

Favoritism in Asymmetric Contests: Head Starts and Handicaps*

René Kirkegaard
Department of Economics
University of Guelph
50 Stone Road East, Guelph,
Ontario, N1G 2W1, Canada.
rkirkega@uoguelph.ca

March 2012

Abstract

I examine a contest with identity-dependent rules in which contestants are privately informed and ex ante heterogenous. A contestant may suffer from a handicap or benefit from a head start. The former reduces the contestant's score by a fixed percentage; the latter is an additive bonus. Although total effort increases if the weak contestant is favored with a head start, the optimal use of handicaps is not as clear-cut. Depending on the nature of the asymmetry, it may or may not be optimal to handicap the strong contestant. Moreover, it is generally optimal to combine the two instruments. For instance, when contestants are sufficiently heterogenous the weak contestant should be given both a head start and a handicap. It may also be possible to induce higher effort and at the same time make both contestants better off ex ante.

JEL Classification Numbers: C72, D44, D82.

Keywords: All-Pay Auctions, Contests, Favoritism, Handicap, Head Start.

*I would like to thank three anonymous referees and an advisory editor as well as J. Atsu Amegashie, Arnaud Dellis, Gregory Pavlov, Al Slivinski, Peter Streufert, Myrna Wooders, John Wooders, and Lixin Ye for comments and suggestions.

1 Introduction

There are countless examples of contests, broadly defined, in which one contestant is explicitly or implicitly favored over his competitors. Regulation is often imposed to manipulate the contest; examples include sports (e.g. golf and horse racing), affirmative action, uneven treatment of internal and external applicants for senior positions, and the preferential treatment occasionally given to domestic firms in procurement. However, the finer details of regulation differ from example to example. Typically, an additional complication is that the advantaged and disadvantaged contestants are heterogenous on some dimension. In fact, it is often precisely this underlying asymmetry among contestants that is used to justify rules that are asymmetric.

The objective of this paper is to formally analyze the consequences of favoritism when contestants are heterogenous. I will emphasize that different types of favoritism may have dramatically different effects, and, accordingly, that they should be utilized differently. I will also provide a rationale for the simultaneous use of several different (and sometimes seemingly contradictory) instruments. An application to real-world contests for research funding is discussed at the end of the paper.

In keeping with the objective of the paper, two simple and realistic forms of favoritism are considered. A contestant has a *head start* if he has an “absolute” advantage over his rivals, in the sense that he wins even if his effort falls short of his rivals’ effort by a pre-specified amount. Alternatively, the contestant can be favored by *handicapping* his rivals, i.e. by discounting their effort by a pre-specified percentage. To be clear, these instruments are analyzed not because they are optimal (which typically they are not), but rather because they are observed in practice and in combination serve to convey the main point, namely that different types of favoritism function in very different ways. In short, it matters how favoritism is modeled.¹

To be more specific, a handicap affects the marginal return to increasing effort, but this is not the case for a head start. As a result, the two instruments lead to different strategic considerations. Indeed, it may be in the contest designer’s interest to combine different instruments, for instance by giving a contestant a head start and a handicap at the same time.

The model and related literature is discussed next. An overview of the main results follow.

MODEL AND LITERATURE. I model the contest as an all-pay auction in which

¹Head starts and handicaps are identity-dependent. Alternatively, the contest can be manipulated by imposing restrictive rules that apply to all contestants. Although the rules are symmetric, the strategic response may be different for weak and strong contestants. The most common examples, namely bidding caps and bidding floors, are reviewed in Section 10.

bidders are independently and privately informed about their valuation of the prize. For tractability, there are two bidders, one “strong” and one “weak”. Bidders are risk neutral and the cost of bidding is linear in the bid.² Together, the assumptions permit the use of mechanism design. Hence, abstract insights from mechanism design are used to analyze the properties of two specific instruments, rather than to derive the “fully optimal” contest (which is characterized in Myerson (1981)).

The literature on favoritism in all-pay auctions is small. Consider first the complete information literature. Konrad (2002) examines a two-bidder model with head starts and handicaps, but he assumes bidders are symmetric. Siegel (2009) examines a more general model in which contestants may be heterogenous. However, both assume the rules are exogenous.

In contrast, Fu (2006) endogenizes handicaps in a model with two heterogeneous bidders. He finds that it is optimal to handicap the strong bidder. I will show that once incomplete information is introduced, the comparative statics change. Thus, Fu’s (2006) conclusion is not robust quantitatively (see Section 3) or indeed necessarily qualitatively (see Section 8). Although it appears not to have been noted before, it is easy to show that the complete information model leads, unambiguously, to the conclusion that the weak bidder should be given a head start and a handicap (see Section 7). In the incomplete information model, however, this property holds only if the asymmetry is sufficiently large. Konrad’s (2002) and Siegel’s (2009) models are discussed in Sections 8 and 10, respectively. In both cases, examples can be provided in which the incomplete information model yields different conclusions than the complete information model. Thus, the complete information model is arguably somewhat fragile and it seems worthwhile to examine its robustness.

With incomplete information, there appears to be no previous papers dealing with head starts, and only three that examine handicaps.³ Lien (1990) and Feess et al (2008) assume the two bidders are homogeneous. Lien (1990) proves that the handicapped bidder wins less often than is efficient if types are drawn from the same uniform distribution. Feess et al (2008) show that this result holds for any distribution function. Clark and Riis (2000) allow bidders to be heterogenous, but assume types are drawn from uniform distributions. In their model, it is profitable to handicap the strong bidder (expected revenue increases) although social surplus

²Moldovanu and Sela (2001) and Gavious, Moldovanu, and Sela (2002) allow the cost of bidding to be non-linear in the bid but they assume that bidders are homogeneous. Clark and Riis (2000) assume bidding costs are non-linear with types that are drawn from different uniform distributions.

³More accurately, Hickman (2011) in fact considers a policy that is equivalent to a head start. However, his model differs in important aspects from the more standard literature on all-pay auctions. Hickman’s (2011) paper is discussed in more detail below.

decreases. The current paper is the first to study the profitability of handicaps in a more general setting than the uniform model.⁴ It is also the first to consider the combination of head starts and handicaps in an incomplete-information model.

Depending on the context, a plethora of different objectives may underlie the manipulation of contests. For the most part, I will assume that the designer is attempting to maximize the effort or bids of the contestants. Section 9 considers objective functions that take bidders' payoffs into account. The main findings of the paper are outlined below, classified by the objective of the designer.

REVENUE MAXIMIZATION. Given suitable regularity assumptions, it is always profitable to give the weak bidder a head start, irrespective of any handicap. Any head start to the strong bidder lowers expected revenue, unless the strong bidder already suffers from a very severe handicap.

Although it is profitable to handicap the strong bidder in Clark and Riis' (2000) model, the impact of a handicap on expected revenue is non-trivial in the general case. Nevertheless, Section 5 describes two environments where it is optimal to handicap the strong bidder. The uniform model is a special case of one of these models. On the other hand, Section 8 provides an example in which it is optimal to handicap the weak bidder. Another example in Section 8 considers an environment with symmetric bidders that violates the typical regularity assumptions. In that example, it is profitable to give a head start to a random bidder. Note that the complete information model cannot produce the conclusions in the examples in Section 8.

There are parallels between Maskin and Riley's (2000) revenue ranking of standard auctions and the profitability of handicaps in all-pay auctions. For example, Clark and Riis's (2000) uniform model coincides with the leading example in Maskin and Riley (2000) of a situation where the first-price auction is more profitable than the second-price auction. However, from a methodological point of view, the approach in the current paper differs from that used by Maskin and Riley (2000). In short, the main challenge in both problems is that one mechanism does not dominate another for all combinations of types. This complication explicitly lead Maskin and Riley (2000) to claim that mechanism design would be of no help in addressing the problem. Nevertheless, Kirkegaard (2011a) has recently shown that mechanism design can in fact be used to rank standard auctions. Indeed, two of Maskin and Riley's (2000) propositions are corollaries of a more general theorem in Kirkegaard (2011a). The mechanism design approach is also used in the current paper.

⁴The uniform model lends itself to a "brute force" approach because the equilibrium allocation and revenue can be characterized in closed form. However, in the general case considered here only an implicit characterization is possible. For this reason, the current paper must necessarily rely on a different methodology than Clark and Riis (2000).

As mentioned, the joint use of head starts and handicaps is also examined. Though the weak bidder should be given a head start, the size of the asymmetry determines who should be handicapped. The weak bidder is more likely to receive both a head start and a handicap the larger the asymmetry between bidders is.

In summary, three scenarios are examined (depending on whether head starts, handicaps, or both instruments are used). Since each scenario has its unique challenges, the method of proof is somewhat different for each scenario, although the analysis is always based on mechanism design. However, Section 7 demonstrates that the problem is “well-behaved” if the asymmetry between bidders is large enough, in which case it becomes easy to unify the analysis and results of the three scenarios.

BIDDERS’ PAYOFFS. It is straightforward to maximize the sum of the bidders’ payoffs in a contest: The prize should be given to the strong bidder, for free. He is, after all, the bidder who is more likely to value the prize highly. However, there is no worse mechanism from the weak bidder’s point of view. Social welfare functions that take the distribution of payoff into account leads to markedly different policy recommendations. For instance, compare the standard all-pay auction with an all-pay auction in which bidder 1 is given a head start *and* a handicap. Bidder 1 is made better off if his type is low (due to the head start), whereas bidder 2 is made better off if his type is high (due to bidder 1’s handicap). In other words, since the two instruments do not cancel each other out, it is possible to make *both* bidders better off *ex ante*. That is, both bidders are favored.⁵ Indeed, both bidders may become better off while expending more resources, meaning that the recipient of the effort, if any, is also better off. It is possible to favor all parties simultaneously.

In a similar vein, the analysis may shed light on questions surrounding affirmative action and how to “level the playing field”. Roemer (1998) suggests “equality of opportunity” as a desirable norm, which means, roughly, that the distribution of “achievements” should be equalized across bidders. Depending on the definition of “achievements” and the nature of the asymmetry between bidders, head starts and handicaps may or may not be flexible enough to create equality of opportunity, although head starts never works.⁶

In the context of admission into university, the probability of enrollment or admission may be a useful measure of achievement. Incidentally, Hickman (2011) uses the

⁵The Oxford English Dictionary defines “favouritism” as “the unfair favouring of one person or group at the expense of another”. However, the definition of “favour” implies only that the favoured party is better off, not that another party is hurt.

⁶Calsamiglia (2009) makes the point that if one contest is manipulated, behavior in other contests in which the same bidders also partake may be influenced. Thus, pursuing equality of opportunity “locally” may not necessarily advance equality of opportunity “globally”.

“enrollment gap” as one measure of the success of various affirmative action policies. In fact, the current paper and Hickman’s (2011) paper can be viewed as being at opposite extremes of a set of models. In this paper, there are only two bidders. Thus, strategic considerations are all-important, and the fact that different bidders have different beliefs about each other obviously impacts equilibrium strategies. Hickman (2011), in contrast, considers a limiting game with a continuum of minority students and non-minority students (i.e. clones of the weak bidder and the strong bidder, respectively) as well as a continuum of distinct prizes. His model is meant to capture the large scale of the education market. With unbiased rules, a minority and non-minority student with the same type take the same action (study equally hard). The reason is that each student is negligible and therefore has the same beliefs about the mass of students he is competing with, whether he is a minority student or not. In one interpretation of his model, this means that the initial allocation of prizes is efficient, i.e. the best placements are given to those that value them the most. There will be an enrolment gap only because the underlying distribution of types is different in the two groups, but not because they behave differently. Note that any intervention will be detrimental to efficiency, contrary to the conclusions of the current paper. Hickman (2011) show that a quota, by design, can eliminate the enrollment gap in his model, but that a head start is less successful. In fact, a head start to the minority students will be detrimental to overall effort in Hickman (2011), but not necessarily in the model in the current paper.

2 Model and Equilibrium

The contest is modelled as an all-pay auction. There are two bidders. Each bidder is characterized by a privately known type which captures the value the bidder places on winning the auction. More generally, the type can be thought of as the certainty equivalent of winning, in the case where the actual value is not certain. Independently of the other bidder, bidder i draws his type from the twice continuously differentiable distribution function F_i with support $[0, \bar{v}_i]$, $i = 1, 2$. The distribution functions have no mass points. The density, $f_i(v) = F_i'(v)$, is bounded above as well as bounded below, away from zero. Finally,

$$h_i(v) = \frac{1 - F_i(v)}{f_i(v)}$$

is strictly decreasing, $i = 1, 2$. $h_i(v)$ is the reciprocal of the hazard rate.

The two bidders are heterogenous. Bidder 1 is “weak” and bidder 2 is “strong”. Formally, F_2 first order stochastically dominates F_1 , $F_2(v) < F_1(v)$ for all $v \in (0, \bar{v}_1)$.

It will also be assumed that $f_1(0) > f_2(0)$ and $\bar{v}_1 < \bar{v}_2$. With the exception of Section 8, the assumptions made so far are imposed throughout the paper.

Next, define $\lambda(v)$ as the solution to $F_1(\lambda) = F_2(v)$, $v \in [0, \bar{v}_2]$. By definition, bidder 1 is as likely to have a type below λ as bidder 2 is to have a type below v ; the two bidders have the same “rank”. Let $\bar{\lambda} = \max \lambda(v)/v$ and $\underline{\lambda} = \min \lambda(v)/v$, with

$$1 > \bar{\lambda} \geq \frac{\bar{v}_1}{v_2} \geq \underline{\lambda}.$$

The first inequality is attributable to first order stochastic dominance and the others to the fact that $\lambda(\bar{v}_2) = \bar{v}_1$. The difference between $\bar{\lambda}$ and $\underline{\lambda}$ can be viewed as a measure of how bidder heterogeneity varies with rank. The behavior of $\lambda(v)/v$ will at times be important. The following definitions will therefore be useful.

Definition 1 F_2 is smaller than F_1 in the star order if $\lambda(v)/v$ is non-decreasing.⁷

Definition 2 F_1 is said to be a scaled down version of F_2 if $F_i(v) = F(\frac{v}{\bar{v}_i})$, $v \in [0, \bar{v}_i]$, $i = 1, 2$, and $\bar{v}_1 < \bar{v}_2$, where F is some distribution function with support $[0, 1]$.⁸

The environment in Definition 2 satisfies Definition 1. Specifically, $\lambda(v)/v = \bar{v}_2/\bar{v}_1$ is constant when F_1 is a scaled down version of F_2 . The uniform model is a special case. The complete information model can also be considered a special case of the model in Definition 2, in which F is degenerate with mass 1 at $v = 1$. However, the complete information model obviously violates the assumption that densities are positive. Definition 1 is satisfied if F_1 is concave and F_2 is convex, for example.

The difference between v and $\lambda(v)$ is also of some relevance.

Definition 3 F_1 is smaller than F_2 in the dispersive order if $v - \lambda(v)$ is non-decreasing, or $\lambda'(v) \leq 1$.

Note that F_1 is less disperse than F_2 if F_1 is a scaled down version of F_2 . More generally, fix F_1 and write F_2 as $F_2(v) = F(v/\bar{v}_2)$, where F is some distribution function with support $[0, 1]$. Then, if F_2 is scaled up sufficiently much (\bar{v}_2 is sufficiently

⁷See Shaked and Shantikumar (2007) for a review of the star order and other stochastic orders of spread, such as the dispersive order (defined below). See Kirkegaard (2011a) for an economic interpretation of these orders. Hopkins (2007) considers first price auctions in which one bidder’s distribution is more dispersed than that of his rival, but where first order stochastic dominance does not apply. See Lu (2010) for another use of the dispersive order in an auction design problem.

⁸This model is isomorphic to one in which bidders draw valuations from the same distribution but have different constant marginal costs of bidding.

large) it must be the case that F_1 is smaller than F_2 in the dispersive order. In a sense, dispersion is almost automatic if the asymmetry is sufficiently large.

In a standard all-pay auction bidder i must decide whether to participate in the auction, and, if so, which non-negative bid, b_i , to submit, $i = 1, 2$. The bidder with the highest bid would then be the winner. Here, however, bidders receive differential treatment. It is convenient to think of bidder i as accumulating a “score”, s_i . If bidder i bids b , his score is $s_i = a_i + r_i b$, where $a_i \geq 0$, $r_i > 0$. Hence, bidder i must decide whether to participate, and, if so, which score to aim for. The winner is the bidder with the highest score. To ensure the existence of an equilibrium, it is assumed that bidder j wins if $a_i > a_j$ and $s_1 = s_2 = a_i$. The tie-breaking rule is inconsequential in all other cases.

Bidder 1 is said to have a head start if $a \equiv a_1 - a_2 > 0$, in which case he wins the auction if both bidders bid zero. He is said to be handicapped if $r \equiv r_1/r_2 < 1$.

Turning to payoffs, bidders are risk neutral and the true cost of a bid of b is b . These assumptions are standard in the auction literature, but a more general treatment would allow for risk aversion and costs that are non-linear in the bid. The cost of obtaining a score of $s \geq a_i$ for bidder i is

$$c_i(s) = \frac{s - a_i}{r_i}. \quad (1)$$

2.1 Equilibrium allocation

Amann and Leininger (1996) characterize the equilibrium and equilibrium allocation of the all-pay auction without head starts and handicaps. It is straightforward to extend the characterization to allow for head starts and handicaps.

To begin, let $k(v)$ denote the function that is implicitly defined by

$$\int_{k(v)}^{\bar{v}_1} \frac{f_1(x)}{x} dx = r \int_v^{\bar{v}_2} \frac{f_2(x)}{x} dx \quad (2)$$

for $v \in [0, \bar{v}_2]$. Note that $k(v)$ is strictly increasing and satisfies $k(\bar{v}_2) = \bar{v}_1$. Moreover, the right and left hand sides both converge to ∞ as v and k , respectively, converge to zero. Hence, $k(0) = 0$ regardless of r . As described in Section 3, this feature of the model simplifies the analysis somewhat.

In equilibrium, bidder 1 with type $k(v)$ obtains the same score as bidder 2 with type v in the absence of head starts. The expression in (2) is easily derived by following the steps in Amann and Leininger (1996), and incorporating handicaps.

The details of how the equilibrium candidate is constructed are therefore omitted (it is verified below that there is no incentive to deviate).⁹

The function $k(v)$ does not depend on a_1 or a_2 . However, it is only valid for those types that submit strictly positive bids. The next step is to identify the types that do not submit strictly positive bids. For the sake of exposition, assume that bidder 1 is the bidder with the head start, $a \geq 0$. Then, bidder 2 may decide to simply bid zero, or equivalently, to not participate in the auction at all. Similarly, bidder 1 may decide to rely on the head start only, and just bid zero (score a_1).

If bidder 2 bids, he must submit a bid of at least $c_2(a_1)$, since any score below a_1 will be unsuccessful. It will never be profitable for bidder 2 to participate if $c_2(a_1) \geq \bar{v}_2$. In the following it is assumed that $c_2(a_1) \in [0, \bar{v}_2)$. Solving

$$vF_1(k(v)) - c_2(a_1) = 0, \quad (3)$$

yields the critical type of bidder 2 who is indifferent between staying out of the auction and entering the auction with a score of a_1 . Let the solution be denoted by v_2^c , and define $v_1^c \equiv k(v_2^c)$.

The main properties of the equilibrium allocation can now be identified. In (v_2, v_1) space, Figure 1 depicts $v_1 = k(v_2)$ as well as the level curve on which $v_2F_1(v_1)$ is constant and equal to $c_2(a_1)$. The former is increasing, by (2), while the latter is decreasing. The intersection of the two satisfies (3) and thus defines v_1^c and v_2^c .

In equilibrium, bidder 2 stays out of the auction if his type is strictly below v_2^c , and enters with a bid of $c_2(a_1)$ (score of a_1) if his type is precisely v_2^c . If his type is higher, he enters the auction and obtains a score equal to that obtained by bidder 1 with type $k(v)$. Bidder 1 enters the auction regardless of his type, but he submits a bid of zero, thereby obtaining a score of a_1 , if his type is v_1^c , or below. If his type is higher he achieves a score to rival that of bidder 2 with type $k^{-1}(v)$. In summary, bidder 2 wins only if his type is above v_2^c and bidder 1's type is below $k(v)$. Bidder 1 wins otherwise.

For completeness, the strategies that support the allocation outlined above are described in the following Proposition, in which it is also proven that they form an equilibrium. However, the analysis to follow is framed in terms of the equilibrium allocation (as described in Figure 1) rather than equilibrium scoring strategies (as described in Proposition 1). An earlier version of the paper describes how strategies change with head starts and handicaps.

⁹Briefly, the first order conditions produce a system of differential equations. This system is autonomous, due to the linearity assumptions (risk neutrality, linear costs). Using this property, Amann and Leininger (1996) show that (2) is the unique solution to the system of differential equations with the appropriate boundary condition. The boundary condition is $k(\bar{v}_2) = \bar{v}_1$; both bidders score the same if they have their highest respective types. See Kirkegaard (2008) for details.

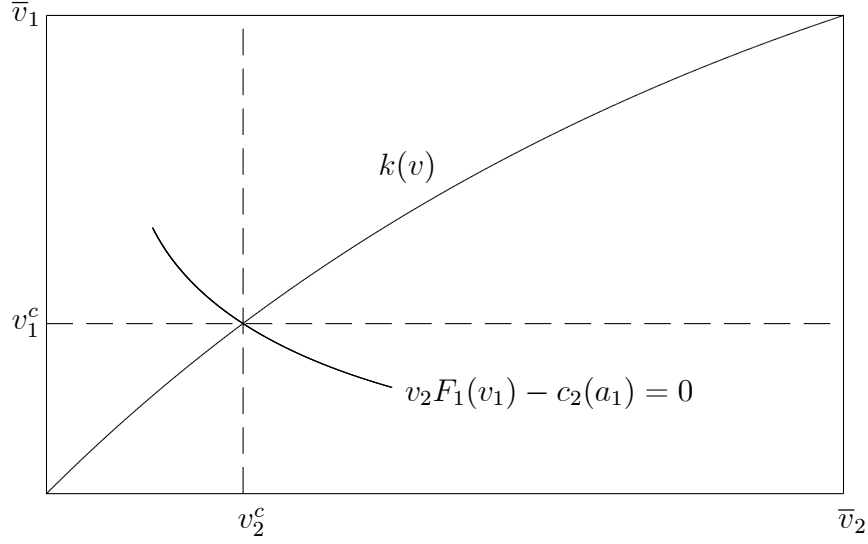


Figure 1: The equilibrium allocation. Bidder 2 wins below $k(v)$, to the right of v_2^c . Bidder 1 wins everywhere else.

Proposition 1 *Assume that bidder 1 has a head start and that $c_2(a_1) \in [0, \bar{v}_2]$. Then,*

$$s_1(v) = \begin{cases} a_1 & \text{if } v \in [0, v_1^c] \\ a_1 + \int_{v_1^c}^v r_2 k^{-1}(x) f_1(x) dx & \text{otherwise} \end{cases}$$

$$s_2(v) = \begin{cases} 0 & \text{if } v \in [0, v_2^c] \\ a_1 + \int_{v_2^c}^v r_1 k(x) f_2(x) dx & \text{otherwise} \end{cases}$$

form equilibrium scoring strategies for bidder 1 and bidder 2, respectively.

Proof. See the Appendix. ■

Note that $k(v)$ is decreasing in r , by (2). That is, as expected bidder 2 will win less often when his handicap increases. For a fixed value of r , the following Lemma “quantifies” k by comparing it to λ . Recall that $r\bar{\lambda} < 1$ if $r = 1$.

Lemma 1 *If $r\bar{\lambda} < 1$ then $k(v) > \lambda(v)$ for all $v \in (0, \bar{v}_2)$. If $r\bar{\lambda} > 1$ then $k(v) < \lambda(v)$ for all $v \in (0, \bar{v}_2)$.*

Proof. See the Appendix. ■

Lemma 1 implies that bidder 2 wins with a probability that exceeds his rank, $F_1(k(v)) > F_1(\lambda(v)) = F_2(v)$, unless r is large. Thus, in the standard all-pay auction, bidder 2 is ex ante more likely to win the auction than bidder 1.

As a starting point for the analysis of optimal head starts and handicaps, observe that the allocation can be manipulated by manipulating the two curves in Figure 1. First, k can be made to move down by increasing $r = r_1/r_2$. Second, the “level curve” can be moved to the right by increasing $c_2(a_1) = (a_1 - a_2)/r_2$. Note that $c_2(a_1)$ measures the cost for bidder 2 of nullifying bidder 1’s head start; it measures bidder 1’s head start in “real” terms. In summary, the “relative handicap”, r , and the “real” head start, $c_2(a_1)$, are all that matters for the equilibrium allocation. For convenience, define $\bar{a} \equiv c_2(a_1)$ as bidder 1’s real head start.

3 Revenue maximization

In most of the paper it will be assumed that \bar{a} and r are chosen with the objective of maximizing total expenditures. This is a reasonable objective if the designer or seller is the recipient of the expenditures, in which case it is equivalent to revenue maximization (the seller is assumed to put zero value on the prize).

The results in this paper are best understood by appealing to mechanism design. To begin, let $q_i(v)$ denote bidder i ’s equilibrium probability of winning if his type is v , $i = 1, 2$. When bidder 1 is given a head start these are

$$q_1(v) = \begin{cases} F_2(v_2^c) & \text{if } v \in [0, v_1^c) \\ F_2(k^{-1}(v)) & \text{otherwise} \end{cases}, \quad (4)$$

and

$$q_2(v) = \begin{cases} 0 & \text{if } v \in [0, v_2^c) \\ F_1(k(v)) & \text{otherwise} \end{cases}, \quad (5)$$

respectively. Myerson (1981) defines $J_i(v) = v - h_i(v)$ as bidder i ’s *virtual valuation*. With this key term, he then proves that ex ante expected gross surplus can be decomposed into ex ante expected expenditures and ex ante expected net surplus,

$$\underbrace{\int_0^{\bar{v}_i} v q_i(v) f_i(v) dv}_{\text{bidder } i\text{'s gross surplus}} = \underbrace{\int_0^{\bar{v}_i} J_i(v) q_i(v) f_i(v) dv}_{\text{bidder } i\text{'s expenditures}} + \underbrace{\int_0^{\bar{v}_i} h_i(v) q_i(v) f_i(v) dv}_{\text{bidder } i\text{'s net surplus}}.^{10} \quad (6)$$

¹⁰This formulation is accurate only if bidders earn zero payoff if their type is the lowest possible. This is the case in all the mechanisms studied in this paper. The reason is the assumption that the lowest possible type is zero. Without this assumption, the weaker bidder would stay out of the standard all-pay auction with positive probability, and the strong bidder would earn excessive rent if his type is the lowest possible. Then, there is an additional reason to disadvantage the strong bidder, namely to eliminate some of his rent. To rule out this confounding incentive and simplify the exposition I impose the assumption that the lowest type is zero.

It follows that the expected total expenditures in the auction is,

$$ER(\bar{a}, r) = \int_0^{\bar{v}_1} J_1(v)q_1(v)f_1(v)dv + \int_0^{\bar{v}_2} J_2(v)q_2(v)f_2(v)dv, \quad (7)$$

which equals the expected value of the winner's virtual valuation.

By assumption, $J_i(v)$ is strictly increasing. Let v_i^* denote the unique value of v for which $J_i(v) = 0$, $i = 1, 2$. Let $\kappa(v)$ denote the strictly increasing function satisfying $J_1(\kappa) = J_2(v)$, whenever it exists. Since $f_1(0) > f_2(0)$ implies that $J_1(0) > J_2(0)$, it must be the case that $\kappa(v) = 0$ for some $v > 0$. Moreover, since $J_2(\bar{v}_2) = \bar{v}_2 > \bar{v}_1 = J_1(\bar{v}_1)$, $\kappa(v) = \bar{v}_1$ for some $v < \bar{v}_2$.

As a point of comparison to the all-pay auction, consider the revenue maximizing mechanism (among mechanism where the good is sold with probability one). In an optimal mechanism bidder 1 should win if his type exceeds κ when his rival has type v , in which case $J_1 > J_2$. Otherwise he should lose. Such a rule maximizes the expected value of the virtual valuation of the winner by ensuring that the winner is the bidder with the highest virtual valuation. Bulow and Roberts (1989) and Bulow and Klemperer (1996) discuss the similarities to optimal monopoly pricing.

Figure 2 compares an optimal mechanism and the standard all-pay auction. The standard all-pay auction is not optimal; $k(v)$ does not coincide with $\kappa(v)$. In particular, bidder 2 wins more often than is optimal “near the bottom” (when both bidders have low types), but not often enough “near the top” (when both have high types).

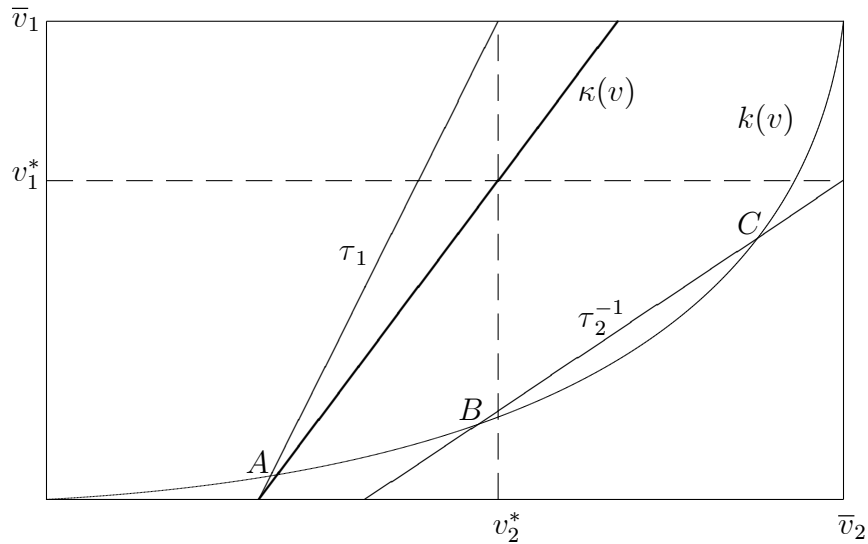


Figure 2: The optimal mechanism ($\kappa(v)$) versus the all-pay auction ($k(v)$).

At the top, virtual valuations coincide with types, or $J_i(\bar{v}_i) = \bar{v}_i$. This also occurs in the standard monopoly pricing problem, where marginal revenue on the first unit coincides with consumers' value of that unit. Since $J_2(\bar{v}_2) = \bar{v}_2 > \bar{v}_1 = J_1(\bar{v}_2)$, the strong bidder should win with probability one in an optimal mechanism for a mass of types close to \bar{v}_2 . However, this is inconsistent with equilibrium behavior in most auctions, including the all-pay auction, since the weak bidder has an incentive to overbid the mass of types. Indeed, $k(v)$ is not flat at the top. A similar problem arises at the bottom, where the weak bidder is supposed to outbid a mass of the strong bidder's low types. This would require the strong bidder to refrain from bidding if his type is sufficiently low. However, the strong bidder, who faces relatively weak competition, participates with probability one in equilibrium.

Now, the purpose of changing \bar{a} and r is to manipulating the allocation to bring it closer to the optimal allocation. However, these affine manipulations of bids cannot implement the optimal auction. For example, k is not horizontal at the top. The focus on linear scoring functions is thus not without loss of generality. On the other hand, linear scoring rules are perhaps easier to understand and use in the real world. See Nti (2004) for a discussion of linear scoring rules under complete information.

Sections 4 and 5, respectively, examine head starts and handicaps in isolation. Section 6 considers the simultaneous use of head starts and handicaps. Unfortunately, the three scenarios present different technical challenges and for that reason require three distinct methods of proof. As a consequence, it is necessary to impose different assumptions in each of the three sections (Section 4 being the only section where the analysis is completely general). In Section 5 and 6, assumptions are thus imposed on the "shape" of the asymmetry – or more precisely on the shape of $\lambda(v)/v$ – and on the "size" of the asymmetry. The purpose of Section 7 is to bring it all together by presenting environments that satisfy all the various assumptions simultaneously and by so doing unifies the analysis. Section 8 considers two environments in which some of the assumptions imposed so far are violated.

4 Head starts

In this subsection, r is assumed to be fixed.

Define $\tau_i(v)$ as the type (if any) that satisfies

$$E [J_i(x)|x \leq \tau_i(v)] = J_j(v), \tag{8}$$

where

$$E [J_i(x)|x \leq \tau_i(v)] = \int_0^{\tau_i(v)} J_i(x) \frac{f_i(x)}{F_i(\tau_i(v))} dx,$$

for $i, j = 1, 2, i \neq j$. For example, if bidder 1 receives a head start, $E[J_1(x)|x \leq v_1^c]$ measures the expected virtual valuation among the types that score a_1 . The significance of $E[J_1(x)|x \leq v_1^c]$ is precisely that it summarizes this set of types, $[0, v_1^c]$.

Since the left hand side of (8) is zero when $\tau_i = \bar{v}_i$ it must be the case that $\tau_i(v_j^*) = \bar{v}_i$. Note also that τ_1 and κ coincide on the horizontal axis, but $\tau_1(v) > \kappa(v)$ whenever τ_1 is defined and $\kappa > 0$. Thus, τ_1 and k must cross. Turning to τ_2 , it can also be verified that $\tau_2(v) > \kappa^{-1}(v)$ whenever both are defined. The following Lemma “quantifies” τ_2 by comparing it to λ , which turns out to be relevant once a head start to the strong bidder is contemplated (see Theorem 2, below). Figure 2 illustrates τ_1 and the inverse of τ_2 .

Lemma 2 $\lambda > \tau_2^{-1}$ whenever τ_2^{-1} is defined.

Proof. See the Appendix. ■

Compared to the optimal mechanism, one of the drawbacks of the all-pay auction is that bidder 2 wins too often when types are small. A head start to bidder 1 rectifies this problem. Hence, the allocation will move closer to what is optimal.

Theorem 1 *Regardless of r , it is always optimal to employ a head start. In particular, regardless of r , it is profitable to give bidder 1 a head start. When it is optimal to give bidder 1 a head start, then the optimal head start is such that $\tau_1(v_2^c) = k(v_2^c)$.*

Proof. Given (7), observe that

$$\frac{\partial ER(\bar{a}, r)}{\partial \bar{a}} = f_2(v_2^c)F_1(v_1^c) \frac{\partial v_2^c}{\partial \bar{a}} \left(\int_0^{v_1^c} J_1(v) \frac{f_1(v)}{F_1(v_1^c)} dv - J_2(v_2^c) \right), \quad (9)$$

the sign of which is determined by the term in parenthesis (v_2^c is increasing in \bar{a}). As \bar{a} approaches 0, v_1^c and v_2^c approaches 0 and this term converges to

$$-\frac{1}{f_1(0)} + \frac{1}{f_2(0)},$$

by L’Hôpital’s rule. This is positive, by assumption. Hence, $ER(\bar{a}, r)$ is strictly increasing in \bar{a} when \bar{a} is small. Thus, it is optimal to use a head start since a head start to the weak bidder, in particular, strictly increases expected revenue. The first order condition is satisfied when the term in parenthesis is zero, which occurs if and only if $v_1^c = \tau_1(v_2^c)$. ■

Consider a marginal increase in bidder 1’s head start. If the allocation changes, it is because bidder 2 won before, but now loses. Bidder 2’s type in this event is

v_2^c , while bidder 1 has a type below $k(v_2^c) \equiv v_1^c$. Thus, the virtual valuation of the winner changes from $J_2(v_2^c)$ to $E[J_1(x)|x \leq k(v_2^c)]$, in expectation. The optimal head start ensures the marginal loss and gain are equated, which necessitates that $\tau_1(v_2^c) = k(v_2^c)$. In Figure 2, the intersection of τ_1 and k (the point A) thus determines v_2^c and by extension \bar{a} . Conditions under which $ER(\bar{a}, r)$ is single-peaked in \bar{a} are discussed in Section 7. This is the case when τ_1 and k intersect only once.

Theorem 1 does not claim that a head start to the strong bidder is not profitable or even optimal. There are two points in Figure 2 where the first order conditions are satisfied for an optimal head start to bidder 2. Point B , which would require the smallest head start, is a local minimum, while point C is a local maximum. Thus, if a head start to bidder 2 is profitable, it must be a large head start. In contrast, any small head start to bidder 1 is profitable. Arguably, the seller needs less information about the distribution functions to profit from a head start to the weak bidder. Moreover, as the next result shows, a head start to the strong bidder can be profitable only if he is handicapped a lot. Thus, it can never be optimal to give the strong bidder a head start in the absence of a handicap. The intuition is that without a handicap the strong bidder is already winning too often at the bottom. Giving him a head start would only make matters worse.

Theorem 2 *Assume that $r\bar{\lambda} < 1$ (e.g. $r = 1$). Then, any head start to the strong bidder lowers expected revenue.*

Proof. By Lemma 1, $k(v) > \lambda(v)$ for all $v \in (0, \bar{v}_2)$, while Lemma 2 states that $\lambda(v) > \tau_2^{-1}(v)$, whenever the latter is defined. Thus, $k(v) > \tau_2^{-1}(v)$, meaning that the two never intersect (to get a crossing, as in Figure 2, the strong bidder must necessarily be severely handicapped). Switching the roles of bidders 1 and 2 in (9) then proves that expected revenue is decreasing in bidder 2's head start. ■

5 Handicaps

In this subsection, it is assumed that there is no head start, or $a = 0$. Hence, changes in r_1 and r_2 do not impact the real head start ($\bar{a} = 0$ for all r_1 and r_2).

Bidder 1 wins more often, regardless of his type, when r increases (that is, $k(v)$ moves down). This moves the allocation closer to the optimal allocation if his type is low, but farther away if his type is high. Hence, there is a trade-off.

Clark and Riis (2000) have proven that it is optimal to handicap the strong bidder ($r > 1$) when types are drawn from different uniform distributions. In general, however, the trade-off makes it difficult to predict which bidder should be handicapped.

Moreover, the approach used by Clark and Riis does not generalize. In short, they rely on closed-form expressions of bidding strategies that can be obtained in the uniform case, but not in the general case. Thus, a different method of proof is required.

In this section, Clark and Riis' (2000) result is generalized in two directions. First, their result is shown to hold whenever F_1 is a scaled down version of F_2 . Recall that Clark and Riis' (2000) uniform model is a special case. Second, the result also holds when the asymmetry is "sufficiently large", regardless of the shape of the asymmetry. Kirkegaard (2011a) recently proved that the first-price auction is more profitable than the second-price auction in these models. As in that paper, the proofs in the current paper are based on mechanism design. Section 8 provides an example in which it is optimal to handicap the weak bidder.

5.1 F_1 is a scaled down version of F_2

Assume F_1 is a scaled down version of F_2 . Let

$$\rho = r \frac{\bar{v}_1}{\bar{v}_2}, \quad (10)$$

which reflects the significance of the handicap once the relative strength of the two bidders is taken into account. Since $\bar{\lambda} = \underline{\lambda} = \bar{v}_1/\bar{v}_2$ in this model, Lemma 1 implies that a measure of symmetry is restored to the game if $\rho = 1$ (or $r = \bar{v}_2/\bar{v}_1 > 1$) since in this case bidders with the same rank win with the same probability, $k(v) = \lambda(v)$. It is convenient to think of ρ as the choice variable, rather than r .

Let $EP_i(\rho)$ denote bidder i 's ex ante expected payment, $i = 1, 2$. The proof of Lemma 3 establishes that $EP_i(\rho)$ is separable, such that it can be written $EP_i(\rho) = \bar{v}_i EP_i^s(\rho)$, where \bar{v}_i captures the strength of bidder i and $EP_i^s(\rho)$ is a "scale-adjusted" payment that filters out the bidder's strength. Hence, $EP_1^s(1) = EP_2^s(1)$ because, as mentioned, bidders with the same rank win with the same probability when $\rho = 1$; adjusting for their scale, the handicap has made bidders symmetric. The next Lemma describes other important features of the scale-adjusted payments.

Lemma 3 *Assume that F_1 is a scaled down version of F_2 , and $a = 0$. Then, $EP_2^{s'}(1) = -EP_1^{s'}(1) > 0$. Moreover, $EP_2^{s'}(\rho) > 0$ for all $\rho \in (0, 1]$ and $EP_1^{s'}(\rho) < 0$ for all $\rho \in [1, \infty)$.*

Proof. See the Appendix. ■

Expected revenue is

$$ER(\rho) = \bar{v}_1 EP_1^s(\rho) + \bar{v}_2 EP_2^s(\rho). \quad (11)$$

The main result of this section now follows easily from Lemma 3.

Theorem 3 Assume that F_1 is a scaled down version of F_2 , and $a = 0$. Then, expected revenue is maximized at some $\rho > 1$ (or, equivalently, $r > \frac{\bar{v}_2}{\bar{v}_1} > 1$).

Proof. See the Appendix. ■

Theorem 3 rules out that $ER(\rho)$ is maximized at any $\rho \leq 1$. However, it is not claimed that there is a unique global maximum.

The magnitude of the optimal handicap is striking. Since $\rho > 1$, bidder 2's winning probability now falls below his rank, or $k(v) < \lambda(v)$. Hence, bidder 2's ex ante winning probability is now less than $\frac{1}{2}$. In a sense, the weak bidder is overcompensated. See Section 9 for an interpretation in the context of affirmative action and Section 5.3 for a comparison to the complete information model.

From (11), the seller is maximizing a weighted average of the bidders' scale-adjusted payments. Since the weight on bidder 2 is the largest, the seller is willing to sacrifice some revenue from bidder 1 to more effectively milk bidder 2.

Figure 3 depicts the scale-adjusted payments in the uniform model. Recall that $r = 1$ is equivalent to $\rho = \frac{\bar{v}_1}{\bar{v}_2} < 1$, whereas the optimal value of ρ is larger than one, but to the left of the peak of EP_2^s . Clearly, bidder 2's expected payment increases when he is handicapped optimally. On the other hand, bidder 1's expected payment may increase or decrease, depending on how close to the peak of EP_1^s he would be without a handicap in place. This type of argument holds whenever $EP_2^s(\rho) < EP_2^s(1)$ for all $\rho \in (0, 1)$, which is guaranteed to hold, by Lemma 3.

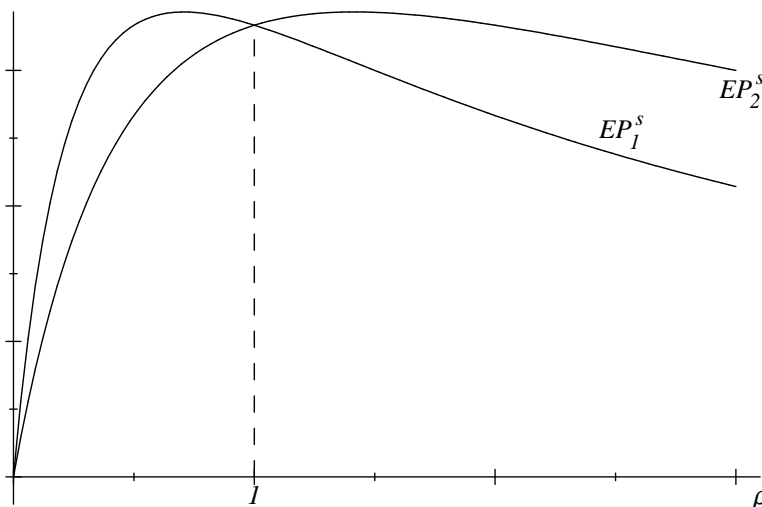


Figure 3: Scale-adjusted payments in the uniform model.

In conclusion, the expected payments of the handicapped bidder increases, while the change in the favored bidder's payment depends on how heterogeneous the bid-

ders are.¹¹ With these observations in mind, it is now possible to consider other types of asymmetry as well.

5.2 Large asymmetries

Figure 3 indicates that if the asymmetry is sufficiently large in the uniform model, then the expected payment of *both* bidders would decline if the weak bidder is handicapped. Thus, it is optimal to handicap the strong bidder. This argument turns out to extend beyond the model in Theorem 3.

Theorem 4 *Assume $\min_x f_1(x) \geq 2 \max_x f_2(x)$. Then, $ER(0, r)$ is strictly increasing in r for all $r \in (0, 1]$. Hence, it is optimal to handicap the strong bidder.*¹²

In the uniform model, the condition in Theorem 4 is satisfied if $\bar{v}_2 \geq 2\bar{v}_1$. Outside the uniform model, $\bar{v}_2 > 2\bar{v}_1$ is necessary but not sufficient. Although the asymmetry needed to invoke Theorem 4 may seem large, it should be noted that the method of proof is very demanding. Specifically, the derivative of *each* bidder's expected payment with respect to r is written as the integral of a function which is shown to be positive for *all* realizations, as long as $r \in (0, 1]$.

The condition in Theorem 4 implies that $f_1(x) \geq f_2(y)$ for all $x \in [0, \bar{v}_1]$ and $y \in [0, \bar{v}_2]$. Hence, $\lambda'(v) = f_2(v)/f_1(\lambda(v)) \leq 1$. Thus, F_1 is less disperse than F_2 .

5.3 Complete information

Fu (2006) examines the complete information model, which as mentioned earlier can be thought of as a special case of the model in Section 5.1. The difference is that the latter assumes densities are strictly positive everywhere. This difference is enough to change the results in important ways.

In Fu's (2006) model, both bidders maximize their expected payment when $\rho = 1$. Lemma 3 clearly states that this property does not hold in the incomplete information model. Thus, $\rho = 1$ is optimal in the complete information model, but not in the

¹¹It can be shown that both bidders have types that bid more aggressively with the handicap, and types that bid less aggressively. Nevertheless, although Clark and Riis (2000) examine aggregate revenue only, they claim that in their model it is profitable to handicap the strong bidder because it encourages the favored (weak) bidder to bid more aggressively. However, it is now clear that the expected payment from the favored bidder may decline. Similarly, an optimal head start increases the disadvantaged bidder's expected payment but decreases the favored bidder's expected payment.

¹²Moreover, if the strong bidder is handicapped only slightly (so that r remains close to one), the expected payment of both bidders increase.

incomplete information model. The addition of incomplete information means that the strong bidder must be handicapped even more severely (Theorem 3). While the two bidders would win equally often in the complete information model with the optimal handicap, the weak bidder wins more often when information is incomplete.

Section 8 describes two other conclusions that can be obtained in an incomplete information model, but never in a complete information model.

6 Head starts and handicaps

Theorems 1 and 2 imply that the weak bidder should be given a head start in the absence of a handicap. However, Section 5 suggests that it may be optimal to severely handicap the strong bidder, while Section 4 in turn suggests that large handicaps should perhaps be accompanied by a head start to the strong bidder. So, how should the seller optimally combine the two instruments?

The next result states that it is the weak bidder who will receive a head start when \bar{a} and r are determined jointly. The proof is by contradiction. In particular, it is shown that if a “locally optimal” head start is given to the strong bidder, then expected revenue can be improved further by lowering r , thus contradicting that the original combination of head starts and handicaps is optimal. Since this method of proof relies on studying the effects of changing r , some of the challenges that complicated the previous section once again come into play. However, if the problem is sufficiently “regular”, these challenges can be circumvented. Below, two sets of conditions on (F_1, F_2) are identified for which the result holds.

Theorem 5 *The revenue-maximizing combination of head starts and handicaps involves a head start to the weak bidder ($\bar{a} > 0$) if either condition (i) or (ii) is satisfied:*

(i) $\kappa(v)$ and $k(v)$ cross exactly once for all r such that $r\bar{\lambda} \geq 1$.

(ii) $\frac{\lambda(v)}{v}$ is non-decreasing and $\int_0^{v_2^*} J_2(x) \frac{f_2(x)}{F_2(v_2^*)} dx \leq J_1(0)$.

Proof. See the Appendix. ■

Note that κ and k depend only on F_1 and F_2 , for any given r . Thus, condition (i) is a condition on the primitives of the game. It is satisfied in the uniform model.¹³

¹³The domain of κ is a subset of $(0, \bar{v}_2)$, while the domain of k is $[0, \bar{v}_2]$. Both are increasing and range from 0 to \bar{v}_1 . Hence, the two functions must cross. In the uniform model, κ is linear whereas k is concave, linear, or convex, depending on the size of r . Hence, κ and k cross exactly once. For similar reasons, the assumption in Theorem 6, below, is also satisfied in the uniform model.

Turning to condition (ii), the first part states that F_2 is stochastically smaller than F_1 in the star order. Thus, it is an assumption on the “shape” of the asymmetry. The role of this assumption is to enable a comparison between (1) $\lambda(v)$ and $k(v)$ when r is high and (2) $\lambda(v)$ and $\kappa(v)$ when v is high. See the Appendix for details.

The second part of condition (ii) concerns the “size” of the asymmetry. The strong bidder must be sufficiently stronger than the weak bidder. For instance, if F_1 is a scaled down version of F_2 , the assumption can be written

$$\frac{\bar{v}_2}{\bar{v}_1} f(0) \frac{v^*(1 - F(v^*))}{F(v^*)} \geq 1, \quad (12)$$

where $v^* = \arg \max x(1 - F(x))$ is independent of \bar{v}_1 and \bar{v}_2 . Clearly, this assumption is satisfied when \bar{v}_2 is sufficiently large compared to \bar{v}_1 . For the uniform model, the assumption is satisfied if $\bar{v}_2 \geq 2\bar{v}_1$, as in Theorem 4.

When a head start is used to bring the allocation near the bottom closer to what is optimal, there is less of an incentive to handicap the strong bidder. Instead, it may be better to use the handicap to bring the allocation closer to what is optimal near the top. Recall that the problem near the top is that the strong bidder does not win often enough. Handicapping the weak addresses this problem. The next result shows that the weak bidder is either given a large head start and a handicap, or the double advantage of a moderate head start and a handicapped opponent. A “large” or “moderate” head start refers to one that exceeds or falls below, respectively, the head start the weak bidder would get without handicaps ($r = 1$). Let \bar{a}^* denote the optimal real head start without a handicap, and let \bar{a}^{**} and r^{**} denote the optimal real head start and handicap, respectively, when the two are determined jointly.

Once again, a regularity assumption is needed. The assumption is that expected revenue is single-peaked in \bar{a} when $r = 1$. The role of the assumption is to rule out large jumps in \bar{a} when r is changed a bit. The assumption is satisfied in the uniform model, for example, and discussed further in the next section.

Theorem 6 *Assume that the revenue-maximizing combination of head starts and handicaps involves a head start to the weak bidder. Assume also that $ER(\bar{a}, 1)$ is single-peaked in \bar{a} . Then, either (i) $\bar{a}^{**} > \bar{a}^*$ and $r^{**} < 1$ or (ii) $0 < \bar{a}^{**} \leq \bar{a}^*$ and $r^{**} \geq 1$.*

Proof. The first assumption implies that (v_2^c, v_1^c) is determined by the intersection of k and τ_1 . The second assumption is equivalent to the assumption that $k(v)$ and $\tau_1(v)$ intersect exactly once when $r = 1$. Hence, if $r < 1$ (k shifts up), k must intersect τ_1 to the right of the intersection for $r = 1$. Consequently, if $r < 1$ then v_2^c

increases, which necessitates $\bar{a}^{**} > \bar{a}^*$. A similar argument proves that v_2^c decreases if $r \geq 1$, which necessitates $\bar{a}^{**} \leq \bar{a}^*$. ■

The intuition behind Theorem 6 is explained further in the next section. This section concludes with an example.

Example 1: Consider the uniform model; $F_i(v) = \frac{v}{\bar{v}_i}$, $v \in [0, \bar{v}_i]$, $i = 1, 2$, with $\bar{v}_2 > \bar{v}_1 = 1$. By Theorems 1 and 2, $\bar{a}^* > 0$ if the seller can use head starts only. By Theorem 3, $r^* > 1$ if the seller can use handicaps only. As explained in the next section, the first part of Theorem 6 is more likely to apply the larger the asymmetry is. For example, if the seller can use both instruments then part (ii) of Theorem 5 holds if $\bar{v}_2 = 2$, but (i) holds if \bar{v}_2 increases to $\bar{v}_2 = 3$.¹⁴ In the latter case, the weak bidder is given a head start and a handicap. When $\bar{v}_2 = 3$, the use of the optimal handicap (and no head start) increases expected revenue by 42.5%. On the other hand, the use of the optimal head start (and no handicap) increases expected revenue by 58.8%. With both instruments, expected revenue increases only a bit more, specifically by 59.1% in total. Thus, the head start is responsible for most of the increase. □

7 Beyond the uniform model

A variety of assumptions were imposed in Sections 5 and 6, as the need arose. The purpose of this section is to describe environments in which the results can be unified. Recall first the properties of the uniform model, illustrated in Example 1.

Corollary 1 *The uniform model has the following properties:*

1. *If only one instrument can be used, then it is optimal to favor the weak bidder.*
2. *If both instruments can be used, then it is optimal to give the weak bidder a head start. In addition, either (i) $\bar{a}^{**} > \bar{a}^*$ and $r^{**} < 1$ or (ii) $0 < \bar{a}^{**} \leq \bar{a}^*$ and $r^{**} \geq 1$.*

Proof. The first part follows directly from Theorems 1 – 3. The second part invokes Theorems 5 and 6. As explained in the text, Theorems 5 and 6 apply to the uniform model because k intersects κ and τ_1 precisely once, regardless of r . ■

The question is now whether the properties in Corollary 1 extend beyond the uniform model. In short, I will argue that this is the case whenever the asymmetry

¹⁴The details of the examples in the paper are omitted, but are available on request.

between bidders is “sufficiently large”. Moreover, in some environments it is possible to quantify the size of the asymmetry that is required.

To begin, any given distribution function F_2 can be written as

$$F_2(v) = F\left(\frac{v}{\bar{v}_2}\right), \quad v \in [0, \bar{v}_2], \quad (13)$$

where F is some, appropriately chosen, distribution function with support $[0, 1]$. Now, fix a distribution F instead, and consider the class of distribution functions that can be obtained from F by scaling it, as in (13).¹⁵ Then, holding F_1 fixed, it turns out that $ER(\bar{a}, 1)$ is single-peaked in \bar{a} if F_2 is obtained by “scaling up” F sufficiently much (that is, if \bar{v}_2 is sufficiently large). Roughly speaking, the problem is automatically well-behaved if the stakes are high enough for the strong bidder.¹⁶

Lemma 4 *Let $F_2(v) = F\left(\frac{v}{\bar{v}_2}\right)$, $v \in [0, \bar{v}_2]$. Then, $ER(\bar{a}, 1)$ is single-peaked in \bar{a} , $\bar{a} \geq 0$, if \bar{v}_2 is sufficiently large.*

Proof. See the Appendix. ■

Lemma 4 and Theorems 1 and 2 imply that there is a unique optimal head start when neither bidder is handicapped and the asymmetry is sufficiently large. It is bidder 1 who is given the head start. However, the primary use of Lemma 4 is that it permits Theorem 6 to be invoked. To prove Corollary 1, Theorem 3 and the first part of Theorem 5 were used. However, neither of these can necessarily be relied upon in the general case. Thus, the intention is to use Theorem 4 and the second part of Theorem 5 in the attempt to generalize Corollary 1.

Clearly, $\max_x f_2(x) = \max_x f\left(\frac{x}{\bar{v}_2}\right) / \bar{v}_2$ is decreasing in \bar{v}_2 . Thus, the condition in Theorem 4 is satisfied when \bar{v}_2 is large. As for the second part of Theorem 5, the inequality in the second part of the condition is also easier to satisfy the larger \bar{v}_2 is. Finally, if $\lambda(v)/v$ is non-decreasing for some (F_1, F_2) pair, then it is also non-decreasing if F_2 is re-scaled (as proven below). Corollary 1 can now be generalized.

Corollary 2 *Let $F_2(v) = F\left(\frac{v}{\bar{v}_2}\right)$, $v \in [0, \bar{v}_2]$. Assume F is smaller than F_1 in the star order. Then, the contest has the same properties as in the uniform model when \bar{v}_2 is sufficiently large.*

¹⁵Assume F has all the properties described in Section 2 (positive and finite density, increasing hazard rate). Note that F_2 first order stochastically dominates F_1 if \bar{v}_2 is large enough.

¹⁶Figure 2 provides some intuition. As \bar{v}_2 increases, k becomes flatter. In contrast, the slopes of τ_1 and κ are bounded below. Thus, k intersects τ_1 and κ precisely once when \bar{v}_2 is large.

Proof. The first part follows from Theorems 1, 2, and 4. To use the second part of Theorem 5, note that

$$\frac{d}{dv} \frac{\lambda(v)}{v} = \frac{f_2(v)v - f_1(\lambda(v))\lambda(v)}{f_1(\lambda(v))v^2} = \frac{f\left(\frac{v}{\bar{v}_2}\right) \frac{v}{\bar{v}_2} - f_1(\lambda(v))\lambda(v)}{f_1(\lambda(v))v^2}. \quad (14)$$

By assumption, (14) is non-negative for all $v \in [0, \bar{v}_2]$ when $\bar{v}_2 = 1$. Since $F\left(\frac{v}{\bar{v}_2}\right) = F_2(v) = F_1(\lambda)$, λ depends only on the ratio $v/\bar{v}_2 \in [0, 1]$. Thus, the numerator in (14) is independent of \bar{v}_2 when it is evaluated at a constant rank (or v/\bar{v}_2 ratio). In other words, the sign of (14) is unaffected by the change in \bar{v}_2 and so $\lambda(v)/v$ remains non-decreasing when F_2 is rescaled. Thus, Theorem 5 applies when \bar{v}_2 is sufficiently large. Then, Lemma 4 can be used to invoke Theorem 6. ■

A special case is examined next. Start with the benchmark uniform model. Then, keeping the supports fixed, one way to increase the asymmetry is by moving the two distributions away from each other by “curving” them in opposite directions, making F_1 concave and F_2 convex. Analytically, the advantage is that the size of the asymmetry required to invoke Theorems 4 – 6 can be quantified.

Corollary 3 *Assume F_1 is concave and F_2 is convex. Then, the contest has the same properties as in the uniform model when*

$$\frac{f_1(\bar{v}_1)}{2} \geq f_2(\bar{v}_2) \text{ and } \frac{\bar{v}_2 - \bar{v}_1}{\bar{v}_1^2} \geq f_2(\bar{v}_2).^{17}$$

Proof. See the Appendix. ■

To appreciate the advantages of handicapping the weak bidder while simultaneously giving him a head start, note that in the limit as $r \rightarrow 0$ the weak bidder is handicapped so much that he will not submit positive bids. Thus, he will score a_1 . Then, from the strong bidder’s point of view, \bar{a} functions as a take-it-or-leave-it price. The strong bidder would then win if his type is above \bar{a} , and lose otherwise. If \bar{a} is chosen judiciously, this mechanism maximizes the payment that is obtainable from the strong bidder. If he is very strong compared to the weak bidder, it is intuitive that it is worthwhile sacrificing revenue on the weak bidder (who pays nothing in the limiting case) to get more out of the strong bidder.¹⁸

¹⁷The last condition can be replaced by the weaker condition that $(\bar{v}_2 - \bar{v}_1)/\bar{v}_1^2 \geq f_2(v_2^*)$. In the uniform model, the second condition is automatically satisfied if the first condition is satisfied. More generally, if the two conditions are satisfied for some pair of distributions, then they are also satisfied if F_2 is scaled up further.

¹⁸For instance, if F_1 is a scaled down version of F_2 , (11) implies that only the strong bidder’s payment is relevant as $\bar{v}_1 \rightarrow 0$ or $\bar{v}_2 \rightarrow \infty$.

As a final example, consider the case where $v_2 = \bar{v}_2$ with probability one (F_2 is degenerate). Then, $\bar{a}^{**} = \bar{v}_2$, $r^{**} = 0$. Since $r^{**} = 0$ the weak bidder has no incentive to bid. It is then optimal for the strong bidder to bid $\bar{a}^{**} = \bar{v}_2$ and win with probability one (given the tie-breaking rule). Here, social surplus is maximized, but bidders obtain zero payoff. Hence, there is no better mechanism from the seller's point of view. Of course, the complete information model (in which F_1 is also degenerate) is a special case. Thus, the complete information model yields the same qualitative conclusion regarding the simultaneous use of both instruments as the incomplete information model with large asymmetries. However, Example 1 shows that things are different when the asymmetry is small.

8 Irregular distributions

This section considers two environments that violate the regularity assumptions imposed so far. In one, it is profitable to handicap the weak bidder in an asymmetric contest. In the other, it is profitable to give a head start to a random bidder in a symmetric contest.

Note that neither conclusion could be obtained in a complete-information model. It follows from Fu (2006) that handicapping the weak bidder in a complete information model unambiguously decreases expected revenue. If bidders are symmetric and information is complete, an unbiased all-pay auction maximizes expected revenue. Thus, a head start to either bidder can never be optimal. This result can also be verified more directly by using the equilibrium characterization in Konrad (2002).

8.1 The weak bidder is potentially uninterested

Inspired by an example in Maskin and Riley (2000), consider the possibility that the weak bidder may not be interested in the prize at all.

Definition 4 *Bidder 1 is potentially uninterested if there exists some $\alpha \in (0, 1)$ such that*

$$F_1(v) = 1 - \alpha + \alpha F_2(v), \quad v \in [0, \bar{v}_2].$$

Maskin and Riley (2000) prove that the second-price auction is more profitable than the first-price auction in a slightly more general model. If bidder 1 is potentially uninterested, then he is believed to be like bidder 2 with probability α , but to be uninterested in the prize with probability $1 - \alpha$. This model is different from the model described in Section 2, due to the mass point $(1 - \alpha)$, the coincidence of \bar{v}_1

and \bar{v}_2 , and the fact that $f_1(0) < f_2(0)$. It is readily checked that $J_1(v) = J_2(v)$, or $\kappa(v) = v$, for all $v \in (0, \bar{v}_2]$. This is the critical feature of this class of models.

The derivation of equilibrium in Section 2 remains valid, meaning that

$$\int_{k(v)}^{\bar{v}_2} \frac{f_2(x)}{x} dx = \frac{r}{\alpha} \int_v^{\bar{v}_2} \frac{f_2(x)}{x} dx. \quad (15)$$

When $r = 1$, $k(v) < v$ for $v \in (0, \bar{v}_2)$. Thus, in the absence of head starts and handicaps, bidder 1 is more aggressive than bidder 2 for comparable types. However, this outcome is unequivocally negative in the current model, since it implies that $k(v) < v = \kappa(v)$, for all $v \in (0, \bar{v}_2)$. Hence, bidder 1 wins too often compared to what is optimal, regardless of his type. However, $k(v) = v = \kappa(v)$ when $r = \alpha < 1$. Handicapping the weak bidder involves no trade-off. Incidentally, such a handicap is efficient as well.

Proposition 2 *Assume that $a = 0$. The seller profits from handicapping the weak bidder ($r < 1$) if he is potentially uninterested. The optimal value of r is $r^* = \alpha$.*

Proof. In the text. ■

8.2 A symmetric model

Assume now that the two bidders are symmetric. They both draw types from some distribution function F , which is assumed to have no mass points. However, contrary to the assumption in Section 2, the virtual valuation is not monotonic. In particular, assume that $J'(0) < 0$. In such a case, Myerson (1981) prescribes “ironing” virtual valuations by using a non-deterministic mechanism. However, as Bulow and Roberts (1989) note, ironing can also be achieved in a symmetric model by treating bidders in an asymmetric, but deterministic, fashion. For this reason, it will be profitable to single out some random bidder for a small head start before bidding commences.

Proposition 3 *Assume bidders are symmetric and $J'(0) < 0$. When $r = 1$, it is profitable to give some random bidder a small head start.*

Proof. Let bidder 1 be the beneficiary of the head start. Since bidders are symmetric and $r = 1$, $v_1^c = v_2^c = v^c$. By (9),

$$\frac{\partial ER(\bar{a}, r)}{\partial \bar{a}} = f(v^c)F(v^c) \frac{\partial v^c}{\partial \bar{a}} \left(\int_0^{v^c} (J(v) - J(v^c)) \frac{f(v)}{F(v^c)} dv \right),$$

which is positive when v^c is small but positive. The reason is that since $J'(0) < 0$, there is a $v^c > 0$ such that $J(v) - J(v^c) > 0$ for all $v \in [0, v^c)$. ■

Note that $J'(0) < 0$ whenever F is “very concave” near the origin. A common example is $F(v) = v^\gamma$, $v \in [0, 1]$ where $\gamma \in (0, 1)$. The next example illustrates that non-monotonic virtual valuations may arise quite naturally in all-pay auctions. Thus, ironing provides another possible justification for the use of head starts.

Example 2: Bidders have types with two components, v and c , where c^{-1} measures the marginal cost of bidding or effort. If a bidder who bids b wins with probability $q(b)$, his expected payoff is $vq(b) - b/c$, which is maximized where $vcq(b) - b$ is maximized. Thus, the bidder’s type is essentially vc . It is usually assumed that either v or c is common knowledge (and identical for all bidders). If the one private component is drawn from a uniform distribution, say, then virtual valuations are monotonic. Now, consider the case where both components are private information and both are drawn independently from a uniform distribution on $[0, 1]$. The distribution of $x = vc$ is now determined by the multiplicative convolution of two well-behaved uniform distributions, or

$$F(x) = \int_0^1 \min\left\{\frac{x}{v}, 1\right\} dv = \int_0^x dv + \int_x^1 \frac{x}{v} dv = x(1 - \ln x),$$

which has a U-shaped virtual valuation. □

9 Bidders’ payoffs

In the preceding analysis no weight was assigned to the welfare of the bidders. The starting point of this Section is the extreme opposite assumption; the only concern is bidders’ ex ante expected payoffs. This is a natural concern in contests in which effort is pure waste and no-one benefits from it. In the remainder of the paper, assume the assumptions that were described in Section 2 are satisfied.

Let EU_i denote bidder i ’s ex ante expected payoff, $i = 1, 2$. I will compare two commonly used welfare functions, the utilitarian welfare function,

$$H = EU_1 + EU_2,$$

and the Rawlsian welfare function,

$$W = \min\{EU_1, EU_2\}.$$

If payments or transfers from one bidder to another is permitted, the mechanism suggested by d’Aspremont and Gérard-Varet (1979) can be used to allocate the prize

efficiently and maximize H subject to budget balancedness. However, monetary transfers are not necessarily possible in a contest setting where bidders expend effort rather than money. In any event, the purpose of the current paper is to examine the use of head starts and handicaps in the context of a contest. For these reasons, attention will be restricted to mechanisms or contests in which $U_i(0) = 0$, $i = 1, 2$, i.e. a bidder with type zero earns zero expected payoff. If $U_i(0) < 0$, the bidder would not want to participate if his type is low (as Myerson and Satterthwaite (1983) point out, the participation constraint may be violated in d'Aspremont and Gérard-Varet's (1979) mechanism). Likewise, since a bidder with type zero places zero value on winning, $U_i(0) > 0$ is possible only if the bidder receives a transfer from the other bidder or from a third party.

In the class of mechanisms in which $U_i(0) = 0$, $i = 1, 2$, the two welfare functions give rise to different results and policy recommendations, as the next result demonstrates. The first part is general, and does not rely on the regulation of the contests taking the form of a head start and/or a handicap.

Proposition 4 *Consider the class of mechanisms in which $U_i(0) = 0$, $i = 1, 2$. Within this class, it holds that: (i) when H is maximized, W is minimized; $W = EU_1 = 0$, and (ii) H can be maximized by giving bidder 2 a real head start that exceeds \bar{v}_1 .¹⁹*

Proof. See the Appendix. ■

Bidder 2 is better off than bidder 1, ex ante, in the standard all-pay auction, or $EU_2 > EU_1$ (formally, this is a consequence of Lemma 1). Thus, in sharp contrast to the utilitarian case, bidder 1 must be made better off if W is to increase.

Consider first the possibility that W is increased by using only one instrument; bidder 1 is given a head start or bidder 2 is handicapped. The next result is a follow-up to Proposition 4. If one instrument is used in moderation to increase W , then H must decrease. That is, bidder 2 is hurt more than bidder 1 benefits. The second part of the Proposition assumes that F_2 is more dispersed than F_1 .

Proposition 5 *(i) If bidder 1 gets a head start, then W is increasing in \bar{a} for small values of \bar{a} , but decreasing in \bar{a} for large values of \bar{a} , regardless of the value of r . The opposite holds for H . (ii) Assume that $\bar{a} = 0$ and that F_2 is more dispersed than F_1 . Then, W is increasing in r and H is decreasing in r as long as $r\bar{\lambda} < 1$.*

¹⁹It is possible to get arbitrarily close to the maximum by letting $r \rightarrow 0$. However, there is an equilibrium existence problem for $r = a = 0$ (as long as the tie-breaking rule allows bidder 1 a strictly positive probability of winning).

Proof. See the Appendix. ■

The second part of the Proposition implies that if F_1 is a scaled down version of F_2 and a handicap is used to increase W , then the optimal handicap is so large that the weak bidder wins more often than the strong bidder, as in Theorem 3.

It is natural to ask whether head starts and handicaps can be combined to construct a mechanism that is better for *both* bidders than the standard all-pay auction.²⁰ Of course, such an intervention would increase both H and W . For this to work, one of the bidders must receive a head start and a handicap at the same time. Otherwise, one of the bidders is unambiguously worse off.

Thus, the objective is to choose (\bar{a}, r) to make both bidders better off than with the status quo where $(\bar{a}, r) = (0, 1)$. It would be surprising if the “endowment” is Pareto efficient; trading a higher \bar{a} for a lower r (or vice versa) seems likely to be mutually beneficial. The following example demonstrates the stronger point that it is possible to increase revenue and make both bidders better off at the same time. Thus, the combination of head starts and handicaps can be used to improve the efficiency (or social surplus) of the auction.

Example 3: Assume that $F_i(v) = \frac{v}{\bar{v}_i}$, $v \in [0, \bar{v}_i]$, $i = 1, 2$, with $\bar{v}_2 = 3 > \bar{v}_1 = 1$. Figure 4 illustrates the indifference curves of the two bidders through the status quo, $(\bar{a}, r) = (0, 1)$. Bidder 1 receives a head start to the right of the origin, bidder 2 to the left. Bidder 1 is better off to the north-east, bidder 2 to the south-west.

Clearly, there are two regions where Pareto improvements occur. In the first region, the weak bidder has a head start and a handicap. In the second region, the strong bidder has a large head start and a large handicap. Thus, less sizeable regulation is needed to achieve a Pareto improvement if the weak bidder is given a head start and a handicap.

Moreover, to the right of the origin it can be shown that the level curve on which the total expected expenditures (revenue) are constant is between the two indifference curves. Thus, it is possible to make all parties, including the recipient of the expenditures, better off ex ante. In other words, bidders may be better off even if they are expending more resources. □

H and W are concerned with expectations (of utility) taken over all types. Apart from their mean, the characteristics of the distributions of utility are unimportant. In comparison, Roemer (1998) advocates “equality of opportunity”. Simplifying a bit, the objective of regulation here is to equalize the distribution of “achievement”

²⁰Awarding the prize by a (possibly unfair) lottery may also make both bidders better off. However, this is not a feasible option if a minimum level of revenue must be obtained. Likewise, lotteries may be politically unpalatable. A lottery can be achieved by letting $r_1 = r_2 = a = 0$.

across bidders. Assume for the moment that achievement means expected utility, and let $U_i(\cdot)$ denote bidder i 's expected utility as a function of his type. Then, the distributions of achievement are equalized if $U_2(v) = U_1(\lambda(v))$ for all $v \in [0, \bar{v}_2]$.

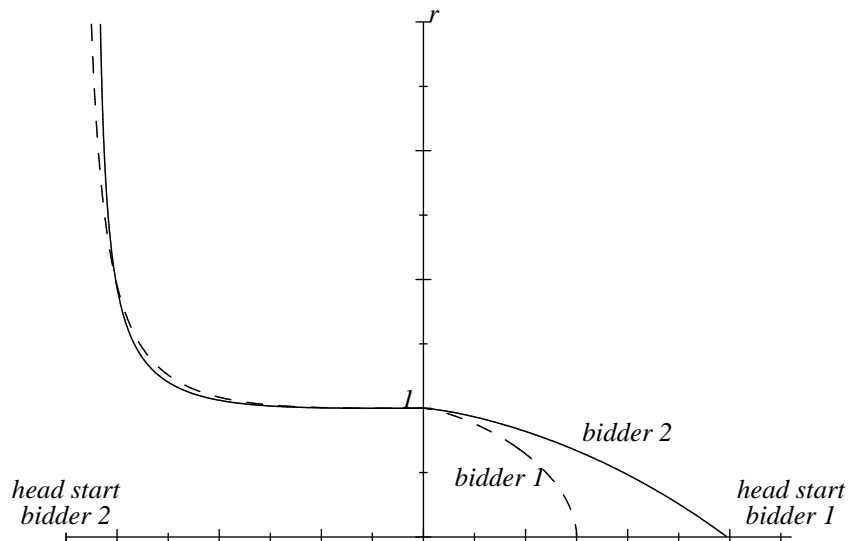


Figure 4: Indifference curves.

A head start will not be able to achieve this objective. The reason is that a head start would exclude the disadvantaged bidder's low types, while at the same time benefitting the low types of his rival. Thus, $U_2(v) \neq U_1(\lambda(v))$ when v is small.

Consider now the use of handicaps. To illustrate, assume for simplicity that $\lambda'(v) < 1$ for all v , which implies that F_1 is less disperse than F_2 . Following Myerson (1981), for any $v \in [0, \bar{v}_2]$,

$$U_2(v) = \int_0^v q_2(x) dx$$

$$U_1(\lambda(v)) = \int_0^{\lambda(v)} q_1(x) dx = \int_0^v q_1(\lambda(x)) \lambda'(x) dx \equiv U_1^s(v),$$

where integration by substitution was used in the second line. To begin, assume for instance that bidders with the same rank win with the same probability, or $q_2(x) = q_1(\lambda(x))$ for all $x \in [0, \bar{v}_2]$. Since $\lambda'(v) < 1$, however, $U_2(v) > U_1(\lambda(v))$ for all $v > 0$. In fact, to have any hope of creating equality of opportunity, the weak bidder must win more often than the strong bidder at comparable ranks. Note that this is in line with the conclusion of Proposition 5 when the objective is to maximize

W .²¹ Similarly, by Theorem 3 a revenue maximizing designer appears to move in the right direction by perhaps creating more equality of opportunity (the caveat is that Theorem 3 does not reveal if the designer is too generous to the weak bidder). However, it is impossible to achieve perfect equality of opportunity with a handicap, or any other instrument for which $q_1(\bar{v}_1) = q_2(\bar{v}_2)$ for that matter. To see this, assume that $U_2(\bar{v}_2) = U_1^s(\bar{v}_2)$ and note that $U_2'(\bar{v}_2) = q_2(\bar{v}_2) > q_1(\bar{v}_1)\lambda'(\bar{v}_2) = U_1^{s'}(\bar{v}_2)$. Hence, there is a neighborhood to the left of \bar{v}_2 where $U_2(v) \neq U_1^s(v)$.²²

Roemer (1998) might not agree with the previous definition of achievement. He is more interested in the final outcome, something that is tangible and in principle measurable. Examples include levels of education or lifetime income, but apparently not utility in the standard sense. Indeed, Roemer (1998, page 26) explicitly states that it is irrelevant whether effort is “pleasant or painful”.²³ By this definition, the probability of winning may be a better measure of achievement. Then, the objective is to equalize $q_2(v)$ and $q_1(\lambda(v))$ for all $v \in [0, \bar{v}_2]$. As noted in Section 5.1 (and Section 5.3 for the complete information model), this can be achieved when F_1 is a scaled down version of F_2 , by setting $\rho = 1$. However, in the general case the objective cannot be met with a linear handicap. Hickman (2011) interprets the difference between $q_2(v)$ and $q_1(\lambda(v))$ as measuring the “enrollment gap” when the contest is a competition for admission into university. Hickman (2011) primarily examines quotas and head starts; his model and results are discussed in the introduction.

10 Discussion and conclusion

Applications and extensions are discussed next. A previous version contains a more thorough and formal examination of the topics in Sections 10.1 and 10.2.

²¹See Roemer (1998, Chapter 4) for a comparison of the equality of opportunity principle with utilitarian and Rawlsian welfare.

²²By similar reasoning, since $U_2'(v) = q_2(v)$ and $U_1^{s'}(v) = q_1(\lambda(v))\lambda'(v)$, $q_1(\lambda(v)) > q_2(v)$ is necessary whenever $q_2(v) > 0$ in order to achieve $U_2(v) = U_1^s(v)$ for all $v \in [0, \bar{v}_2]$. As already mentioned, it is necessary that bidder 1 wins more often than bidder 2, holding the rank fixed.

²³Roughly speaking, Roemer (1998) views the bidder’s propensity for effort (here captured by his type, v) as a primitive of his observable identity (e.g. his upbringing or social circumstances). One of his examples is that of Asian students whose upbringing may emphasize academic effort more. If an Asian student (bidder 2) is competing with a non-Asian student (bidder 1) for admission into university, then neither should be held accountable for the preferences or propensity for effort instilled in them by their different upbringing.

10.1 Application: Contests for research funding

The competition for research funding can be viewed as a contest. A significant amount of time and effort may go into writing a grant proposal, but this investment is not always rewarded.

Think of the researcher as a bidder. The prize is a grant. The bidder is characterized by his observable CV and perhaps his seniority.²⁴ The researcher's "bid" is his proposal, which is costly in terms of effort. His type is how much he values the grant or how much effort he is willing to expend to be successful.

In Canada, the following rules were used until 2011 to evaluate an application for a Standard Research Grant submitted to the Social Sciences and Humanities Research Council (SSHRC): "the score on the record of research achievement accounts for 60 per cent of the overall score, and the score on the program of research accounts for 40 per cent of the overall score".²⁵ Hence, an individual with a good research record will out-score a competing proposal by a less successful scholar, even if the programs of research are of comparable quality. She has a head start.

In contrast, if the applicant obtained his Ph.D. within the last five years of the application deadline he would be considered a "new scholar". In this case, the two components are weighted "such that either a 60/40 or 40/60 ratio will apply, depending on which will produce the more favorable overall score."

Now, compare an established scholar and a new, unproven, scholar. The established scholar has a head start due to her past research achievements. On the other hand, she is handicapped because her score is less sensitive to her effort.

Often, a grant is more valuable early in a scholars' career, meaning that new scholars perhaps value a grant more than established scholars. That is, new scholars

²⁴In practice, there may be an incentive to devote time and effort to build up one's CV this year to increase one's chances next year. Here, however, I consider the CV to be exogenous when the decision to write a proposal is made. See Konrad (2002) for a two-stage model in which preliminary actions in the first stage affects the contest, held in stage two. Arozamena and Cantillon (2004) study another two-stage model. See Avery (1998) and Hörner and Sahuguet (2007) for models in which early action may transmit information to other contestants in a later competition.

²⁵See http://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/standard_grants_subventions_ordinaires-eng.aspx. Since 2011, the Standard Research Grant has been replaced by Insight Grants, but the application for these are evaluated in a similar way. The biggest difference is that Insight Grants do not give any explicit advantages to new scholars. The shorter Insight Development Grants reserves 50% of funds for new scholars. The Australian Research Council evaluates applications based on a fixed 40/60 weighting. A "target level of funding" is identified for projects by Early Career Researchers. See www.arc.gov.au/pdf/DP10_FundingRules.pdf. In the UK, the Economic and Social Research Council has a dedicated program (the First Grants scheme) for new researchers. See www.esrc.ac.uk.

are the “strong” contestants; given the pressure faced by the untenured, they are willing to suffer more (expend more effort) to win a grant. Likewise, established scholars may have more demands on their time and therefore higher opportunity costs. If this is so, then SSHRC gives the weak bidder (the established scholar) a head start and a handicap. Sections 7 and 9 provide two possible explanations for why this design may be worthwhile.

Likewise, within the two categories (established/new scholars) it is the researchers who have already proven their worth, as captured by their better CVs, who have a head start. Relatively speaking, the disadvantaged contestants are those that still have something to prove, i.e. those that may value a grant especially highly.

10.2 Many bidders

For tractability, it was assumed that the contest has only two bidders; Parreiras and Rubinchik (2010) demonstrate the technical difficulties of analyzing even the standard all-pay auction with many bidders. However, with multiple bidders there may be additional reasons for manipulating the contest. For instance, Baye, Kovenock, and de Vries (1993) show that it may pay to exclude strong bidders. Such exclusion can in principle be achieved by giving very large head starts to all rival bidders.

In a companion paper, Kirkegaard (2011b) studies handicaps in all-pay auctions with more than two heterogeneous bidders who have non-linear cost functions. In that setting, some of the weaker bidders may become worse off if a strong bidder is handicapped. In other words, handicapping a contestant may not be favorable to all his rivals when there are more than two bidders in total. In contrast, it follows from Siegel’s (2009) analysis that a bidder in a generic complete information contest can never be hurt (though he may not gain) when an opponent is handicapped.

10.3 Other instruments

A contest can be manipulated in many other ways than by using head starts and handicaps.

For instance, bidding caps are sometimes imposed. Caps in asymmetric contests are analyzed in Che and Gale (1998) and Sahuguet (2006). Che and Gale (1998) consider a complete information model, while Sahuguet (2006) consider the uniform model. In both cases, a suitably chosen cap is profitable. In the incomplete information model, the most obvious effect of a cap is to pool types at the top. That is, both bidders bid at the cap if their type is high enough, say if it exceeds some \hat{v}_1 and \hat{v}_2 , respectively. Of course, the incentive to bid at the cap is different for bidder

1 and bidder 2. It will generally be the case that $\hat{v}_1 \neq k(\hat{v}_2)$, where $k(\cdot)$ is the tying function from the standard all-pay auction. This gives rise to an interesting secondary effect. Specifically, the allocation also changes among the types that bid below the cap. The reason is that for types in $[0, \hat{v}_i]$, the new tying function, \hat{k} , is determined in part by the boundary condition $\hat{k}(\hat{v}_2) = \hat{v}_1 \neq k(\hat{v}_2)$. Hence, \hat{k} and k will differ for any $v \in (0, \hat{v}_2]$. In other words, like a handicap, a cap affects the allocation globally. Thus, a formal analysis of caps would meet similar challenges as those in Section 5.

Kirkegaard (2008) considers a related instrument. He allows one bidder the opportunity to preempt the all-pay auction by submitting a sufficiently high bid (or bribe). In some ways, this identity-dependent instrument is a counterpart to an identity-dependent head start, except that the primary effect is at the top. However, for reasons similar to those described above, the allocation once again changes globally. Kirkegaard (2008) shows that giving the strong bidder the opportunity to preempt the contest is profitable in the uniform model. Indeed, it may be possible to make the seller and both bidders better off ex ante, as in Example 3.

Another common instrument is a reserve price or bidding floor, which excludes both bidders' low types. In the model in the current paper, such exclusion has the attractive feature that it denies types with negative virtual valuation the prize. As in Sections 6 and 7, once a reserve price has been used to change things at the bottom, there is more of an incentive to use a handicap to fix things at the top by handicapping the weak bidder. However, in the absence of a handicap, the reserve price generally causes a mass of the strong bidder's relatively low types to bid exactly at the reserve price. It would be desirable to exclude these types altogether, because they are likely to have negative virtual valuations. This can be achieved by increasing the reserve price but at the same time give the weak bidder a head start that nullifies the increase in the reserve. The weak bidder then uses the same strategy as before, but it now becomes less desirable for the strong bidder to bid at the reserve. Thus, a reserve price and a head start may go hand-in-hand.

10.4 Concluding remarks

By contrasting head starts and handicaps, I showed that the specifics of the instruments used to favor a contestant are highly important in determining the consequences of favoritism. Even if they favor different contestants, different instruments need not cancel each other out. In fact, they may be profitably combined in order to pursue objectives such as expenditure maximization or the maximization of bidders' payoff. In particular, giving the weak bidder a head start and a handicap may increase revenue and/or make both bidders better off.

References

- [1] Amann, E. and Leininger, W., 1996, Asymmetric All-Pay Auctions with Incomplete Information: The Two-Player Case, *Games and Economic Behavior*, 14: 1-18.
- [2] Arozamena, L. and Cantillon, E., 2004, Investment Incentives in Procurement Auctions, *Review of Economic Studies*, 71: 1-18.
- [3] Asker, J. and Cantillon, E., 2008, Properties of scoring auction, *RAND Journal of Economics*, 39 (1): 69-85.
- [4] Avery, C., 1998, Strategic Jump Bidding in English Auctions, *Review of Economic Studies*, 65: 185-210.
- [5] Baye, M.R., Kovonock, D., and de Vries, C.G., Rigging the Lobbying Process: An Application of the All-Pay Auction, *American Economic Review*, 83 (1): 289-294.
- [6] Bulow, J., and Klemperer, P., 1996, Auctions Versus Negotiation, *American Economic Review*, 86 (1): 180-194.
- [7] Bulow, J. and Roberts, J., 1989, The Simple Economics of Optimal Auctions, *Journal of Political Economy*, 97: 1060-1090.
- [8] Calsamiglia, C., 2009, Decentralizing Equality of Opportunity, *International Economic Review*, 50 (1): 273-290.
- [9] Che, Y-K, 1993, Design competition through multidimensional auctions, *RAND Journal of Economics*, 24 (4): 668-680.
- [10] Che, Y-K and Gale, I.L., 1998, Caps on Political Lobbying, *American Economic Review*, 88 (3): 643-651.
- [11] Clark, D.J. and Riis, C., 2000, Allocation efficiency in a competitive bribery game, *Journal of Economic Behavior & Organization*, 42: 109-124.
- [12] d'Aspremont, C. and Gérard-Varet, L.-A., 1979, Incentives and Incomplete Information, *Journal of Public Economics*, 11: 25-45.
- [13] Feess, E., Muehlheusser, G., and Walzl, M., 2008, Unfair contests, *Journal of Economics*, 93 (3): 267-291.

- [14] Fu, Q., 2006, A Theory of Affirmative Action in College Admissions, *Economic Inquiry*, 44 (3): 420-428.
- [15] Gaviious, A., Moldovanu, B., and Sela, A., 2002, Bid costs and endogenous bid caps, *RAND Journal of Economics*, 33 (4): 709-722.
- [16] Hickman, B.R., Human Capital Investment, Race Gaps, And Affirmative Action: A Game-Theoretic Analysis of College Admissions, mimeo.
- [17] Hörner, J. and Sahuguet, N., 2007, Costly Signalling in Auctions, *Review of Economic Studies*, 74: 173-206.
- [18] Hopkins, E., 2007, Rank-Based Methods for the Analysis of Auctions, mimeo.
- [19] Kirkegaard, R., 2008, Comparative Statics and Welfare in Heterogeneous All-Pay Auctions: Bribes, Caps, and Performance Thresholds, *The B.E. Journal of Theoretical Economics*, Volume 8 (Topics), Article 21.
- [20] Kirkegaard, R., 2011a, A Mechanism Design Approach to Ranking Asymmetric Auctions, *Econometrica*, forthcoming.
- [21] Kirkegaard, R., 2011b, Preferential Treatment may Hurt: Another Application of the All-Pay Auction, mimeo.
- [22] Konrad, K.A., 2002, Investment in the absence of property rights; the role of incumbency advantages, *European Economic Review*, 46: 1521-1537.
- [23] Lien, D., 1990, Corruption and Allocation Efficiency, *Journal of Development Economics*, 33: 153-164.
- [24] Lu, J., 2010, Entry Coordination and Auction Design with Private Costs of Information Acquisition, *Economic Inquiry*, 48 (2): 274-289.
- [25] Maskin, E. and Riley, J., 2000, Asymmetric Auctions, *Review of Economic Studies*, 67: 413-438.
- [26] Myerson, R.B., 1981, Optimal Auction Design, *Mathematics of Operations Research*, 6: 58-73.
- [27] Myerson, R.B. and Satterthwaite, M.A., 1983, Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory*, 29: 265-281.

- [28] Nti, K.O., Maximum efforts in contests with asymmetric valuations, *European Journal of Political Economy*, 20: 1059-1066.
- [29] Parreiras, S.O. and Rubinchik, A., 2010, Contests with Three or More Heterogeneous Agents, *Games and Economic Behavior*, 68: 703-715.
- [30] Roemer, J.E., 1998, *Equality of Opportunity*, Harvard University Press.
- [31] Sahuguet, N., 2006, Caps in asymmetric all-pay auctions with incomplete information, *Economics Bulletin*, 3 (9): 1-8.
- [32] Shaked, M. and Shanthikumar, J.G., 2007, *Stochastic Orders*, Springer
- [33] Siegel, R., 2009, All-Pay Contests, *Econometrica*, 77 (1): 71-92.

Appendix: Proofs

Proof of Proposition 1. It will be confirmed that no bidder has an incentive to deviate from the strategies in the Proposition, given that the other bidder follow his equilibrium strategy.

As a preliminary step, note from (2) that

$$k'(v) = \frac{r_1 k(v) f_2(v)}{r_2 v f_1(k(v))} \quad (16)$$

for all $v \in (0, \bar{v}_2]$. Thus, when $v \geq v_1^c$,

$$\begin{aligned} s_1(v) &= a_1 + \int_{v_1^c}^v r_2 k^{-1}(x) f_1(x) dx = a_1 + \int_{k^{-1}(v_1^c)}^{k^{-1}(v)} r_2 y f_1(k(y)) k'(y) dy \\ &= a_1 + \int_{v_2^c}^{k^{-1}(v)} r_1 k(y) f_2(y) dy = s_2(k^{-1}(v)), \end{aligned}$$

where the second equality follows from integration by substitution. In conclusion, $s_1(v) = s_2(k^{-1}(v))$, which confirms that bidder 1 with type $k(v)$ scores the same as bidder 2 with type v . Since strategies are monotonic, a score of $s_1(z)$, $z \in (v_1^c, \bar{v}_1]$, gives bidder 1 with type v expected payoff of

$$\begin{aligned} EU_1(v, z) &= v F_2(k^{-1}(z)) - c_1(s_1(z)) = v F_2(k^{-1}(z)) - \int_{v_2^c}^{k^{-1}(z)} k(x) f_2(x) dx \\ &= v F_2(v_2^c) + \int_{v_2^c}^{k^{-1}(z)} (v - k(x)) f_2(x) dx. \end{aligned}$$

In comparison, a score of only a_1 (bidding zero, as if the type is zero) yields expected payoff of $EU_1(v, 0) = v F_2(v_2^c)$, and it follows that

$$EU_1(v, z) - EU_1(v, 0) = \int_{v_2^c}^{k^{-1}(z)} (v - k(x)) f_2(x) dx.$$

If $v \geq v_1^c$, expected payoff is therefore maximized by letting $z = v$ or scoring $s_1(v)$ as prescribed by the equilibrium strategy. Likewise, if $v < v_1^c$, bidding zero or scoring a_1 is optimal. Note that scoring higher than $s_1(\bar{v}_1) = s_2(\bar{v}_2)$ cannot be optimal either, because such an action would only increase costs without increasing the winning probability beyond that obtained by a score of $s_1(\bar{v}_1)$.

Turning to bidder 2 with type v , a score of $s_2(z)$, $z \in [v_2^c, \bar{v}_2]$, produces expected payoff of

$$\begin{aligned} EU_2(v, z) &= vF_1(k(z)) - c_2(s_1(k(z))) = vF_1(k(z)) - \int_{v_1^c}^{k(z)} k^{-1}(x)f_1(x)dx - c_2(a_1) \\ &= vF_1(v_1^c) + \int_{v_1^c}^{k(z)} (v - k^{-1}(x)) f_1(x)dx - c_2(a_1) \\ &= \int_{v_1^c}^{k(z)} (v - k^{-1}(x)) f_1(x)dx, \end{aligned}$$

where the last equality follows from the definition of v_1^c in (3). In comparison, if bidder 2 scores zero, his payoff is zero. Thus, bidder 2 has no incentive to deviate from his equilibrium strategy either. ■

Proof of Lemma 1. Compare the slopes of k and λ ,

$$k'(v) = \left(r \frac{k(v)}{v} \right) \times \frac{f_2(v)}{f_1(k(v))}, \quad \lambda'(v) = \frac{f_2(v)}{f_1(\lambda(v))},$$

where the former can be derived from (2) and the latter from the definition of λ . If k coincides with λ , the term in parenthesis in $k'(v)$ is at most $r\bar{\lambda}$, while the other term equals $\lambda'(v)$. Thus, if $r\bar{\lambda} < 1$ then $k(v)$ is flatter than $\lambda(v)$ whenever $k(v) = \lambda(v)$. Consequently, $k(v) = \lambda(v)$ at most once on $(0, \bar{v}_2]$. However, since $k(\bar{v}_2) = \lambda(\bar{v}_2)$ they can never coincide on $(0, \bar{v}_2)$. Since $k'(\bar{v}_2) < \lambda'(\bar{v}_2)$ it must then hold that $k(v) > \lambda(v)$ for all $v \in (0, \bar{v}_2)$. The proof of the second part is analogous. ■

Proof of Lemma 2. Assume $\lambda(\tau_2(v)) = v$ or $F_2(\tau_2(v)) = F_1(v)$ for some $v > 0$. Since

$$\int_0^{\tau_2} J_2(x)f_2(x)dx = -\tau_2(1 - F_2(\tau_2)),$$

(8) implies that

$$\begin{aligned} -\tau_2(1 - F_2(\tau_2)) &= J_1(v)F_1(v) = \int_0^v J_1(x)f_1(x)dx \\ &> \int_0^v J_1(x)f_1(x)dx = -v(1 - F_1(v)) = -v(1 - F_2(\tau_2)) \end{aligned}$$

or $v > \tau_2$. However, this leads to the contradiction that $F_1(v) > F_1(\tau_2) > F_2(\tau_2)$. Thus, there is no $v > 0$ for which bidder 2 with type τ_2 has the same rank as bidder 1

with type v . Moreover, since $\tau_2(0) > 0$ or $F_1(0) = 0 < F_2(\tau_2(0))$, bidder 2 with type $\tau_2(0)$ has higher rank than bidder 1 with type 0. Combining the two observations proves the Lemma. ■

Proof of Lemma 3. The first step is to show that $EP_i(\rho)$ can be written in the form $EP_i(\rho) = \bar{v}_i EP_i^s(\rho)$, as claimed. If bidder i has type v_i , his “scale-adjusted” type is $v_i^s = \frac{v_i}{\bar{v}_i}$. Note that bidder 1 and bidder 2 have the same rank if their scale-adjusted types are the same, i.e. $F_1(v_1) = F_2(v_2)$ if $v_1^s = v_2^s$. Using Definition 2, equation (2) can be written

$$\int_k^{\bar{v}_1} \frac{\frac{1}{\bar{v}_1} f\left(\frac{x}{\bar{v}_1}\right) \frac{1}{x}}{\frac{x}{\bar{v}_1}} dx = r \int_v^{\bar{v}_2} \frac{\frac{1}{\bar{v}_2} f\left(\frac{x}{\bar{v}_2}\right) \frac{1}{x}}{\frac{x}{\bar{v}_2}} dx.$$

Integrating both sides by substitution yields

$$\int_{k^s}^1 \frac{f(x)}{x} dx = \rho \int_{v^s}^1 \frac{f(x)}{x} dx, \quad (17)$$

where k^s is the scale-adjusted type of bidder 1 who ties with the scale-adjusted type of bidder 2, v^s . Note that $k^s(v^s) = v^s$, or $k(v) = \lambda(v)$, if $\rho = 1$.

If bidder i has type v , his virtual valuation is

$$J_i(v) = v - \frac{1 - F\left(\frac{v}{\bar{v}_i}\right)}{\frac{1}{\bar{v}_i} f\left(\frac{v}{\bar{v}_i}\right)} = \bar{v}_i \left(\frac{v}{\bar{v}_i} - \frac{1 - F\left(\frac{v}{\bar{v}_i}\right)}{f\left(\frac{v}{\bar{v}_i}\right)} \right) = \bar{v}_i J(v^s),$$

where

$$J(v^s) = v^s - \frac{1 - F(v^s)}{f(v^s)}.$$

Next, define $q_i^s(v^s) = q_i(v^s \bar{v}_i) = q_i(v)$ as bidder i 's scale-adjusted winning probability. Bidder i 's ex ante expected payment can be written as

$$\begin{aligned} EP_i &= \int_0^{\bar{v}_i} J_i(v) q_i^s\left(\frac{v}{\bar{v}_i}\right) f_i(v) dv = \int_0^{\bar{v}_i} \bar{v}_i J\left(\frac{v}{\bar{v}_i}\right) q_i^s\left(\frac{v}{\bar{v}_i}\right) \frac{1}{\bar{v}_i} f\left(\frac{v}{\bar{v}_i}\right) dv \\ &= \bar{v}_i \int_0^1 J(v^s) q_i^s(v^s) f(v^s) dv^s, \end{aligned}$$

where the last step follows from integrating by substitution. Defining

$$EP_i^s = \int_0^1 J(v^s) q_i^s(v^s) f(v^s) dv^s \quad (18)$$

as bidder i 's scale-adjusted expected payment, bidder i 's expected payment is $EP_i = \bar{v}_i EP_i^s$. Since k^s determines the winning probability, $q_i^s(v^s)$, and by extension EP_i^s , it follows from (17) that EP_i^s depends only on ρ , and not at all on \bar{v}_i .

The results stated in the lemma are proven next. The expected scale-adjusted payment from bidder 2 is

$$\begin{aligned} EP_2^s &= \int_0^1 J(v^s)F(k^s(v^s))f(v^s)dv^s = \int_0^1 v^s(1 - F(v^s))f(k^s(v^s))k^{s'}(v^s)dv^s \\ &= \int_0^1 (1 - F(v^s))\rho k^s(v^s)f(v^s)dv^s, \end{aligned}$$

where the second line comes from integration by parts and the third from implicit differentiation of (17).

Equation (17) reveals how k^s depends on ρ

$$\frac{\partial k^s(v^s)}{\partial \rho} = -\frac{k^s(v^s)}{f(k^s(v^s))} \int_{v^s}^1 \frac{f(x)}{x} dx,$$

and it follows that

$$\frac{\partial EP_2^s(\rho)}{\partial \rho} = \int_0^1 k^s(v^s)(1-F(v^s))f(v^s)dv^s - \int_0^1 (1-F(v^s)) \left(\rho \frac{k^s(v^s)f(v^s)}{f(k^s(v^s))} \int_{v^s}^1 \frac{f(x)}{x} dx \right) dv^s. \quad (19)$$

From (17),

$$\rho \frac{k^s(v^s)f(v^s)}{f(k^s(v^s))} = k^{s'}(v^s)v^s$$

meaning that (19) can be written

$$\frac{\partial EP_2^s(\rho)}{\partial \rho} = \int_0^1 k^s(v^s)(1 - F(v^s))f(v^s)dv^s - \int_0^1 v^s(1 - F(v^s))g(v^s)dv^s \quad (20)$$

where

$$g(v^s) = k^{s'}(v^s) \int_{v^s}^1 \frac{f(x)}{x} dx.$$

The antiderivative of g is

$$\begin{aligned} G(v^s) &= \int_0^{v^s} g(y)dy = \left[k^s(y) \int_y^1 \frac{f(x)}{x} dx \right]_0^{v^s} + \int_0^{v^s} \frac{k^s(y)}{y} f(y)dy \\ &= k^s(v^s) \int_{v^s}^1 \frac{f(x)}{x} dx + \int_0^{v^s} \frac{k^s(y)}{y} f(y)dy, \end{aligned}$$

since L'Hopital's rule can be used to show that $k^s(y) \int_y^1 f(x)/x dx$ converges to zero as y converges to zero. The last term in (19) can now be rewritten by integrating by parts

$$\begin{aligned}
& \int_0^1 (1 - F(v^s)) \left(\rho \frac{k^s(v^s) f(v^s)}{f(k^s(v^s))} \int_{v^s}^1 \frac{f(x)}{x} dx \right) dv^s \\
&= \int_0^1 v^s (1 - F(v^s)) g(v^s) dv^s \\
&= \int_0^1 G(v^s) (v^s f(v^s) - (1 - F(v^s))) dv^s \\
&= \int_0^1 k^s(v^s) v^s \int_{v^s}^1 \frac{f(x)}{x} dx f(v^s) dv^s - \int_0^1 k^s(v^s) (1 - F(v^s)) \int_{v^s}^1 \frac{f(x)}{x} dx dv^s \\
&\quad + \int_0^1 \left(\int_0^{v^s} \frac{k(x)}{x} f(x) dx \right) (v^s f(v^s) - (1 - F(v^s))) dv^s.
\end{aligned}$$

Rearranging yields

$$\begin{aligned}
& 2 \int_0^1 k^s(v^s) (1 - F(v^s)) \int_{v^s}^1 \frac{f(x)}{x} dx \left(\frac{\rho f(v^s)}{f(k^s(v^s))} \right) dv^s + A(\rho) \\
&= \int_0^1 k^s(v^s) v^s \int_{v^s}^1 \frac{f(x)}{x} dx f(v^s) dv^s + \int_0^1 k^s(v^s) (1 - F(v^s)) f(v^s) dv^s,
\end{aligned}$$

where integration by parts was used to obtain the last part on the right hand side and where

$$A(\rho) = \int_0^1 k^s(v^s) (1 - F(v^s)) \int_{v^s}^1 \frac{f(x)}{x} dx \left(1 - \frac{\rho f(v^s)}{f(k^s(v^s))} \right) dv^s.$$

Thus,

$$\begin{aligned}
& \int_0^1 k^s(v^s) (1 - F(v^s)) \int_{v^s}^1 \frac{f(x)}{x} dx \left(\frac{\rho f(v^s)}{f(k^s(v^s))} \right) dv^s \\
&= \frac{1}{2} \left(\int_0^1 k^s(v^s) v^s \int_{v^s}^1 \frac{f(x)}{x} dx f(v^s) dv^s + \int_0^1 k^s(v^s) (1 - F(v^s)) f(v^s) dv^s - A(\rho) \right).
\end{aligned}$$

Inserting this back into (19) yields

$$\begin{aligned}
\frac{\partial EP_2^s(\rho)}{\partial \rho} &= \frac{1}{2} \left(\int_0^1 k^s(v^s)(1 - F(v^s))f(v^s)dv^s - \int_0^1 k^s(v^s)v^s \int_{v^s}^1 \frac{f(x)}{x} dx f(v^s)dv^s + A(\rho) \right) \\
&= \frac{1}{2} \left(\int_0^1 k^s(v^s)v^s \left(\frac{1 - F(v^s)}{v^s} - \int_{v^s}^1 \frac{f(x)}{x} dx \right) f(v^s)dv^s + A(\rho) \right) \\
&= \frac{1}{2} \left(\int_0^1 k^s(v^s)v^s \left(\int_{v^s}^1 \frac{f(x)}{v^s} dx - \int_{v^s}^1 \frac{f(x)}{x} dx \right) f(v^s)dv^s + A(\rho) \right)
\end{aligned}$$

Note that the first term is strictly positive. If $\rho = 1$ then $k^s(v^s) = v^s$. Thus, $A(1) = 0$ and $EP_2^{s'}(1) > 0$ regardless of the curvature F . Consider now $\rho \in (0, 1)$. To prove that $A(\rho)$ is positive in this case, it suffices to show that the integrand is positive for all values of v^s . In other words, the intention is to show that

$$1 - \frac{\rho f(v^s)}{f(k^s(v^s))} = f(v^s) \left(\frac{1}{f(v^s)} - \frac{\rho}{f(k^s(v^s))} \right)$$

is positive. Using (17) to solve for ρ , the previous expression is proportional to

$$\frac{1}{f(v^s)} \int_{v^s}^1 \frac{f(x)}{x} dx - \frac{1}{f(k^s(v^s))} \int_{k^s(v^s)}^1 \frac{f(x)}{x} dx.$$

Since $\rho < 1$, $k^s(v^s) \geq v^s$ and so it is sufficient to prove that

$$C(v^s) \equiv \frac{1}{f(v^s)} \int_{v^s}^1 \frac{f(x)}{x} dx = \frac{1 - F(v^s)}{f(v^s)} \int_{v^s}^1 \frac{1}{x} \frac{f(x)}{1 - F(v^s)} dx$$

is decreasing in v^s . The first term is non-negative and decreasing, by assumption. Since the second term is also non-negative and decreasing, $C'(v^s) \leq 0$. Thus, $A(\rho) \geq 0$ for all $\rho \in (0, 1]$ whenever $F(v)$ satisfies the regularity assumption. Hence, $EP_2^{s'}(\rho) > 0$ for all $\rho \in (0, 1]$.

The fact that $EP_1^{s'}(1) = -EP_2^{s'}(1) < 0$ regardless of the curvature of F and that $EP_1^s(\rho) < 0$ for all $\rho \in [1, \infty)$ when F is regular can be proven in a similar manner. However, this symmetry is obvious; (17) and (18) imply that the roles of bidder 1 and bidder 2 can be reversed by changing ρ to $\frac{1}{\rho}$. Thus, $EP_1^s(\rho) = EP_2^s(\frac{1}{\rho})$. ■

Proof of Theorem 3. Since $\bar{v}_2 > \bar{v}_1$, the first part of Lemma 3 implies that $ER'(1) > 0$, which in turn implies the existence of a local maximum at some $\rho > 1$.²⁶

²⁶Note that $k^s(v^s) \rightarrow 0$ as $\rho \rightarrow \infty$, or $q_1^s(v^s) \rightarrow 1$, $q_2^s(v^s) \rightarrow 0$. This implies that $EP_i^s(\rho) \rightarrow 0$ and thus $ER(\rho) \rightarrow 0$ as $\rho \rightarrow \infty$. In other words, the optimal value of ρ is finite.

Next, I will show $ER(\rho)$ cannot achieve a global maximum at any $\rho < 1$. Since $J(v)$ is monotonic, by assumption, $EP_1^s(\rho) + EP_2^s(\rho)$ is maximized at $\rho = 1$. The reason is that $\rho = 1$ implies a symmetric allocation, which is optimal in a mechanism with homogeneous bidders (given the good is sold with probability one). It follows that if $EP_i^s(\rho) > EP_i^s(1)$ then $EP_j^s(\rho) < EP_j^s(1)$, $j \neq i$; at most one of the scale-adjusted payments can exceed the scale-adjusted payments for $\rho = 1$. Likewise, $ER'(1) \neq 0$ means that $\rho = 1$ is not optimal, and thus $EP_i^s(\rho) > EP_i^s(1)$ for some $i = 1, 2$ at the optimal ρ . Finally, to maximize (11), it must be the case that $EP_2^s(\rho) \geq EP_1^s(\rho)$. Otherwise, changing ρ to $\frac{1}{\rho}$ reverts the roles of the bidders in (17) and therefore switches EP_2^s and EP_1^s . This would increase expected revenue because the coefficient of EP_2^s is the larger. Putting these observations together imply that $EP_2^s(\rho) > EP_2^s(1) = EP_1^s(1) > EP_1^s(\rho)$ when $ER(\rho)$ is maximized. The first inequality necessitates $\rho > 1$, by Lemma 3. ■

Proof of Theorem 4. Using the same procedure as in Lemma 3, the derivative of bidder 2's expected payment with respect to r can be written

$$\begin{aligned} \frac{\partial EP_2(r)}{\partial r} &= \frac{1}{2} \int_0^{\bar{v}_2} k(v)v \left(\int_v^{\bar{v}_2} \frac{f_2(x)}{v} dx - \int_v^{\bar{v}_2} \frac{f_2(x)}{x} dx \right) f_2(v) dv \\ &\quad + \frac{1}{2} \int_0^{\bar{v}_2} k(v)(1 - F_2(v)) \int_v^{\bar{v}_2} \frac{f_2(x)}{x} dx \left(1 - \frac{rf_2(v)}{f_1(k(v))} \right) dv. \end{aligned}$$

The first term is positive. By assumption, $f_1(k(v)) \geq \min f_1(x) \geq \max f_2(x) \geq f_2(v)$. Hence, when $r \leq 1$, the second term is positive as well. Thus, bidder 2's expected payment declines if he is favored ($r < 1$).

Consider now the effects of handicapping bidder 1. Switching the roles of bidders 1 and 2 yields

$$\begin{aligned} \frac{\partial EP_1(r)}{\partial r} &= \frac{1}{2} \int_0^{\bar{v}_1} k^{-1}(v)v \left(\int_v^{\bar{v}_1} \frac{f_1(x)}{v} dx - \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx \right) f_1(v) dv \\ &\quad + \frac{1}{2} \int_0^{\bar{v}_1} k^{-1}(v)(1 - F_1(v)) \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx \left(1 - \frac{rf_1(v)}{f_2(k^{-1}(v))} \right) dv. \end{aligned}$$

with some abuse of notation, since r should here be interpreted as a handicap against bidder 1. With this interpretation in mind, the interest is on values of r greater than 1, which is equivalent to favoring bidder 2. When $r \geq 1$, the second term in the derivative is negative. Collecting terms,

$$\frac{\partial EP_1(r)}{\partial r} = \frac{1}{2} \int_0^{\bar{v}_1} k^{-1}(v)T(v)f_1(v)dv$$

where

$$\begin{aligned} T(v) &= 1 - F_1(v) - v \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx + (1 - F_1(v)) \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx \left(\frac{1}{f_1(v)} - \frac{r}{f_2(k^{-1}(v))} \right) \\ &\leq 1 - F_1(v) - v \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx + (1 - F_1(v)) \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx \left(\frac{1}{\min f_1(x)} - \frac{1}{\max f_2(x)} \right), \end{aligned}$$

whenever $r \geq 1$. Note that this bound on T is zero when $v = \bar{v}_1$. For brevity, let $C < 0$ denote the term in the parentheses. The derivative of the bound with respect to v is then

$$- \int_v^{\bar{v}_1} \frac{f_1(x)}{x} dx (1 + f_1(v)C) - (1 - F_1(v)) \frac{f_1(v)}{v} C,$$

which is positive if $1 + f_1(v)C \leq 0$. This is the case for all v if $1 + \min f_1(x)C \leq 0$ or

$$2 \max f_2(x) - \min f_1(x) \leq 0,$$

which is true by the assumption in the theorem. Hence, $T(v)$ is bounded above by a function that is negative, and so $T(v) \leq 0$ for all v . It now follows that bidder 1's expected payment decreases whenever he is handicapped. In conclusion, both bidders pay less, in expectation, if bidder 1 is handicapped. The theorem follows. ■

Proof of Theorem 5. Assume (i) is satisfied. Theorem 1 implies that $\bar{a} \neq 0$. It remains to show that $\bar{a} \not\leq 0$. By Theorem 2, $\bar{a} < 0$ necessitates that $r\bar{\lambda} \geq 1$. Following the argument in the proof of Theorem 1, any $\bar{a} < 0$ that is a candidate for a maximum must produce a (v_2^c, v_1^c) pair at the intersection of τ_2^{-1} and k in Figure 2 (where k implicitly depends on r). I next show that a more profitable combination of head starts and handicaps exists. Since τ_2^{-1} is below κ , any intersection of τ_2^{-1} and k takes place in the region below κ . Since $k(0) = 0$ but $\kappa(v) = 0$ for some $v > 0$, it also follows that the unique (by assumption) intersection of k and κ must occur to the left of any intersection between τ_2^{-1} and k . In other words, k and κ do not intersect to the right of the intersection between τ_2^{-1} and k , which means that k is below κ from this point on. In this region, below κ , expected revenue increases if bidder 2 wins more often. This can be achieved by lowering r (shifting k upwards) while at the same time adjusting a to keep v_1^c constant. The shaded area in Figure A.1 captures the combinations of types for which the object is awarded to bidder 2 with the new mechanism but not the old mechanism. Since bidder 2's virtual valuation exceeds that of bidder 1 in this region, expected revenue has increased.

Assume now that (ii) is satisfied instead. The first part of the condition implies that $\lambda'(v)v \geq \lambda(v)$ or $vf_2(v) \geq \lambda(v)f_1(\lambda(v))$ since $\lambda'(v) = \frac{f_2(v)}{f_1(\lambda(v))}$. Recalling that

$v \geq \lambda(v)$ and $1 - F_1(\lambda(v)) = 1 - F_2(v)$ by definition of $\lambda(v)$, a comparison of J_1 and J_2 at identical ranks then reveals that

$$J_2(v) - J_1(\lambda(v)) = v \left[1 - \frac{1 - F_2(v)}{vf_2(v)} \right] - \lambda(v) \left[1 - \frac{1 - F_1(\lambda(v))}{\lambda(v)f_1(\lambda(v))} \right] \geq 0$$

whenever $J_2(v) \geq 0$ or $v \geq v_2^*$. Since J_1 is an increasing function, by assumption, it must be the case that $\kappa(v) \geq \lambda(v)$ for all $v \geq v_2^*$ (recall that κ satisfies $J_1(\kappa(v)) = J_2(v)$). The significance of the second part of condition (ii) is the implication that $\tau_2(0) \geq v_2^*$. Since $\lambda > \tau_2^{-1}$ whenever τ_2^{-1} is defined (Lemma 2), the conclusion is that $\kappa(v) \geq \lambda(v) > \tau_2^{-1}(v)$ whenever τ_2^{-1} is defined.

Now compare k and λ . As noted above, $r\bar{\lambda} \geq 1$ is necessary for a head start to bidder 2 to be profitable. As before, any candidate (v_2^c, v_1^c) pair must be found at the intersection of τ_2^{-1} and k , but at such a point k must be below λ (since τ_2^{-1} is). Starting from such a point, it can then be shown that $k(v) \leq \lambda(v)$ for any $v \in (v_2^c, \bar{v}_2]$. This property follows from the assumed monotonicity of $\frac{\lambda(v)}{v}$ and an argument similar to the one in the proof of Lemma 1. Consequently, $\kappa(v) \geq k(v) > \tau_2^{-1}(v)$ for all $v \geq v_2^c$. The argument that lead to Figure A.1 can now be repeated to conclude the proof. ■

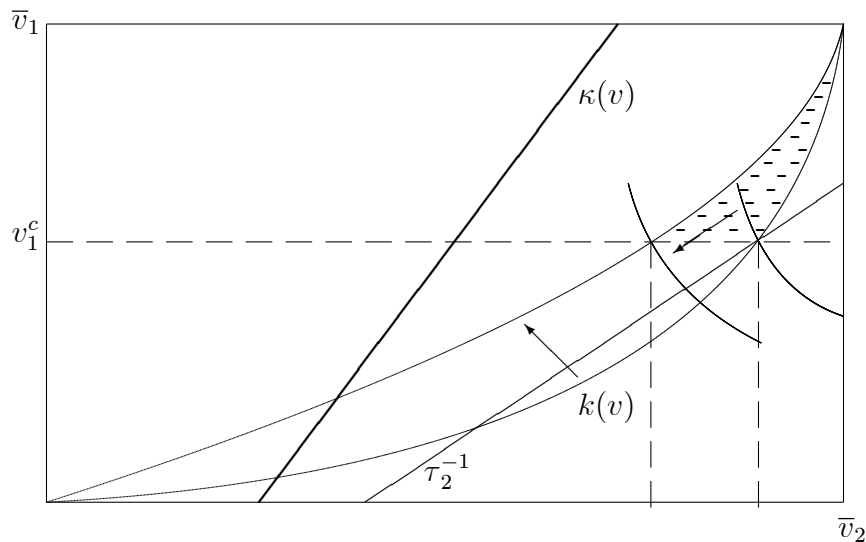


Figure A.1: The weak bidder gets the head start, $a^{**} > 0$.

Proof of Lemma 4. First, F_2 automatically first order stochastically dominates F_1 if \bar{v}_2 is sufficiently large. From Theorem 1, it follows that it is sufficient to show that τ_1 and k cross exactly once in order to prove that $ER(\bar{a}, 1)$ is single-peaked in \bar{a} , $\bar{a} \geq 0$. Using the definition of τ_1 ,

$$\frac{d}{dv} \ln F_1(\tau_1(v)) = \frac{f_1(\tau_1(v))}{F_1(\tau_1(v))} \tau_1'(v) = J_2'(v) (J_1(\tau_1(v)) - J_2(v))^{-1},$$

whenever $\tau_1(v) > 0$. The term in the parenthesis is strictly positive, since J_1 is strictly increasing. Once again by definition of τ_1 ,

$$J_2(v) = -\tau_1(v) \frac{1 - F_1(\tau_1(v))}{F_1(\tau_1(v))},$$

and so

$$\frac{d}{dv} \ln F_1(\tau_1(v)) = J_2'(v) \left(\frac{\tau_1(v)}{F_1(\tau_1(v))} - \frac{1 - F_1(\tau_1(v))}{f_1(\tau_1(v))} \right)^{-1}.$$

Since h_2 is monotonic, by assumption, $J_2'(v) \geq 1$. Thus,

$$\frac{d}{dv} \ln F_1(\tau_1(v)) \geq \left(\max_x \left(\frac{x}{F_1(x)} - \frac{1 - F_1(x)}{f_1(x)} \right) \right)^{-1}. \quad (21)$$

By assumption, f_1 is finite and strictly positive. Hence, both ratios are finite and strictly positive. The difference must be finite, and so the term on the right of (21) is strictly positive. In other words, the slope of $\ln F_1(\tau_1(v))$ is bounded below, and the bound just derived is independent of F_2 . On the other hand, for $k > 0$ and $r = 1$,

$$\frac{d}{dv} \ln F_1(k(v)) = \frac{f_1(k(v))}{F_1(k(v))} k'(v) = \frac{k(v)}{F_1(k(v))} \frac{f_2(v)}{v} = \frac{k(v)}{F_1(k(v))} \frac{1}{(\bar{v}_2)^2} \frac{f\left(\frac{v}{\bar{v}_2}\right)}{\frac{v}{\bar{v}_2}}. \quad (22)$$

As explained in Section 4, k and τ_1 intersects at least once. At such an intersection, both k and v are strictly positive, while $J_2(v)$ is strictly negative and equal to $E[J_1(x)|x \leq \tau_1(v)]$. Since (see Section 5)

$$J_2(v) = \bar{v}_2 J\left(\frac{v}{\bar{v}_2}\right) = E[J_1(x)|x \leq \tau_1(v)]$$

where J is negative and strictly increasing, $\frac{v}{\bar{v}_2}$ must increase to produce the same τ_1 when \bar{v}_2 increases. In other words, when \bar{v}_2 is large, τ_1 is not defined for small values of $\frac{v}{\bar{v}_2}$. Thus, at any intersection of τ_1 and k , the third term in (22) is bounded above.

The first term is also bounded above (since f_1 is bounded below, $k/F_1(k)$ does not explode as k goes to zero). However, the second term goes to zero as \bar{v}_2 increases. Thus, when \bar{v}_2 is sufficiently large, $\ln F_1(k(v))$ is almost flat at any point where it intersects $\ln F_1(\tau_1(v))$. Since the slope of the latter is bounded below, it follows that the two functions intersect no more than once. This concludes the proof. ■

Proof of Corollary 3. As a preliminary step, note that $J_i(\bar{v}_i) = \bar{v}_i$ and

$$J'_i(v) = 2 + \frac{f'_i(v)}{f_i(v)} \frac{1 - F_i(v)}{f_i(v)}.$$

By concavity and convexity, respectively, it follows that

$$J_1(v) \geq 2v - \bar{v}_1 \text{ and } J_2(v) \leq 2v - \bar{v}_2.$$

By definition, $\tau_1(v) = 0$ when v is such that $J_1(0) = J_2(v)$ or

$$-\bar{v}_1 \leq J_1(0) = J_2(v) \leq 2v - \bar{v}_2,$$

from which it can be inferred that $v \geq (\bar{v}_2 - \bar{v}_1)/2$ is necessary for τ_1 to be defined.

Since F_2 is convex, $J'_2(v) \geq 2$. Then, from the proof of Lemma 4,

$$\frac{d}{dv} \ln F_1(\tau_1(v)) = J'_2(v) \underbrace{\left(\frac{\tau_1(v)}{F_1(\tau_1(v))} - \frac{1 - F_1(\tau_1(v))}{f_1(\tau_1(v))} \right)}_{\geq 0}^{-1} > 2 \frac{F_1(\tau_1(v))}{\tau_1(v)}$$

whenever $\tau_1 < \bar{v}_1$. Assuming $r = 1$ and using (22) in Lemma 4, $\ln F_1(k(v))$ is strictly flatter than $F_1(\tau_1(v))$ at any point of intersection (which must occur at $k = \tau_1 < \bar{v}_1$) if

$$2 \left(\frac{F_1(k(v))}{k(v)} \right)^2 \geq \frac{f_2(v)}{v}.$$

Since F_1 is concave, $F_1(k)/k$ is decreasing in k . Likewise, $f_2(v)$ is increasing in v , by assumption. Finally, v is no smaller than $(\bar{v}_2 - \bar{v}_1)/2$, or τ_1 would not be defined. Thus, the inequality above is satisfied if

$$2 \left(\frac{1}{\bar{v}_1} \right)^2 \geq \frac{f_2(\bar{v}_2)}{(\bar{v}_2 - \bar{v}_1)/2}$$

or

$$\frac{\bar{v}_2 - \bar{v}_1}{\bar{v}_1^2} \geq f_2(\bar{v}_2),$$

as reported in the corollary (in fact, since τ_1 is not defined for $v > v_2^*$, $f_2(\bar{v}_2)$ can be replaced by $f_2(v_2^*)$, which is smaller). If this condition is met, then $ER(\bar{a}, 1)$ is single-peaked.

Since F_1 is concave and F_2 convex, the condition in Theorem 4 is equivalent to the first condition in the corollary. Moreover, the curvature assumptions imply that $\lambda(v)$ is convex and so $\lambda(v)/v$ is non-decreasing. To invoke the second part of Theorem 5, it is also necessary that

$$J_1(0) \geq \int_0^{v_2^*} J_2(x) \frac{f_2(x)}{F_2(v_2^*)} dx = -\frac{v_2^*(1 - F_2(v_2^*))}{F_2(v_2^*)},$$

but since $J_1(0) \geq -\bar{v}_1$, it is sufficient that

$$\frac{v_2^*(1 - F_2(v_2^*))}{F_2(v_2^*)} \geq \bar{v}_1.$$

Recall that $v_2^* = \arg \max x(1 - F_2(x))$ or $J_2(v_2^*) = 0$. Let m denote the median, i.e. $F_2(m) = \frac{1}{2}$. By convexity, $f_2(m) \geq (2m)^{-1}$ (otherwise, $f_2(x) \leq f_2(m) < (2m)^{-1}$ for all $x \in [0, m]$, but that contradicts $F_2(m) = \frac{1}{2}$). Thus,

$$J_2(m) = m - \frac{1 - \frac{1}{2}}{f_2(m)} \geq 0,$$

which in turn implies $v_2^* \leq m$. By definition, $v_2^*(1 - F_2(v_2^*)) \geq m(1 - F_2(m))$, and so

$$\frac{v_2^*(1 - F_2(v_2^*))}{F_2(v_2^*)} \geq \frac{m(1 - F_2(m))}{F_2(v_2^*)} \geq \frac{m(1 - F_2(m))}{F_2(m)} = m \geq \frac{1}{2}\bar{v}_2,$$

where the last inequality follows from the convexity of F_2 . To satisfy the condition in the second part of Theorem 5, it is thus sufficient that

$$\frac{1}{2}\bar{v}_2 \geq \bar{v}_1,$$

but, as explained after Theorem 4, this is implied by the first condition stated in the corollary. Theorems 1, 2, 4, 5(ii), and 6 now apply, which proves the corollary. ■

Proof of Proposition 4. Since $U_i(0) = 0$, $i = 1, 2$, (6) implies that H can be written as

$$H = \int_0^{\bar{v}_1} h_1(v)q_1(v)f_1(v)dv + \int_0^{\bar{v}_2} h_2(v)q_2(v)f_2(v)dv. \quad (23)$$

Thus, maximizing H is equivalent to maximizing the expected value of the reciprocal of the hazard rate of the winner (given $U_i(0) = 0$). However, h_1 and h_2 are decreasing, by assumption. In comparison, Myerson (1981) shows that when virtual valuation is not everywhere increasing, the revenue maximizing mechanism “irons” virtual valuation by pooling together an interval of types and assigning them the same probability of winning. Applying the same logic to H , it is optimal to pool together all of a bidder’s types, since $h_i(v)$ is nowhere increasing. The “ironed” value of $h_i(v)$ on $[0, \bar{v}_i]$ is

$$\int_0^{\bar{v}_i} h_i(v) f_i(v) dv = \int_0^{\bar{v}_i} (1 - F_i(v)) dv = \int_0^{\bar{v}_i} v f_i(v) dv,$$

which is the expected value of bidder i ’s type, EV_i . Since all types are pooled together, bidder i ’s probability of winning is independent of his type. Thus, the objective is to maximize $q_1 EV_1 + q_2 EV_2$. The unique solution is to give the object with probability one to bidder 2 since he has the highest expected valuation (due to first order stochastic dominance). As a result, bidder 1 never wins and his payoff is zero, the lowest possible individually rational payoff. This proves part (i). A head start can be used to implement the optimal mechanism. If bidder 2’s head start exceeds \bar{v}_1 , bidder 1 is deterred from participating, $q_1 = 0$. This proves part (ii). ■

Proof of Proposition 5. (i) Assume bidder 1 gets a head start, $a > 0$. Given (4) and (5), the derivative of H with respect to a is

$$f_2(v_2^c) F_1(v_1^c) \frac{\partial v_2^c}{\partial a} \left(\int_0^{v_1^c} h_1(v) \frac{f_1(v)}{F_1(v_1^c)} dv - h_2(v_2^c) \right),$$

which is analogous to (9). Since $h_2(0) > h_1(v)$ for any $v \in [0, \bar{v}_1]$, the derivative is negative when a (or v_2^c) is small. The first term in the parenthesis approaches EV_1 as a (or v_1^c and v_2^c) grows, whereas the second term approaches 0. Thus, the derivative is positive when a is large. $W = \min\{EU_1, EU_2\}$ must be inverse u-shaped because EU_1 and EU_2 move in opposite directions and cross once.

(ii) If F_2 is more dispersed than F_1 then $h_2(v) > h_1(\lambda(v))$ for all $v \in [0, \bar{v}_2)$, since $1 - F_2(v) = 1 - F_1(\lambda(v))$. Since h_i is decreasing in its argument, the function $\kappa^h(v)$ defined by $h_1(\kappa^h(v)) = h_2(v)$ is therefore below $\lambda(v)$ when $v < \bar{v}_2$. However, $\kappa^h(\bar{v}_2) = \lambda(\bar{v}_2) = \bar{v}_1$. Figure A.2 illustrates the relationship between $\kappa^h(v)$, $\lambda(v)$, and $k(v)$ when $r\lambda < 1$ (see Lemma 1). Since $h_2 > h_1$ above κ^h , bidder 2 should ideally win when the combination of types is *above* κ^h in order to maximize H . In other words, H would be minimized if k sits on top of κ^h ; the objective is to move the allocation away from κ^h . Now, an increase in r moves k closer to κ^h , but k remains

highest as long as $r\bar{\lambda} < 1$. Thus, the allocation changes because bidder 1 wins more often in the region above κ^h . This unambiguously lowers the expected value of the reciprocal of the hazard rate of the winner; H decreases.

Turning to W , recall that $EU_1 < EU_2$ when $r = 1$. Thus, $W = \min\{EU_1, EU_2\}$ is increasing in r when $EU_1 < EU_2$. This is the case for any r such that $r\bar{\lambda} < 1$. Since $k > \lambda$, bidder 2 wins with probability $F_1(k(v)) > F_1(\lambda(v)) = F_2(v)$. Hence, bidder 2's ex ante expected utility can be bounded below,

$$EU_2 > \int_0^{\bar{v}_2} h_2(x)F_2(x)f_2(x)dx = \int_0^{\bar{v}_2} (1-F_2(x))F_2(x)dx = \int_0^1 (1-z)z \frac{1}{f_2(F^{-2}(z))} dz,$$

where the substitution $z = F_2(x)$ was used to obtain the last equality. Since F_2 is more dispersed than F_1 , the density of the former is by definition smaller than the density of the latter at the same rank, or $f_2(F_2^{-1}(z)) < f_1(F_1^{-1}(z))$. Hence,

$$EU_2 > \int_0^1 (1-z)z \frac{1}{f_2(F^{-2}(z))} dz > \int_0^1 (1-z)z \frac{1}{f_1(F^{-1}(z))} dz > EU_1,$$

where the upper bound on EU_1 (the last inequality) was obtained analogously to the bound on EU_2 (bidder 1 wins with a probability below his rank when $r\bar{\lambda} < 1$). ■

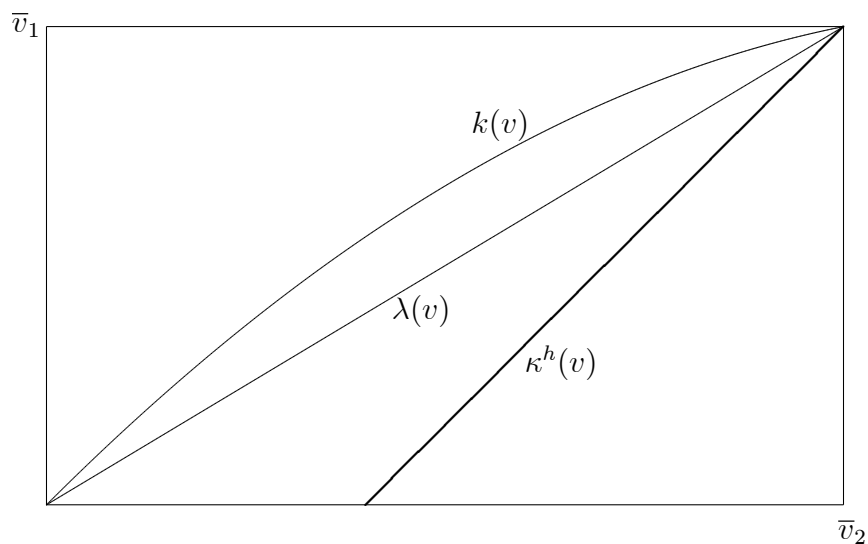


Figure A.2: $\kappa^h(v)$, $\lambda(v)$, and $k(v)$ when $r\bar{\lambda} < 1$.