y < a_1$ in $(\tilde{x}_{j-1}, \tilde{x}_j)$, so $G_i(x) = \tilde{G}_{i-1}(x)$ for every $i > 1$ and $x \leq y$. As a result, every score in $(\tilde{x}_{j-1}, y) \subseteq (a_2, a_1)$ is a best response in \mathbf{G} for players $2, 3, \dots, j$ (because, by definition of \tilde{x}_{j-1} , every score in (\tilde{x}_{j-1}, y) is a best response in $\tilde{\mathbf{G}}$ for player $j - 1$, and, by Corollary 4, for every player $i < j - 1$). Because in \mathbf{G} players $j + 1, j + 2, \dots, m + 2$ do not have best responses in $(a_2, y) \subseteq (\tilde{a}_1, \tilde{x}_j)$ (by definition of \tilde{x}_k , for $k \geq j$, because $\tilde{x}_j \leq \tilde{x}_k$), and at y the algorithm proceeds to Stage 3 and no player's CDF increases until a_1 , in \mathbf{G} players $j + 1, j + 2, \dots, m + 2$ do not compete in (a_2, a_1) . This shows that in \mathbf{G} , among players $2, \dots, m + 2$, precisely the players in \mathcal{A}_2 compete in (a_2, a_1) .

To see that (ii) holds in \mathbf{G} , consider any two players in the same partition element in $\mathcal{D}_2 \setminus \{1\}$. There is a maximal $k < m + 2$ for which both players are in the same partition element in \mathcal{D}_k , which implies that both players are in $\mathcal{A}_k = \mathcal{D}_k \setminus \mathcal{D}_{k+1}$. Since $k \geq 2$, Corollary 4 shows that the CDFs of the two players coincide at some score $x < a_1$, and remain equal for all higher scores. Part (ii) of Lemma 13 then shows that the players' best response sets in $(x, 1)$ are the same, which suffices because of part (b) of Theorem 6.

Now consider two players $i < j$ that belong to different partition elements in $\mathcal{D}_2 \setminus \{1\}$. To complete the proof, we must show that the players' sets of best responses in $(a_1, 1)$ are two distinct intervals with upper bound 1. By part (b) of Theorem 6, it suffices to show that the players' sets of best responses in $(a_1, 1)$ are distinct. There are two cases to consider. The first case is $i = 2$ and $\{2\}$ is in \mathcal{D}_2 . That $\{2\}$ is in \mathcal{D}_2 implies that $\mathcal{A}_2 = \phi$ (because $\mathcal{A}_2 = \mathcal{D}_2 \setminus \mathcal{D}_3$, $\{2\}$ is in \mathcal{D}_3 , and if $\mathcal{A}_2 \neq \phi$, then \mathcal{A}_2 includes 2 and at least one other player $k > 2$). As described above, the choice of a_1 when $\mathcal{A}_2 = \phi$ is such that in \mathbf{G} we have $\mathcal{CP}(a_2) = \{1, 2\}$, which implies that $\mathcal{CP}(a_1) = \{1, 2\}$.⁴⁵ This, together with the observation that Stage 4 is executed at a_1 , means that scores immediately above a_1 are best responses for player 2 and are not best responses for any player $j > 2$. The second case to consider is $i > 2$ or $\{2\}$ is not in \mathcal{D}_2 . If $i = 2$, choose another player i other than 2 from the partition element in $\mathcal{D}_2 \setminus \{1\}$ that

⁴⁵If $i > 2$ is in $\mathcal{CP}(a_1)$, then by Corollary 4 $G_i(a_1) = G_2(a_1)$, which implies that $G_i(a_2) = G_2(a_2)$ (because Stage 3 is executed at a_2). But then part (ii) of Lemma 13 shows that i is in $\mathcal{CP}(a_2)$.

contains 2 (we already know that players in the same partition element in $\mathcal{D}_2 \setminus \{1\}$ have the same best responses in $(a_1, 1)$). Because i and j are in separate partition elements in \mathcal{D}_2 , they are in separate partition elements in \mathcal{D}_3 (\mathcal{D}_2 is a coarsening of \mathcal{D}_3), and therefore $i - 1$ and $j - 1$ are in separate partition elements of $\tilde{\mathcal{D}}_2 \setminus \{1\}$. Therefore, in $\tilde{\mathbf{G}}$ their best response sets in $(\tilde{a}_1, 1)$ are, respectively, $[\tilde{x}_{i-1}, 1) \setminus \{\tilde{a}_1\}$ and $[\tilde{x}_{j-1}, 1)$ for some $\tilde{x}_{i-1} < \tilde{x}_{j-1}$. Because j is not in \mathcal{A}_2 (otherwise $i < j$ would also be in \mathcal{A}_2 , and therefore both would belong to the same partition element in $\mathcal{D}_2 \setminus \{1\}$), as described above the choice of a_1 is such that the score y in $[a_2, a_1)$ at which player 1 first becomes active in \mathbf{G} (and Stage 3 is executed) satisfies $y < \tilde{x}_{j-1}$,⁴⁶ and $G_k(x) = \tilde{G}_{k-1}(x)$ for every $k > 1$ and $x \leq y$. Now, because $\tilde{G}_{i-1}(\tilde{x}_{j-1}) = \tilde{G}_{j-1}(\tilde{x}_{j-1})$ (by Corollary 4) and \tilde{G}_{i-1} strictly increases in $(\tilde{x}_{i-1}, \tilde{x}_{j-1})$, we have $\tilde{G}_{i-1}(y) < \tilde{G}_{j-1}(y)$, so $G_i(y) < G_j(y)$. Since \mathbf{G} does not change on $[y, a_1)$ (Stage 3 is executed), and neither player i nor j has an atom at a_i , we have $G_i(a_i) < G_j(a_i)$. Suppose that in \mathbf{G} $[x, 1) \setminus \{a_1\}$ is the set of best responses in $(a_1, 1)$ for players i and j . Then G_i and G_j do not increase on $[a_1, x]$, so $G_i(x) < G_j(x)$, but Corollary 4 shows that $G_i(x) = G_j(x)$. Therefore, in \mathbf{G} the sets of best responses in $(a_1, 1)$ for players i and j differ.

C.7 Proof of Proposition 4

Because of the nature of the proposition, I relax the assumption that $a_{m+1}^k = 0$ and $V = 1$, which implies that the power of every player $i \leq m + 1$ in the k^{th} contest is $a_i^k - a_{m+1}^k$.

Consider first the sequence of reduced contests that involve only players a_1^k, \dots, a_{m+1}^k . Choose some x in $(0, V)$ and some M so that for all $k \geq M$ we have $a_1^k < x$ and $a_1^k - a_{m+1}^k < 1 - (1 - (x - a_{m+1}^k)/V)^{1/m}$. Because the cost function of player $m + 1$ is strictly increasing above $a_{m+1}^k < x$, if x is a best response for player $m + 1$, then it is also a best response for another player $i < m + 1$. Therefore, Lemma 2 implies that $(1 - \prod_{j \neq i} (1 - G_j^k(x))) V = x - a_{m+1}^k$ for some $i < m + 1$. This implies that $\max_{j \neq i} G_j^k(x) \geq 1 - (1 - (x - a_{m+1}^k)/V)^{1/m}$. Proposition 2 shows that $\max_{j \neq i} G_j^k(x) = G_{m+1}^k(x)$, so $G_{m+1}^k(x) \geq 1 - (1 - (x - a_{m+1}^k)/V)^{1/m}$. Also, $G_{m+1}^k(a_1^k) \leq a_1^k - a_{m+1}^k$, otherwise player 1 could get more than his power by choosing a_1^k . Because $a_i^k - a_{m+1}^k < 1 - (1 - (x - a_{m+1}^k)/V)^{1/m}$, we have that $G_{m+1}^k(a_1^k) < G_{m+1}^k(x)$. This implies that G_{m+1}^k increases in (a_1^k, x) , so players $1, \dots, m + 1$ are active immediately above x (by part (b) of Theorem 6). Therefore, for every player $j \leq m + 1$ (5) shows that

$$G_j^k(x) = 1 - \left(1 - \frac{x - a_{m+1}^k}{V}\right)^{\frac{1}{m}} \xrightarrow{k \rightarrow \infty} 1 - \left(1 - \frac{x}{V}\right)^{\frac{1}{m}}.$$

Returning to the original sequence of contests, by Corollary 1 no player $j > m + 1$ participates in any contest in the sequence, so for every k we have $G_j^k(x) = \begin{cases} 0 & \text{if } x < a_j^k \\ 1 & \text{if } x \geq a_j^k \end{cases}$, which implies

⁴⁶If $\mathcal{A}_2 = \phi$, then $y = a_2$. If $\mathcal{A}_2 \neq \phi$, then y is in (a_2, a_1) .

that G_j^k converges weakly to $1_{[0,\infty)}$.

References

- [1] **Amann, Erwin, and Wolfgang Leininger.** 1996. "Asymmetric All-Pay Auctions with Incomplete Information: The Two-Player Case." *Games and Economic Behavior*, 14(1): 1-18.
- [2] **Barut, Yasar, and Dan Kovenock.** 1998. "The Symmetric Multiprize All-Pay Auction with Complete Information." *European Journal of Political Economy*, 14(4): 627-644.
- [3] **Baye, Michael R., Dan Kovenock, and Casper de Vries.** 1993. "Rigging the Lobbying Process: An Application of All-Pay Auctions." *American Economic Review*, 83(1): 289-94.
- [4] **Baye, Michael R., Dan Kovenock, and Casper de Vries.** 1996. "The All-Pay Auction with complete-information." *Economic Theory*, 8(2): 291-305.
- [5] **Casas-Arce, Pablo, and F. Asís Martínez-Jerez.** 2010. "The All-Pay Auction with Handicaps." Mimeo.
- [6] **Chapman, Charles P.** 2002. *Real Mathematical Analysis*, New York: Springer-Verlag.
- [7] **Che, Yeon-Koo, and Ian Gale.** 1998. "Caps on Political Lobbying." *American Economic Review*, 88(3): 643-51.
- [8] **Che, Yeon-Koo, and Ian Gale.** 2003. "Optimal Design of Research Contests." *American Economic Review*, 93(3): 646-71.
- [9] **Che, Yeon-Koo, and Ian Gale.** 2006. "Caps on Political Lobbying: Reply." *American Economic Review*, 96(4): 1355-60.
- [10] **Clark, Derek J., and Christian Riis.** 1998. "Competition over More than One Prize." *American Economic Review*, 88(1): 276-289.
- [11] **Dasgupta, Partha.** 1986. "The Theory of Technological Competition." In *New Developments in the Analysis of Market Structure*, ed. Joseph E. Stiglitz and G. F. Mathewson, 519-47. Cambridge: MIT press.
- [12] **Ellingsen, Tore.** 1991. "Strategic Buyers and the Social Cost of Monopoly." *American Economic Review*, 81(3): 648-57.
- [13] **Fu, Qiang, Jingfeng Lu, and Yuanzhu Lu.** 2009. "How to Motivate Innovation: Subsidies or Prizes?" Mimeo.
- [14] **González-Díaz, Julio.** 2009. "First-Price Winner-Takes-All Contests." *Optimization*, forthcoming.

- [15] **Hillman, Arye L., and John G. Riley.** 1989. “Politically Contestable Rents and Transfers.” *Economics and Politics*, 1(1): 17-39.
- [16] **Hillman, Arye L., and Dov Samet.** 1987. “Dissipation of Contestable Rents by Small Numbers of Contenders.” *Public Choice*, 54(1): 63-82.
- [17] **Kaplan, Todd R., Israel Luski, and David Wettstein.** 2003. “Innovative activity and sunk cost.” *International Journal of Industrial Organization*, 21(8): 1111-1133.
- [18] **Kaplan, Todd R., and David Wettstein.** 2006. “Caps on Political Lobbying: Comment” *American Economic Review*, 96(4): 1351-4.
- [19] **Kirkegaard, René.** 2009. “Favoritism in Asymmetric Contests: Head Starts and Handicaps.” Mimeo.
- [20] **Kirkegaard, René.** 2010. “Preferential Treatment May Hurt: Another Application of the All-Pay Auction.” Mimeo.
- [21] **Konrad, Kai A.** 2002. “Investment in the absence of property rights: the role of incumbency advantages.” *European Economic Review*, 46(8): 1521-1537.
- [22] **Konrad, Kai A.** 2004. “Inverse campaigning.” *Economic Journal*, 114(492): 69-82.
- [23] **Konrad, Kai A.** 2007. “Strategy in Contests - an Introduction.” Berlin Wissenschaftszentrum Discussion Paper SP II 2007 – 01.
- [24] **Koshy, Thomas.** 2009. *Catalan Numbers with Applications*, New York: Oxford University Press.
- [25] **Lazear, Edward P., and Sherwin Rosen.** 1981. “Rank-Order Tournaments as Optimum Labor Contracts.” *Journal of Political Economy*, 89(5): pp. 841-64.
- [26] **Moldovanu, Benny, and Aner Sela.** 2001. “The Optimal Allocation of Prizes in Contests.” *American Economic Review*, 91(3): 542-58.
- [27] **Moldovanu, Benny, and Aner Sela.** 2006. “Contest Architecture.” *Journal of Economic Theory*, 126(1): 70-96.
- [28] **Nitzan, Shmuel.** 1994. “Modeling Rent-Seeking Contests.” *European Journal of Political Economy*, 10(1): 41-60.
- [29] **Parreiras, Sergio, and Anna Rubinchik.** 2009. “Contests with Many Heterogeneous Agents.” *Games and Economic Behavior*, 68(2): 703-715.
- [30] **Siegel, Ron.** 2009. “All-Pay Contests.” *Econometrica*, 77(1): 71-92.

- [31] **Siegel, Ron.** 2010. “Asymmetric Contests with Conditional Investments.” *American Economic Review*, 100(5): 2230-2260.
- [32] **Siegel, Ron.** 2011. “Asymmetric Contests with Interdependent Valuations.” Mimeo.
- [33] **Tullock, Gordon.** 1980. “Efficient Rent Seeking.” In *Toward a theory of the rent seeking society*, ed. James M. Buchanan, Robert D. Tollison, and Gordon Tullock, 269-82. College Station: Texas A&M University Press.
- [34] **Varian, Hal.** 1980. “A Model of Sales.” *American Economic Review*, 70(4): 651-58.