FIRST PRICE AUCTIONS WITH GENERAL INFORMATION STRUCTURES: IMPLICATIONS FOR BIDDING AND REVENUE

By

Dirk Bergemann, Benjamin Brooks, and Stephen Morris

August 2015

COWLES FOUNDATION DISCUSSION PAPER NO. 2018

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

First Price Auctions with General Information Structures: Implications for Bidding and Revenue*

Dirk Bergemann Benjamin Brooks Stephen Morris

August 10, 2015

Abstract

This paper explores the consequences of information in sealed bid first price auctions. For a given symmetric and arbitrarily correlated prior distribution over valuations, we characterize the set of possible outcomes that can arise in a Bayesian equilibrium for some information structure. In particular, we characterize maximum and minimum revenue across all information structures when bidders may not know their own values, and maximum revenue when they do know their values. Revenue is maximized when buyers know who has the highest valuation, but the highest valuation buyer has partial information about others' values. Revenue is minimized when buyers are uncertain about whether they will win or lose and incentive constraints are binding for all upward bid deviations.

We provide further analytic results on possible welfare outcomes and report computational methods which work when we do not have analytic solutions. Many of our results generalize to asymmetric value distributions. We apply these results to study how entry fees and reserve prices impact the welfare bounds.

Keywords: First price auction, information structure, Bayes correlated equilibrium, private values, interdependent values, common values, revenue, surplus, welfare bounds, reserve price, entry fee.

JEL Classification: C72, D44, D82, D83.

*Bergemann: Department of Economics, Yale University, dirk.bergemann@yale.edu; Brooks: Becker Friedman Institute and Department of Economics, University of Chicago, babrooks@uchicago.edu; Morris: Department of Economics, Princeton University, smorris@princeton.edu. We acknowledge financial support through NSF Grant ICES 1215808. We thank Jason Hartline, Phil Reny, and Satoru Takahashi for helpful conversations. We are also grateful to seminar participants at Hebrew University, MIT, Northwestern, Queen Mary, Stanford, and TAU.

1 Introduction

The first price auction is an important institution used in a wide variety of settings. Its theoretical properties have been extensively studied for the last fifty years; it is a leading example in the expanding theoretical and empirical literature on auctions. Yet it may be argued that its properties are not well-understood outside of relatively special cases. Under complete information, the auction reduces to the straightforward problem of Bertrand competition. Under incomplete information, most work on private value first price auctions is carried out under the assumption of independent private values.[1] While the independence assumption is strong, a far more serious limitation of this environment is that bidders know nothing about others' valuations. In first price auctions, unlike in second price auctions, bidders' information about others' values is of first order strategic importance. Existing work with private values but neither complete nor no information about others' values—discussed further below—only deals with special cases. Results with interdependent values also rely on the assumption of one dimensional types, which means that there is a one-to-one map between buyers' information about their own values and their information about others' values. Significant further restrictions are also required on the nature of the correlation.[2]

In this paper, we derive results about equilibria that hold across *all* common prior information structures. Our results cover two cases. In the *known values* case, bidders are assumed to know their own values for the object being auctioned. We study what can happen for all information structures specifying bidders' information about other bidders' values, for any given joint prior distribution of values. This model thus generalizes the classical analysis of private value auctions.[3] In the *unknown values* case, we allow for the possibility that bidders do not necessarily know their own values. Unknown values therefore embeds all possible interdependent values environments, e.g., the case of pure common values. For each scenario, our analytic results focus on the welfare outcomes of revenue and the bidders' surpluses. We identify the information structures and Bayesian equilibria that generate extreme points of the set of possible surplus and revenue pairs. We also use computational methods to explore features of the model for which analytical results are unavailable.

It is useful start with a classical example, which we shall return to throughout the paper to illustrate our results. Suppose that there are two bidders in a first price auction, with values independently and uniformly distributed between 0 and 1. Figure 1 describes the set

[1] See Kaplan and Zamir (2014) for a review of recent developments in the theory of first price auctions.

[2] The classic analysis of Milgrom and Weber (1982) relies on a strong affiliation restriction.

[3] Throughout, we will use the term *private values* to denote an information structure in which each bidder receives a signal which is equal to their own value. Bidders therefore form beliefs about others' values (and hence others' signals) through Bayesian updating from the prior. In contrast, a known values model may involve the bidders obtaining additional signals about others' values and beliefs.

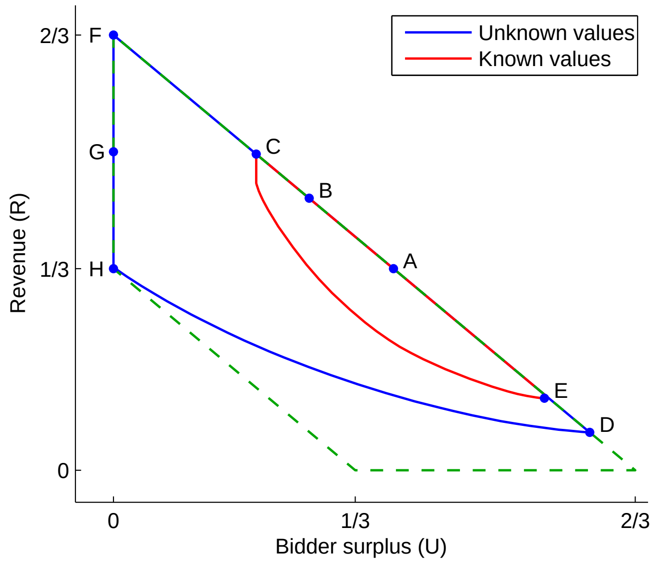

Figure 1: The set of revenue/total bidder surplus pairs that can arise in a BCE. Computed for uniform distribution with grids of 10 valuations and 50 bids between 0 and 1. The axes have been re-scaled to match moments with the continuum limit; for the discretized example, the efficient surplus and minimum surplus are respectively 41/60 and 19/60, as opposed to their limit values of 2/3 and 1/3.

of pairs of expected revenue and expected total bidder surplus that can arise in this setting. Feasibility alone will always impose some elementary restrictions. The sum of revenue and bidder surplus cannot be driven above the expectation of the highest value—2/3 in the example. Since the first price auction without reserve price always allocates the object to some bidder, the sum of revenue and bidder surplus cannot fall below the expected lowest valuation—1/3 in this example. Also, it is not feasible for revenue to fall below zero, and individual rationality of bidders implies that bidder surplus cannot fall below zero as well. Thus, the green trapezoid represents restrictions on revenue and bidder surplus that are implied by feasibility and participation alone. Now, under either complete information or under independent private values, in which bidders know their own value but not the other bidder's value, revenue equals the expectation of the lower value (1/3), while expected bidder surplus is the difference between the efficient surplus (2/3) and revenue (1/3). This outcome is represented as point A in Figure 1. Also shown in Figure 1 are the sets of surplus-revenue pairs that can arise in equilibrium for any information structure, with the larger blue region corresponding to the unknown values case and the smaller red region corresponding to the known values case. Note that while we use this leading example to motivate our results, our results apply to the case with many bidders with general and arbitrarily correlated distributions of values.

We will start by giving some intuition for how revenue could be increased (and bidder surplus decreased) while remaining on the efficiency frontier, so that we move to the north-west from point A in Figure 1. Suppose that in the uniform example each bidder learns if they have the high or the low value. In addition, if the low value was relatively close to the high value (in particular, greater than one half of the high value), then the high value bidder learns the low value precisely. But if the low value is low relative to the high value (in particular, less than half of the high value), then the high value bidder only learns that the low value is less than half his value. In the equilibrium we construct for this information structure, the low value bidder will always lose and any bid that is weakly less than the value is optimal, regardless of what else he knows about the high value. If the high value bidder knows the low value, she will—as in the complete information case—bid the low value, which is supported by the low type's randomizing close to his value. If the high type does not know the low value, her belief will be that it is uniformly distributed between 0 and half her (high) value v, and the low type will be bidding his value. But the high type's best response in this case is to bid $v/2$: (i) bidding $v/2$ will give probability 1 of winning, so there is no incentive to bid higher; (ii) bidding $b < v/2$ will give a probability of winning of $2b/v$ and thus expected surplus $(v - b)\, 2b/v$, which is maximized by setting $b = v/2$. One can verify that this gives rise to point B in Figure 1, where revenue is $5/12$ and bidder surplus is $1/4$.[4] An intuition from this example is that giving bidders information about their rank and giving the highest value bidder only partial information about lower values increases revenue and reduces bidder surplus by inducing bidders to always bid more than under complete information while maintaining efficiency.

Our first main result is a characterization of the maximal revenue in the known values scenario. To establish this result, we first note an easy lower bound on bidder surplus. It is always an option for a bidder to choose his bid as a function of his value alone and not condition on his information about other bidders' values. Even if the bidder is as pessimistic as possible about others' strategies—believing that other bidders will all bid their values—such a bidder can guarantee himself a certain expected surplus as a function of his value. Now, an ex-ante lower bound on that bidder's surplus is the expectation of that value dependent minimum surplus.[5] Moreover, an upper bound on revenue is the efficient surplus minus the sum of all bidders' surplus lower bounds. In fact, we construct an information structure and equilibrium—generalizing that of the previous paragraph—where this bound

[4] A formal version of this example is discussed in Appendix B.

[5] We are grateful to Satoru Takahashi for suggesting this bound to us.

is exactly attained: Each bidder observes if he has the highest value or not, and those who do not have the highest value always lose and bid their value. If the highest valuation bidder knew nothing more, the optimal shading of his bid would mean that he would sometimes lose to lower valuation bidders, which would undermine the proposed strategies for those with low valuations. If the highest valuation bidder knew the second highest value precisely, on the other hand, he would bid that value and get more surplus than his lower bound. We show, however, that there is an intermediate amount of information for the highest valuation bidder such that he always bids more than the second highest value, and therefore wins the auction, but is always indifferent between his equilibrium bid and the bid associated with his surplus lower bound for that value. Thus, the allocation is efficient, all bidders get only their lower bound surplus, and revenue attains its upper bound. This argument and result holds for any number of bidders and any distribution of private values. In the uniform example, this gives revenue of 1/2 and bidder surplus of 1/6 and is represented by point C in Figure 1.

We now give some intuition for how revenue could be decreased (and bidder surplus increased) while again remaining on the efficiency frontier, so that we move to the southeast from point A in Figure 1. Suppose that in the uniform example, each bidder observed a signal of the other bidder's value. With high probability, each bidder's signal is equal to the true value of the other bidder. However, with low probability, both bidders observe signals which are below both bidders' values, with the signal of the high value bidder above the signal of the low value bidder. There is an equilibrium where each bidder bids the minimum of his valuation and his signal of the other bidder's value. If it were not for the low probability event, this would correspond to equilibrium in the complete information case. But with small probability, the high value bidder will be bidding less than the low value bidder's value and winning. Since the allocation is still efficient and winning bids are shifted downwards, revenue must decrease. In order for this to be an equilibrium, a bidder must be unsure if he has the high value or the low value; and we must ensure that a bidder does not have an incentive to deviate to a higher bid so that he wins even if the other bidder has a higher value.[6] This will be true only if the probability of the low signal event is small enough. An intuition from this example is that in order to raise revenue, bidders must be uncertain about their rank and the informational content of the signal; and the key constraint on decreasing revenue is the incentive to bidders to increase their bids in order to win even if they have the low value.

Our second main result is a characterization of the minimal revenue in the unknown values scenario. We describe what we know about minimum revenue in the known values

[6] A fully specified and formal version of this example is discussed in Appendix B.

case below, but we focus on the unknown values case because we have a sharp and insightful characterization. Here our approach is to study a relaxation of the original problem where we ignore bidders' incentives to deviate to lower bids and focus only on bidders' incentives to deviate to higher bids. Moreover, we restrict attention to what we call *uniform upward deviations*, in which, for some bid b, the bidder deviates up to b whenever his strategy specifies that he should bid less than or equal to b. In an equilibrium, it must be the case that such a deviation does not increase bidder surplus in ex-ante terms. We show that revenue is minimized among solutions to this relaxed program only if all these uniform upward incentive constraints are binding. This in turn implies that minimum revenue arising in the relaxed program can be characterized as the solution to a differential equation. We can then establish that if a certain revenue can be attained in a solution to the relaxed program, we can also construct an information structure and strategy profile where that minimum revenue is attained in a Bayes Nash equilibrium, so that all other deviations (including deviations to lower bids) are not optimal. A key feature of this construction is that bidders are uncertain about whether they have the highest valuation or not and they are indifferent between all bids between their equilibrium bid and the maximum bid that is made in equilibrium. Moreover, this information structure and strategy profile always induce an efficient allocation, implying that bidder surplus is maximized. In the uniform example, minimum revenue is approximately 0.096 and maximum bidder surplus is 0.571, and is represented by point D in Figure 1.

The unknown values minimum revenue immediately implies a lower bound on the known values minimum revenue. As we can see in the uniform example, this lower bound is not tight. We also explore the extent to which our methodology can be applied to understand minimum revenue in the known values case. When bidders can have only one of two possible values,[7] we can use the relaxed approach to solve for the minimum revenue exactly. Beyond binary values, there are certain cases for which we can solve a generalized relaxed program to provide a lower bond on revenue. However, this bound will generally not be tight. Point E in Figure 1 describes minimum revenue in the uniform example in the known values case, but is identified by computation and not an analytic characterization.

The set of points in Figure 1 where bidder surplus is driven to zero under the unknown values scenario are also of interest. If bidders know nothing about their own and others' values, and the distribution of values is symmetric, then the problem reduces to Bertrand competition based on expected values. Bidder surplus will be zero and, because the allocation does not depend on ex-post values, the outcome will be inefficient. In the uniform example, each bidder will bid his expected value of 1/2 and get zero expected surplus. This corresponds

[7]Studied earlier by Fang and Morris (2006) and Āzacis and Vida (2015).

to point G in Figure 1. Alternatively, consider an information structure when both bidders are told only the highest value, without being told to whom it belongs. If we suppose that ties are broken in favor of the bidder with the highest value, then there is an equilibrium where everyone bids that highest value, bidder surplus is zero, and revenue is equal to the efficient surplus. Indeed, we show in the body of the paper how this outcome can be approximated even when ties are broken uniformly. This corresponds to point F in Figure 1. Finally, in the two player and independent value case, if each bidder is told only the other bidder's value, there is a symmetric equilibrium where bids are monotonic in bidder's information about others' values, bidders are indifferent to all bids, the allocation is maximally inefficient. In the uniform example, this corresponds to point H in Figure 1.

In analyzing what can happen in all equilibria for all information structures under the two scenarios, we can restrict attention to information structures where bidders' signals are identified with the bids that they are going to make in equilibrium. A more abstract way of making this observation is that if we fix an incomplete information game, including a description of some initial information, there is an equivalence between the set of what can happen in all equilibria where players observe more information and a class of incomplete information correlated equilibria. Bergemann and Morris (2013) labeled this class of incomplete information correlated equilibria *Bayes correlated equilibria* (BCE), and Bergemann and Morris (2015) consider properties of BCE in general games.[8] This paper can be seen as an application of this methodology in a setting that is significantly more challenging than previous applications. Our characterization of maximum revenue in the known values scenario exploits insights from Bergemann, Brooks, and Morris (2015a), but characterization of minimum revenue in the unknown values case develops new arguments. A significant advantage of this approach is that solving for all equilibria under all information structures corresponds to solving a linear programming program. We exploit this structure in our computation of Bayes correlated equilibria.[9] In fact, we were able to identify analytic solutions for general settings by discovering structure in computations. An abstract observation about BCE in Bergemann and Morris (2015) is that more information always reduces the set of BCE: intuitively, more information can only impose more incentive constraints. The inclusion of the set of possible welfare outcomes under the known values scenario in the corresponding set for unknown values is a clean illustration of that insight.

[8]This builds on the work of Forges (1993) whose "Bayesian solution" characterizes what can happen if players can observe more information but only information that is measurable with respect to the join of players' information. Thus, BCE and the Bayesian solution coincide under known values but diverge under unknown values.

[9]A series of computation tools for computing BCE in general games is available through the authors' websites.

We have motivated our results with an example involving a continuum of values and a continuum of bids. In the formal analysis of this paper, however, we focus on the case of a finite set of values and a continuum of bids.[10] This modelling choice allows us to state our main results and arguments in the most transparent way. In our computations, we work with finite values and finite bids, but we can assume fine grids in each case. The extension of our results to continuum values and continuum bids raises complications in the statement of our results but does not change the structure of equilibria. We will occasionally refer again to continuum bid case, and the uniform example discussed above, to develop intuition for our results.

Our primary focus in this paper is on developing insights about how information can affect outcomes in the first price auction and on the qualitative properties of the information structure that lead to different outcomes. There are many different uses for which our methodology could be used. For example, it can be used for counterfactual exercises, for identifying restrictions on possible value distributions, or for comparison with other mechanisms. We present one such illustration in this paper, examining what happens to the maximum possible and minimum possible revenue in the first price auction if a reserve price is added. A striking finding for the uniform example is that minimum possible revenue is single peaked and is maximized at a reserve price which is higher that the optimal reserve price under the Myersonian optimal auction. The maximum possible revenue is decreasing in the reserve price is decreasing in the reserve price, but the decrease in maximum possible revenue is small compared with the increase in minimum possible revenue to its maximum. In this sense, high reserve prices deliver more robust revenue performance.

The present inquiry is related to a number of papers which have studied behavior in the first price auction under alternative information structures. First, there is a small number of papers considering the privates value case where bidders have partial information about other bidders' independent values. Landsberger et al. (2001) consider the case where bidders observe their values and also their rank (i.e., whether they have the highest valuation). They showed that revenue increases when bidders know their ranks, although it induces inefficiency. Kim and Che (2004) consider the case where bidders are divided into groups and each bidder knows the valuations of those in his group, but knows nothing about the valuations of those outside the group. Fang and Morris (2006) and Āzacis and Vida (2015) analyze the two bidder binary value case where each bidder observes a conditionally independent signal of the other bidder's valuation. In each case, revenue falls and bidder surplus increases.

[10] This requires a technical extension of the definition and characterization of Bayes correlated equilibrium from finite action to continuum action games, an extension that is described in our note Bergemann, Brooks, and Morris (2015b).

The rest of this paper proceeds as follows. In Section 2, we describe our basic model of the first price auction, with a fixed distribution of ex-post values but many possible specifications of beliefs. In Section 3, we characterize tight bounds on maximum revenue and minimum bidder surplus over all possible specifications of bidders' beliefs consistent with a fixed ex-post distribution of values, first for the no-information model and then for the known values model. In Section 4, we repeat the analysis for the objectives of minimum revenue and maximum bidder surplus. Section 5 presents extensions of the model, and Section 6 concludes. Omitted proofs are contained in the Appendix.

2 Model

There are N potential buyers of a single unit of a good that can be produced at zero marginal cost. The bidders are indexed by $i \in \{1, \ldots, N\}$. Bidders' values are drawn from a compact set $V \subset \mathbb{R}_+ = [0, \infty)$. For all of our formal results, we will assume that V is finite and denumerated as

$$V = \{v^1, \ldots, v^K\},$$

where K is the number of possible valuations. The profile of valuations is $v = (v_1, \ldots, v_N) \in V^N$. There is a fixed common prior joint distribution over values which we denote by $p \in \Delta\left(V^N\right)$, which is common knowledge among the bidders.[11] Throughout most of the paper, we will assume that p is symmetric. Many of our results extend to asymmetric distributions of values, which we will discuss further in Section 5.5. The buyers are participating in a first price auction in which they submit real-valued bids $b_i \in \mathbb{R}_+$.[12] For a profile of bids $b \in \mathbb{R}_+^N$, we denote the set of high bidders by $W(b) = \{i \,|b_i = \max b\}$. Bidder i receives the good with probability

$$q_i(b) = \frac{\mathbb{I}_{i \in W(b)}}{|W(b)|},$$

where the indicator function $\mathbb{I}_X$ is equal to 1 if the event X occurs and zero otherwise. In other words, the high bidder receives the good and ties are broken uniformly.

We assume that bidders may receive additional information about the profile of values, beyond knowing the prior distribution. This additional information comes in the form of signals, which we will call types, that are correlated with the profile of valuations. A *type*

[11] For a metric space X, $\Delta(X)$ denotes the set of Borel measures on X.

[12] The assumption that bids are non-negative is without loss of generality. We can equivalently assume that the set of possible bids is bounded below and contains the convex hull of V.

space is a collection $\mathcal{T} = \left(\{T_i\}_{i=1}^{N}, \mu\right)$, where the T_i are compact metric spaces and $\mu : V^N \to \Delta(T)$ maps profiles of values into Borel probability measures over $T = \times_{i=1}^{N} T_i$. The interpretation is that T_i is the set of bidder i's types and μ describes the conditional distribution of types given values. Write $\mu_i(\cdot|v_i) \in \Delta(T_i)$ for a version of the conditional distribution of buyer i's type given buyer i's value, i.e., μ_i almost surely satisfies

$$\mu_i(X|v_i) = \frac{\sum_{v_{-i} \in V^{N-1}} p(v_i, v_{-i}) \mu(X \times T_{-i}|v_i, v_{-i})}{\sum_{v_{-i} \in V^{N-1}} p(v_i, v_{-i})}.$$

for all Borel subsets $X \subseteq T_i$. Let $\phi_i(v_i|t_i)$ be a version of the conditional distribution of v_i given t_i. We will say that $\mathcal{T}$ is a *known values* type space if for all $t_i \in T_i$ the support of the posterior belief $\phi_i(\cdot|t_i)$ is a singleton for all t_i, i.e., $\operatorname{supp} \phi_i(\cdot|t_i) = \{v_i\}$. Otherwise, $\mathcal{T}$ is an *unknown values* type space.

For a fixed type space $\mathcal{T}$, the first price auction is a game of incomplete information in which bidders' strategies are measurable mappings $\beta_i : T_i \to \Delta(B)$. Let $\mathcal{B}_i$ denote the set of strategies for buyer i. Fixing a profile of strategies $\beta \in \mathcal{B} = \times_{i=1}^{N} \mathcal{B}_i$, bidder i's surplus from the auction is

$$U_i(\beta) = \sum_{v \in V^N} p(v) \int_{t \in T} \mu(dt|v) \int_{b \in \mathbb{R}_+^N} q_i(b)(v_i - b_i) \beta(db|t),$$

where β is the unique measure induced on the product σ-algebra by the product measure $\beta_1 \times \cdots \times \beta_N$ (cf. Cohn, 1980, Theorem 5.1.4).

We shall restrict attention to strategies which satisfy a version of weakly undominated. In particular, we require that for every type t_i, the support of the possibly randomized bid does not exceed the largest possible valuation in the support of his posterior beliefs given t_i, or:

$$\operatorname{supp} \beta_i(t_i) \subseteq [0, \max \operatorname{supp} \phi_i(\cdot|t_i)]. \tag{1}$$

This restriction is weaker than weakly undominated; in the case of complete information, it allows bidder i to bid up to and including his valuation v_i, whereas the conventional definition of weakly undominated would also rule out bidding v_i, and thereby exclude the unique equilibrium strategy in which two (or more) bidders who share the highest valuation bid their valuation.

The profile $\beta \in \mathcal{B}$ is a *Bayes Nash equilibrium* (BNE) if and only if, for all i,

$$\begin{aligned}&\sum_{v \in V^N} p(v) \int_{t \in T} \mu(dt|v) \int_{b \in \mathbb{R}_+^N} q_i(b) \left(v_i - b_i\right) \beta(db|t) \\ &\qquad \geq \sum_{v \in V^N} p(v) \int_{t \in T} \mu(dt|v) \int_{b \in \mathbb{R}_+^N} q_i(b) \left(v_i - b_i\right) (\beta_i', \beta_{-i})(db|t)\end{aligned} \tag{2}$$

for all $\beta_i' \in \mathcal{B}_i$.

Let $F : V^N \to \Delta\left(\mathbb{R}_+^N\right)$ denote a mapping from profiles of values to probability distributions over profiles of bids. We say that F is a *Bayes correlated equilibrium* (BCE) if, for all Borel measurable deviations $\sigma_i : \mathbb{R}_+ \to \mathbb{R}_+$,

$$\begin{aligned}&\sum_{v \in V^N} p(v) \int_{b \in \mathbb{R}_+^N} q_i(b)(v_i - b_i) F(db|v) \\ &\qquad \geq \sum_{v \in V^N} p(v) \int_{b \in \mathbb{R}_+^N} q_i(\sigma_i(b_i), b_{-i}) \left(v_i - \sigma_i(b_i)\right) F(db|v).\end{aligned} \tag{3.1}$$

In addition, F is a *known values BCE* if for all $v_i \in V$ and for all $\sigma_i : \mathbb{R}_+ \to \mathbb{R}_+$ that are Borel measurable,

$$\begin{aligned}&\sum_{v_{-i} \in V^{N-1}} p(v_i, v_{-i}) \int_{b \in \mathbb{R}_+^N} q_i(b) \left(v_i - b_i\right) F\left(db|v_i, v_{-i}\right) \\ &\qquad \geq \sum_{v_{-i} \in V^{N-1}} p(v_i, v_{-i}) \int_{b \in \mathbb{R}_+^N} q_i(\sigma_i(b_i), b_{-i}) \left(v_i - \sigma_i(b_i)\right) F(db|v_i, v_{-i}).\end{aligned} \tag{3.2}$$

We note that the set of BCE is the intersection of the set of mappings from V^N to Borel measures over bid profiles with an (infinite) family of linear inequalities (3). Thus, the set of BCE is convex and bounded, although because the payoff function is discontinuous, the set of BCE is not compact.

Let us interpret these conditions. A BCE is an extension of the notion of correlated equilibrium to games of incomplete information, where players' actions are allowed to be correlated, not just with one another, but also with the underlying payoff relevant states. The incentive compatibility condition is that conditional on the equilibrium action, that action must be optimal. The equilibrium can be viewed as a set of private bid recommendations offered by a disinterested mediator who knows the true profile of values and privately recommends bids to each bidder. A BCE is a recommendation rule for the mediator such that the bidders would want to follow the recommendation. The distinction between unknown and known value BCE is that the known value definition requires that each recommendation is followed even when each bidder knows his own value.

BCE are useful for our analysis because of the epistemic relationship between the set of BCE and the set of outcomes that can be induced by a BNE for some type space $\mathcal{T}$. In particular, let us say that the conditional distributions $F : V^N \to \Delta\left(\mathbb{R}_+^N\right)$ are *induced* by the type space $\mathcal{T}$ and the BNE β if for all v and Borel subsets $X \subseteq \mathbb{R}_+^N$,

$$F(X|v) = \int_{t \in T} \beta(X|t)\mu(dt|v).$$

In earlier work, Bergemann and Morris (2015) define the notion of Bayes correlated equilibria and investigate the relationship between BCE and Bayes Nash equilibria in canonical finite player, finite action, finite state games. In the language of the current setting, their Theorem 1 can be restated as follows:

Theorem 1 (Bayes Correlated Equilibrium)**.** *F is induced by some (known value) type space $\mathcal{T}$ and some BNE β if and only if F is an unknown (known) value BCE.*

In a companion note, Bergemann, Brooks, and Morris (2015b), we extend this and other results from finite games to infinite games, in particular infinite actions and infinite states.

Our goal is to analyze how welfare varies across type spaces and BNE for a fixed distribution over values. Because of Theorem 1, we can equivalently ask how welfare varies across BCE. The welfare outcomes that we will investigate are:

$$(i) \text{ bidder surplus: } U(F) = \sum_{i=1}^{N} \sum_{v \in V^N} p(v) \int_{b \in \mathbb{R}_+^N} q_i(b)\left(v_i - b_i\right) F(db|v);$$

$$(ii) \text{ revenue: } R(F) = \sum_{i=1}^{N} \sum_{v \in V^N} p(v) \int_{b \in \mathbb{R}_+^N} q_i(b) b_i F(db|t);$$

$$(iii) \text{ total surplus: } S(F) = \sum_{i=1}^{N} \sum_{v \in V^N} p(v) \int_{b \in \mathbb{R}_+^N} q_i(b) v_i F(db|t).$$

We will study these objectives for both known and unknown value BCE.

3 Maximum Revenue and Minimum Bidder Surplus

In this section we explore the limits of how large revenue can be and how low bidder surplus can be across all possible information structures. We begin with the known values model. Bidders can always guarantee themselves positive surplus due to the information about their own values. We will characterize the exact limits of how much surplus the bidders need to receive and how much revenue the seller can earn. We illustrate the result with an example of two bidders with uniformly distributed values. We then briefly consider the unknown values model, for which there are trivial information structures that yield zero bidder surplus. Moreover, we can construct beliefs such that the seller extracts all the surplus.

3.1 Known Values

In the known values model, each bidder i is assumed to know his value v_i and that the profile of valuations is drawn from the common prior p. As each bidder always know his value for the object, any weakly undominated strategy profile requires that the bidders never bid above their values. Thus, each bidder knows that their opponents cannot be using a more aggressive strategy than bidding their values. If this were in fact the strategy that others are using, bidder i would face a bid distribution with cumulative distribution

$$P_i^{(2)}(b|v_i) \triangleq \sum_{\{v_{-i} \in V^{N-1} | \max_{j \neq i} v_j \leq b\}} p(v_i, v_{-i}).$$

Against this most aggressive bidding behavior by his competitors, bidder i would optimally bid

$$b_i^*(v_i) \triangleq \max \left\{ \arg\max_{b \in \mathbb{R}_+} \left\{ (v_i - b) P_i^{(2)}(b|v_i) \right\} \right\}, \tag{4}$$

where we take without loss of generality the largest optimal bid in case there are multiple solutions. Note that the optimal bid must be attained at a value $v^{k_i^*(v_i)} \in V$. It follows that bidder i with value v_i must receive in any equilibrium at least the surplus

$$\underline{U}_i(v_i) \triangleq (v_i - b_i^*(v_i)) \, P_i^{(2)}(b_i^*(v_i)|v_i).$$

For if the equilibrium surplus were lower, bidder i could deviate upwards to $b_i = b_i^*(v_i) + \epsilon$ and guarantee himself surplus arbitrarily close to $\underline{U}_i(v_i)$. This implies that in ex-ante terms, bidder i must receive at least

$$\underline{U}_i \triangleq \sum_{v \in V^N} p(v) \underline{U}_i(v_i).$$

The main result of this section is to argue that this lower bound is in fact tight: there are BCE in which bidder i receives *exactly* $\underline{U}_i$ in surplus. Moreover, it is possible for all bidders to receive this surplus at the same time, and in an equilibrium in which the allocation is socially efficient. Thus, these equilibria simultaneously minimize bidder surplus and maximize the revenue of the seller at the level

$$\overline{R} = \overline{S} - \sum_{i=1}^{N} \underline{U}_i. \tag{5}$$

We should emphasize two aspects of the best response $b_i^*(v_i)$. First, the the strategy profile in which each bidder bids according to $b_i^*(v_i)$ clearly does not form an equilibrium profile. After all, each bidder's best response $b_i^*(v_i)$ will generally involve shading, whereas the conjecture that justifies $b_i^*(v_i)$ is that others are bidding their values. Second, the using the strategy $b_i^*(v_i)$ against the conjecture of all other bidders bidding their value would not yield an efficient allocation either, as bidder i would sometimes loose against bidders with lower values. In the equilibrium we construct below, the losing bidders will know that they face a bidder with a higher valuation, and will have an incentive to bid their value, thus justifying the conjecture by the winning bidder implicit in (4). Moreover, the winning bidder will know that he has the highest valuation and will have some additional information about the conditional distribution about the losing bidders' valuation that will lead him to bid higher than suggested by $b_i^*(v_i)$, and in fact sufficiently high to maintain an efficient outcome against the losing bidders. Thus, we will re-establish efficiency and increase the bids of the winning bidder to attain the upper bound in the seller's revenue as suggested by (5).

Theorem 2 (Max revenue and min bidder surplus for known values)**.** *For any distribution p, there exists a BCE in which $U_i(F) = \underline{U}_i$ for all i and $R(F) = \overline{R}$. This BCE simultaneously minimizes bidders' surpluses and maximizes revenue over all BCE.*

We construct the information structure that attains the revenue upper bound of Theorem 2 in Appendix A, and here we will give an informal description. The BCE that maximizes the revenue of the seller—and concurrently minimizes the surplus of the bidders—can alternatively be represented by a Bayes Nash equilibrium with an information structure that is close, but in one important aspect, distinct from the complete information structure. In the information structure that we construct, every losing bidder is informed about the valuations of all of the bidders. The winning bidder, on the other hand, is informed about the fact that his valuation is the highest valuation, but he is not informed about the realized second highest valuation. More precisely, relative to just knowing the conditional distribution of $v^{(2)}$ given his own valuation v_i, he receives some additional information about the distribution of losing valuations that is sufficient to make him indifferent between bidding $b_i^*(v_i)$ and bidding v^l with $b_i^*(v_i) < v^l$. Namely, when the winning bidder is supposed to make a bid $b_i = v^l$, the conditional distribution of second highest distribution "oversamples" v^l relative to its true frequency in the conditional distribution of $v^{(2)}$. In particular, the ratio of mass of type v^l to the mass of valuations $v^{l'}$ with $l' < l$ is only $1 : \alpha^l$, with $\alpha^l < 1$, as opposed to $1 : 1$ according to the prior distribution. The weight α^l is chosen to satisfy the following indifference condition:

$$(v_i - v^l)\left(\alpha^l P_i^{(2)}(v^{l-1}|v_i) + p_i^{(2)}(v^l|v_i)\right) = (v_i - b_i^*(v_i))\alpha^l P_i^{(2)}(b_i^*(v_i)|v_i). \tag{6}$$

As long as $b_i^*(v_i) < v^l$, this condition implies that $\alpha^l < 1$, because at α^l, we know that the winning bidder would not chose $b_i = v^l$ but rather $b_i = b_i^*(v_i)$. But clearly, for every $v^l \leq v_i$, we can find $\alpha^l \in [0, 1)$ such that oversampling the highest valuation v^l and keeping all the lower valuations in the same proportion as they appear in the conditional distribution of $v^{(2)}$ guarantees the indifference. Thus, when the winning bidder makes a bid $b_i = v^l$, he knows that with sufficiently high probability the second highest valuations is v^l and hence he does not have an incentive to lower his bid below v^l. As in the complete information structure, the winning bidder knows that with high probability his bid is just sufficient to fend off the second highest valuation v^l, but importantly, there is some residual uncertainty, expressed by $\alpha^l > 0$, that the second highest valuation is below v^l. We can now see how this information structure induces larger revenues for the seller than the complete information structure. Namely, the residual uncertainty about whether the second highest valuation is indeed v^l or below v^l induces the winner to bid v^l, even though a lower bid, one that would just match the second highest valuation would sometimes have sufficed to win the auction.

We note that whenever we oversample v^l relative to its frequency in the conditional distribution of $v^{(2)}$, we maintain the relative proportions of valuations below v^l as they appear in the conditional distribution of $v^{(2)}$. Thus, the indifference condition (6) suggests an algorithm to construct the information structure for the winning bidder. Namely, start with the highest possible bid for winning bidder v_i, clearly $b_i = v_i$, and lower the winning bid b_i gradually until b_i reaches $b_i^*(v_i)$. At each bid increment $b_i = v^l$, the remaining second highest valuations v^l are exhausted and all that remains are valuations below v^l (importantly in their original proportion according to the conditional distribution of $v^{(2)}$). When the algorithm finally reaches the lowest bid recommendation $b_i = b_i^*(v_i)$, by construction $\alpha^{k_i^*(v_i)} = 1$, and all the remaining second highest valuations are recovered.

From the point of view of the bidder with the second highest valuation $v^{(2)}$, who is one of the losing bidders, the distribution of the winning bids against which he loses can also be recovered from the above construction of the BCE. Namely, the losing bidder with valuation $v^{(2)}$ will face bids ranging from $v^{(2)}$ all the way up to $v^{(1)}$. The probability of a bid v^l is given by

$$x^l = \prod_{l'=l+1}^{K} (1 - \alpha^{l'})$$

if $v^l = v^{(2)}$, and by

$$y^l = \alpha^l x^l$$

if $v^l > v^{(2)}$, and we can easily verify that $x^l + \sum_{l'>l} y^{l'} = 1$ for all l in the relevant range, that is $b_i^*(v_i) \leq v^l \leq v_i$.

The proof of Theorem 2 uses ideas from our related paper on monopoly price discrimination, Bergemann, Brooks, and Morris (2015a) (BBM). Essentially, a bidder who is facing a fixed distribution of opponents' bids is in a situation comparable to that of a monopolist who is facing a fixed distribution of consumers' valuations, each of whom demands a single unit of a good that can be produced at zero cost. In the case of the auction, submitting a bid b will result in a surplus of $v - b$ when others' bids are less than b. The monopolist, on the other hand, earns a revenue of p when consumers' valuations v are greater than p. Thus, we have simply reversed the sign of how the agent's action, the bid or the price enters the objective function, but it is still the case that the relevant surplus is a linear function of that action times the mass in a one-sided tail of a distribution that is, from the bidding agent's perspective, exogenous. The construction of the information structure in Theorem 2 corresponds to a particular information structure discussed in BBM, where we change the conditional distribution of losing valuation only *locally*, namely at the upper end of the distribution. The general results of BBM are proved using a geometric argument to construct extremal distributions that changed the composition of the valuations everywhere, that is *globally*.

In the monopoly case, partial information about consumers' values facilitates third degree price discrimination, whereby the monopolist offers different prices to different segments of consumers. BBM show that it is possible to structure information for the monopolist by creating pools of consumers so that there are enough low valuation consumers to justify dropping the price, but each pool also contains a fair number of high-valuation consumers who benefit when prices fall. In fact, regardless of the ex-ante distribution of consumers' valuations, it is always possible to construct these pools so that prices drop enough so that all consumers purchase the good (as long as they value it above marginal cost), but the monopolist is just indifferent to dropping the price, so that the monopoly profits do not increase relative to the no-discrimination outcome.

In the current setting, partial information about others' bids means that the bidder receives information that is correlated with the other bidders' information, and therefore with their equilibrium actions. Because of the similar features of the payoff function, the same logic that allows for efficient price discrimination with zero benefit to a monopolist means that it is possible to structure information for the high-valuation bidder so that he always outbids his opponents, but does not benefit from the additional information. Thus, the bidder is just as well off as if he knew nothing besides the prior distribution over others' values, and best responded to the conditional distribution of others' valuations. Importantly, while the distribution of consumers' values in the monopoly setting is exogenous, the distribution of others' bids must be generated by best responses which have to be supported in equilibrium. However, this can be accomplished in a straightforward manner when maximizing revenue and minimizing bidder surplus, namely, by guaranteeing that the losing bidders will lose to bids that are greater or equal to their own known values.

3.2 Unknown Values

If the distribution p is symmetric, then each bidder has the same ex-ante valuation for the good, which is

$$\mathbb{E}[v_i] = \sum_{v \in V^N} p(v) v_i.$$

In the absence of any additional information, there is a simple equilibrium in which every bidder bids $b_i = \mathbb{E}[v_i]$. All bidders tie and earn a surplus of 0 in equilibrium. This equilibrium is, however, inefficient, since the winner need not be the bidder with the maximum ex-post valuation.

Perhaps not too surprisingly, there are also equilibria in weakly undominated strategies such that the outcome is very nearly efficient, and yet bidder surplus is arbitrarily close to zero. This result reflects, in some sense, the weakness of weakly undominated strategies as a refinement under unknown values. It may be that beliefs are such that bidders are willing to bid a large amount because they think that the bid is less than their value conditional on winning, although conditional on losing their value might be quite a bit lower. This aggressive bidding behavior when losing supports, in turn, very aggressive bidding on the part of the winning bidder. Indeed, we construct such an equilibrium in the proof of the following theorem:

Theorem 3 (Max revenue and min bidder surplus with unknown values)**.** *For all $\epsilon > 0$, there exists a weakly undominated BCE F such that $R(F) > \overline{S} - \epsilon$ and $U(F) < \epsilon$.*

4 Minimum Revenue and Maximum Bidder Surplus

We now turn our attention to the lower limit of revenue and the upper limit of total bidder surplus. The main result for this section is a tight characterization of minimum revenue and maximum bidder surplus for the unknown values model. At the end of the section, we will apply our methods to the known values model, which turns out to be significantly more complicated. We will give a tight characterization for some cases and explain the limitations of our techniques.

4.1 Unknown Values

4.1.1 Main Idea

The analysis of minimum revenue and maximum bidder surplus requires the development of a number of new ideas. Before diving in, we will provide some intuition for where we are headed using our familiar example with two bidders and independent standard uniform valuations. We will construct an equilibrium in that setting with extremely unfavorable revenue properties, but which generates very high surplus for the bidders. As with the analysis of maximum revenue, our formal results will be for discrete models with finitely many values. In the continuum limit, our constructions converge to the equilibrium that we now describe.

In this equilibrium, the winner will be the bidder with the highest valuation, and, similar to the complete information Bertrand equilibrium, the winning bid will be a deterministic and strictly increasing function of the *loser's* value. We will call this winning bid function $\beta(v)$. Note that revenue can be calculated from the bidding function by integrating over losing values:

$$R = \int_{v=0}^{1} \beta(v) 2(1-v) dv.$$

For now, we take this structure granted, and we look for a lower bound on revenue across this class of equilibria. The approach, which will be fully developed over the course of this section, is to minimize revenue over all possible winning bid functions.

Since bidding behavior is only partially specified by β, we cannot evaluate all possible deviations and verify that β is consistent with an equilibrium. For the purposes of this exercise, however, we will only require that β deter a subset of deviations: for all $v \in [0,1]$, bidders should not want to deviate by bidding $\beta(v)$ whenever they would have bid some $x \leq \beta(v)$ in equilibrium. We refer to this as a *uniform deviation up to* $\beta(v)$ (uniform in the sense of deviating to $\max\{x, \beta\}$ for all equilibrium bids x). Notice that this is a deviation in the strategic normal form of the game, since it involves deviating after all signals which induce bids less than $\beta(v)$ in equilibrium, and we will evaluate its merits in ex-ante terms.

Fortunately, β does contain enough information to evaluate uniform upward deviations. Such a deviation results in two distinct effects on the deviator's surplus. First, the deviator gains some surplus by winning on events when she would have lost the auction in equilibrium, and the other bidder would have won. This gain in surplus is therefore

$$\int_{x=0}^{v} (x - \beta(v))(1 - x)dx.$$

On the other hand, the deviator would sometimes have won with bids less than $x \leq \beta(v)$ in equilibrium. After the deviation, she will still win, but now the deviator will have to pay more for the good. This loss is

$$\int_{x=0}^{v} (\beta(v) - \beta(x))(1 - x)dx.$$

The uniform deviation up to $\beta(v)$ is not attractive if the gain is less than the loss, or if

$$\begin{aligned} \int_{x=0}^{v} (x - \beta(v))(1 - x)dx &\leq \int_{x=0}^{v} (\beta(v) - \beta(x))(1 - x)dx \\ \iff \beta(v) &\geq \frac{2}{1 - (1 - v)^2} \int_{x=0}^{v} (x + \beta(x))(1 - x)dx. \end{aligned} \tag{7}$$

Now consider a candidate β for which (7) is slack for some v. This means that it is possible to have the winner bid less when the loser's value is v, without giving bidders an incentive to uniformly deviate up to $\beta(v)$. Moreover, inspection of the right-hand side of (7) indicates that lowering $\beta(v)$ further relaxes the constraint for higher v, by decreasing the integral. This indicates that at an optimum, the incentive constraint (7) should hold as an equality for all $v \in [0, 1]$.

Equivalently, we can differentiate (7) to conclude that the optimal β should satisfy the following differential equation:

$$\beta'(v) = (v - \beta(v))\frac{1 - x}{1 - (1 - v)^2}.$$

This equation has a unique solution given by

$$\beta(v) = \frac{1}{\sqrt{1 - (1 - v)^2}} \int_{x=0}^{v} \frac{x(1 - x)}{\sqrt{1 - (1 - x)^2}} dx \tag{8}$$

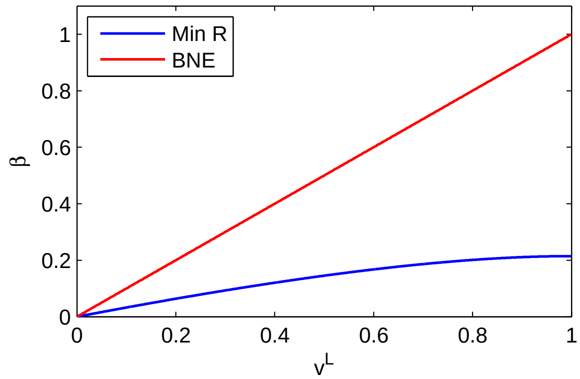

Figure 2: Winning bid as a function of the minimum value in the no information minimum revenue BCE and the complete information BNE, with two bidders whose values are i.i.d.standard uniform.

which is in fact the optimal bidding function. This β is depicted in Figure 2, with comparison to the known values BNE bidding function, which entails each buyer bidding half of their own value (rather than a function of the losing buyer's value).

We can then compute the revenue to be

$$R = \frac{1}{3} - 2\left[\int_{x=0}^{1} \sqrt{1-(1-v)^2}dv - \int_{v=0}^{1}(1-(1-v)^2)dv\right] \approx 0.0959,$$

which is significantly lower than the 1/3 revenue obtained in the known values BNE. Moreover, the allocation is efficient, so total surplus is 2/3. The bidders must obtain all of the surplus net of revenue, which is about 0.5708.

At this point, we have constructed part of a bidding strategy and verified that it is consistent with a subset of the incentive constraints that have to be satisfied in equilibrium. We shall subsequently see that it is possible to fill in the missing pieces, in particular the distribution of losing bids, in order to extend this construction to a full equilibrium. Moreover, not only will this equilibrium minimize revenue when the winner's bid is a deterministic function of the loser's value, but it will globally minimize revenue. Finally, in this equilibrium, it will be the case that bidders are indifferent to *all* deviations in which they bid $\beta(v)$ instead of some $x < \beta(v)$.

The fact that the uniform upward incentive constraints are binding is closely related to how the equilibrium is able to induce such low revenue. Minimizing revenue entails pushing down the distribution of winning bids as far as it will go. Intuitively, the force that prevents the bid distribution from being pushed down further is the bidders' temptation to deviate upwards: if too many winning bids were packed close together at the low end, increasing one's bid would result in a too large of increase in the probability of winning relative to the additional payment. This suggests that upward constraints should be binding at the minimum revenue BCE. Indeed, in our construction, *all* upward incentive constraints are binding, as long as the deviation is to a bid in $[0, \beta(1)]$.

At first glance, it may seem remarkable that this construction pushes down revenue to such a degree while still maintaining efficiency of the allocation. One might have thought that a way to lower revenue would be to destroy some surplus by misallocating the good, thereby leaving fewer rents for both the buyers and the seller. In fact, the opposite is true: efficiency of the allocation is *necessary* to minimize revenue. The reason is that the attractiveness of an upward deviation depends on what the bidder believes is the valuation conditional on losing. By making the allocation efficient, the expected valuation conditional on losing the auction is minimized, thus giving bidders weaker incentives to deviate up.

We will see that similar equilibria can be constructed in general. Our analysis will work with discrete distributions of values, but the equilibria retain several key features of the uniform example:

1. The allocation is efficient.

2. The winning bid recommendation only depends on the loser's value.

3. Bidders with higher valuations lose to higher bids.

For the uniform example, the last property manifests itself in the monotonicity of the function β. A prominent difference between the discrete and continuous models is that in the continuum limit, the winner's bid is a deterministic function of the loser's value, whereas the discrete model will involve some randomization over winning bids conditional on the loser's value. Nonetheless, we shall see that there is an "ordered supports" property that characterizes the discrete solution, so that the winner's bid is drawn from a higher support when the loser's value is higher. After the derivation, we will discuss how our solution behaves in the continuum limit.

4.1.2 The Relaxed Program

The first step in our general characterization of minimum revenue for unknown values is to establish a bound on how low revenue can go. To economize on notation while we develop ideas, we will first focus on the case with two bidders and symmetric BCE. The analysis will subsequently be generalized to many bidders and asymmetric equilibria.

As we have foreshadowed, the constraints that characterize minimum revenue are those corresponding to upward deviations. Thus, at a first cut, we will drop all incentive constraints except for those corresponding to deviating from a bid x to a higher bid b. This still leaves a two-dimensional family of constraints indexed by recommendations and upward deviations. We will reduce the problem even further by only checking incentive compatibility with regard to the one-dimensional family of uniform upward deviations that we looked at with the uniform example, namely, deviations in which a buyer deviates from the equilibrium bid x to $\max\{x, b\}$, for some fixed b.

In addition, we will simplify the problem by only looking at certain marginal distributions of a BCE. Recall that the BCE specifies a joint distribution of bids for every profile of valuations. Revenue, however, can be calculated from the distribution of winning bids alone. Let

$$H(b|v_1, v_2) = \int_{\{(x_1,x_2)\in\mathbb{R}_+^2 | x_i \le b\}} q_1(x_1, x_2) F(dx_1, dx_2 | v_1, v_2) \tag{9}$$

denote the probability that buyer 1 wins with a bid less than b, when the profile of values is (v_1, v_2). Since we are assuming symmetry of the equilibrium, this quantity is independent of the identities of the bidders, and we can make the notation generic by writing $H(b|v, v')$ for the probability that a type v wins against a type v' with a bid less than or equal to b. Note that these probabilities condition on the event that the profile of values is (v, v'), but not on events where v or v' is the winner. Thus, $H(b|v, v') + H(b|v', v)$ is the probability that either bidder wins with a bid less than or equal to b, conditional on the profile of values being (v, v'). Feasibility requires that

$$H(b|v, v') + H(b|v', v) \in [0, 1], \tag{10}$$

for all $b \in \mathbb{R}_+$ and for all $(v, v') \in V^2$.

Using these winning bid distributions, we can compute revenue as the expected winning bid:

$$R = \sum_{(v,v')\in V^2} p(v, v') \int_{b=0}^{v^K} b H(db|v, v') = v^K - \sum_{(v,v')\in V^2} p(v, v') \int_{b=0}^{v^K} H(b|v, v') db$$

where the equivalence comes from integration by parts. Thus, minimizing revenue is aligned with maximizing

$$\sum_{(v,v')\in V^2} p(v, v') \int_{b=0}^{v^K} H(b|v, v') db, \tag{11}$$

which is the area under the winning bid distributions.

In addition to calculating revenue, we can use these objects to evaluate uniform upward deviations. As with the uniform example, we can decompose the effect on bidder utility into a gain and a loss. The buyer loses surplus of

$$\sum_{(v,v')\in V^2} p(v,v')\int_{x=0}^{b}(b-x)H(dx|v,v'),$$

since, by deviating to a higher bid, the buyer will still win on events where he would have won by following the equilibrium strategy, but he will pay more for the good. This quantity integrates by parts to

$$\sum_{(v,v')\in V^2} p(v,v')\int_{x=0}^{b} H(x|v,v')dx,$$

and notice the similarity with the objective of maximizing (11). The gain associated with the deviation comes from winning when the buyer would have lost by following the equilibrium strategy. These are precisely the outcomes where the *other* buyer would have won with a bid less than or equal to b:

$$\sum_{(v,v')\in V^2} p(v,v')(v'-b)H(b|v,v').$$

A uniform deviation up to b is not attractive if the loss exceeds the gain, or if

$$\sum_{(v,v')\in V^2} p(v,v')(v'-b)H(b|v,v') \leq \sum_{(v,v')\in V^2} p(v,v')\int_{x=0}^{b} H(x|v,v')dx. \tag{12}$$

Thus, our relaxed program is to maximize (11), subject to (10) and (12) for all $b\in\mathbb{R}_+$. Note that we are even dropping the requirement that the distributions $H(b|v,v')$ be monotonic, though fortunately the solution to this relaxed formulation will turn out to be weakly increasing. Because this program is a relaxation of the original problem of minimizing revenue over all BCE, the solution must generate a lower bound for revenue across all unknown value BCE.

The relaxed program can be visualized as an optimal control problem. Bids b take the role of "time", which runs from $b=0$ up to $b=v^K$. The winning bid distributions $H(b|v,v')$ are the "controls" and are constrained by (12). The right-hand side of that constraint is a state variable,

$$S(b|v,v') = \int_{x=0}^{b} H(x|v,v')dx,$$

whose law of motion is the current value of the control, i.e., $\dot{S}(b|v,v') = H(b|v,v')$. All of the controls and states are initialized to zero at $b = 0$, and the revenue minimizing winning bid distributions will maximize the sum of the terminal values of the states, which is $\sum_{v,v'} p(v,v')S(v^K|v,v')$.

We now describe the solution to the relaxed program. First, at an optimal solution, it must be that the incentive constraint (12) almost surely holds as an equality whenever $H(b|v,v')+H(b|v',v) < 1$ for some (v,v'). If this were not the case, then it would be possible to increase one of these controls without violating (12) at b. Moreover, increasing $H(b|v,v')$ merely relaxes the right-hand side of (12) even further (and increases the objective), so we know that the increased control will still be feasible (and lowers revenue).

Second, the allocation of the good that is implied by the winning bid distributions must be efficient, i.e., $H(b|v,v') = 0$ if $v < v'$. This can be seen by inspecting the left-hand side of (12). Notice that the control $H(b|v,v')$ is multiplied by $v' - b$, i.e., the valuation of the *losing* buyer. This corresponds to our intuition from the uniform example: the gain from deviating upwards depends on one's valuation on the event that one loses. All things equal, this gain is smaller if losing valuations are lower, which is when the allocation is efficient.

More formally, suppose that $H(b|v',v) > 0$ even though $v' < v$. Then we can define a new solution $\tilde{H}$, which coincides with the old solution except for this pair of values, and with

$$\begin{aligned}
\tilde{H}(b|v',v) &= 0\\
\tilde{H}(b|v,v') &= H(b|v,v') + H(b|v',v).
\end{aligned}$$

It is clear that this new solution will respect feasibility and result in the same right-hand side in (12), as well as the same objective (11). However, the left-hand side of (12) will have strictly decreased, because

$$\begin{aligned}
\sum_{(v,v')\in V^2} p(v,v')(v'-b)\tilde{H}(b|v,v') &= \sum_{(v,v')\in V^2} p(v,v')(v'-b)H(b|v,v') + (v'-v)H(b|v',v)\\
&< \sum_{(v,v')\in V^2} p(v,v')(v'-b)H(b|v,v'),
\end{aligned}$$

since $v' < v$. Thus, there is now some extra slack in (12) that could be used to increase the controls faster, thereby reducing pushing down the winning bid distributions and decreasing revenue.

At this point, we notice that the winning bid does not appear anywhere in either the objective or the constraints, except as an index for the winning bid distributions. It is only the loser's value that directly enters the program, through the incentive constraint. This suggests that we can collapse the program down further. In particular, let us write

$$p^L(v) = \frac{1}{2}p(v,v) + \sum_{v'>v} p(v,v')$$

for the probability that v is the losing value, and let

$$H(b|v) = \frac{1}{p^L(v)} \sum_{v' \geq v} p(v,v')H(b|v',v)$$

denote the distribution of the winning bids made against a losing bidder with valuation v. We can then rewrite the relaxed program as

$$\max \sum_{v \in V} p^L(v) \int_{b=0}^{v^K} H(b|v)db; \tag{13}$$

subject to

$$H(b|v) \in [0,1]; \tag{14}$$

$$\sum_{v \in V} p^L(v)(v-b)H(b|v) \leq \sum_{v \in V} p^L(v) \int_{x=0}^{b} H(x|v)dx. \tag{15}$$

Our earlier conclusion still applies: that (15) should bind whenever $H(b|v) < 1$ for some v.

At this point, we know that (15) has to bind throughout the support of winning bids, and moreover, that support is connected, i.e., an interval. Note that at a given b, there is a certain amount of slack on the right-hand side of (15), and a number of controls on the left-hand side that could use that slack. The question then is: which controls should be used at each b? It turns out that there is a simple answer, which is to use the $H(b|v)$ with the smallest v such that $H(b|v) < 1$. In particular, the optimal solution should satisfy the *ordered supports* property:

$$H(b|v) < 1 \implies H(b|v') = 0 \, \forall v' > v. \tag{16}$$

The logic here is closely related to the reason why the solution should be efficient. It is clear that all $H(b|v)$ contribute symmetrically to the objective and to the right-hand side of (15), but $H(b|v)$ for smaller v contribute less to the left-hand side of (15). Thus, all things equal, it is more efficient to first use $H(b|v)$ with lower v to relax the incentive constraint faster, which subsequently provides more room for controls at higher b.

The final piece to characterizing the solution is an initial condition. We argue in the Appendix that $H(b|v)$ must be zero for $b < v^1$. Intuitively, the lowest types of buyer are in Bertrand competition with one another for the good, and they bid up the price to v^1. In addition, it is clear that (15) does not constrain $H(b|v^1)$ at all, since the coefficient $v^1 - b$ is non-positive for all $b \geq v^1$. Thus, this control can immediately jump up to 1 at $b = v^1$.

Given this initial condition, the rest of the solution is constructed inductively using (16). Suppose that $H(b|v^l) = 1$ for all $l < k$ and $H(b|v^l) = 0$ for $l \geq k$, and (15) is binding at b. We then solve (15) with equality for $H(b|v^k)$, setting $H(b|v^l) = 1$ for $l < k$ and $H(b|v^l) = 0$ for $l > k$. The $H(b|v^k)$ that solves this equation will blow up as b gets close to v^k, since $v^k - b$ converges to zero. Thus, there will be a finite b^k at which $H(b|v^k)$ hits 1. At this point, we set $H(b|v^k) = 1$ for all $b > b^k$, and we continue the process with $H(b|v^{k+1})$. At $k = K$, the algorithm terminates and all of the winning bid distributions have been specified.

This completes our characterization of the two-bidder unknown values relaxed program. The solution is efficient, has binding upward incentive constraints in the support of winning bids, and satisfies the ordered supports property. Moreover, the analysis generalizes in a natural manner to the case of many bidders, but there are two key differences.

First, in a symmetric equilibrium with two bidders, conditional on a profile of winner's and loser's valuations, each buyer is equally likely in ex-ante terms to be the buyer who wins the auction. With N bidders, it must be that each bidder only wins a fraction $1/N$ of the time and loses a fraction $(N-1)/N$ of the time. Thus, the many bidder analogue of the incentive constraint (15) has a weight of $N-1$ on the left-hand side, since an upward deviator is more likely to be a losing buyer and therefore gains more from winning when she was supposed to lose.

Second, in the final form of the two bidder relaxed program, the choice variables are distributions of winning bids conditional on the loser's value. In some sense, we needed to keep track of the statistical relationship between winning bids and losing values in order to calculate the gains from deviating upwards. With $N > 2$, there are multiple losing bidders, each of whom has a different valuation, so in principle we might have to keep track of how winning bids are related to all of the losing values. It turns out, however, that we can just keep track of how winning bids are related to summary statistics of the losing values. In particular, suppose that some bidder deviates up and wins when they were supposed to lose when the true profile of values is $v = (v_1, \ldots, v_N)$. By symmetry, they must also now win whenever the profile is a permutation of v, $\xi(v)$ for $\xi \in \Xi$. In the event that they win when they would have lost in equilibrium, the upward deviator is therefore equally likely to have any losing value in v. Thus, the expected valuation on this event is the *average* losing value, given the profile v.

As a result, the generalized relaxed program will keep track of the distribution of winning bids conditional on an average losing value m, which we denote by $H(b|m)$. Specifically, let

$$\mu(v) = \frac{1}{N-1}\left(\sum_i v_i - \max v\right) \tag{17}$$

denote the average losing value when the profile of values is v. This statistic takes on values in a finite set $M \subset \mathbb{R}_+$, and is distributed according to

$$p^L(m) = \sum_{\{v \in V^N | \mu(v) = m\}} p\left(v\right). \tag{18}$$

The winning bid distributions over which we maximize are

$$H(b|m) = \frac{1}{p^L(m)} \sum_{\{v \in V^N | v_1 = \max v, \mu(v) = m\}} p(v) |\arg\max v| H(b|v_1, v_{-1}).$$

The general form of the relaxed program is to maximize

$$\sum_{m \in M} p^L(m) \int_{b=0}^{v^K} H(b|m) db \tag{19}$$

subject to

$$H(b|m) \in [0,1] \forall m \in M \tag{20}$$

$$(N-1) \sum_{m \in M} p^L(m)(m-b) H(b|m) \leq \sum_{m \in M} p^L(m) \int_{x=0}^{b} H(x|m) dx. \tag{21}$$

The solution to this relaxed program generates a lower bound on revenue:

Lemma 4. *The $H(b|m)$ induced by any symmetric BCE must satisfy* (20) *and* (21)*. Moreover, expected revenue under F must be at least v^K minus* (19)*.*

The solution to the generalized relaxed program has the same structure as with two bidders, as described by the following proposition:

Proposition 5. *The solution to the unknown values relaxed problem is the unique* $\{H(\cdot|m)\}_{m\in M}$ *such that* (21) *holds with equality for* $b > \min M$ *whenever* $H(b|m) < 1$ *for some* $m \in M$ *and that satisfies the ordered supports property:*

$$H(b|m) < 1 \implies H(b|m') = 0 \, \forall m' > m. \tag{22}$$

To recap, we have studied a particular relaxed program for minimizing revenue. In this program, we dropped all incentive constraints except those associated with deviations of the form "bid b whenever you would normally bid any $x \leq b$," and we optimized over winning bid distributions rather than the entire joint distribution of bids. The resulting solution implies a very particular structure for winning bid distributions that will minimize revenue. We note that this lower bound on revenue also generates an upper bound on bidder surplus, since bidders cannot receive more than the efficient surplus minus minimum revenue.

4.1.3 Construction of a BCE

At this point, Proposition 5 gives us a lower bound on minimum revenue and an upper bound on bidder surplus, though we do not yet know if these bounds are tight. There is good reason to think they might not be: we have only specified a particular marginal distribution of a BCE, namely the distribution of winning bids, and we have also not checked if the myriad of other constraints can be satisfied.

It turns out, however, that the solution to the relaxed program can always be extended to a BCE, thus verifying sharpness of the bounds derived above. Let $\{H(b|m)\}$ be the solution described in Proposition 5. Our procedure for extending this solution to a full BCE is as follows:

(i) Draw a profile of values v according to $p(v)$, and pick a bidder with a high value to be the winner, breaking ties uniformly;

(ii) Assign the winner a bid b drawn from the distribution $H(b|\mu(v))$;

(iii) Draw bids b for the remaining bidders independently from the cumulative distribution $L(b'|b, \mu(v))$ (to be specified shortly).

The distribution of losing values contains quite a bit more information than just the winning bid distributions with which we have worked thus far. In particular, $L(b'|b, m)$ can be used to calculate the distribution of the winner's bid conditional on a losing recommendation b'. Whereas before we could only evaluate uniform upward deviations, now we will be able to evaluate *pointwise upward deviations*, in which a bidder deviates from a particular bid $x < b$

up to b. Note that the net gains from a uniform deviation up to b is simply the expectation of the net gains from pointwise deviations from x up to b for all $x \leq b$.

In particular, consider a bidder who is told to bid b and deviates up to $b' > b$. The gain from winning when one would have lost is

$$\sum_{m' \in M} (m' - b') p^L(m') \frac{N-1}{N} \int_{x=0}^{b'} L(db|x, m') H(dx|m') \tag{23}$$

and the loss from paying more when the deviator would have won anyway is

$$(b' - b) \sum_{m' \in M} \frac{p^L(m')}{N} H(db|m'). \tag{24}$$

Thus, the pointwise upward incentive constraint is that

$$(N-1) \sum_{m' \in M} (m' - b') p^L(m') \int_{x=0}^{b'} L(db|x, m') H(dx|m') - (b' - b) \sum_{m' \in M} p^L(m') H(db|m') \leq 0. \tag{25}$$

This inequality must hold almost surely if these losing and winning bid distributions constitute a BCE.

In fact, we claim that if the L's and H's constitute a BCE, the incentive constraint (25) must hold as an equality for all $b' > b$. In other words, not only are bidders indifferent to uniform upward deviations, but they are also indifferent to almost all pointwise upward deviations. Why? Suppose there were a set X of bids less than b which arose with positive probability in equilibrium and for which the bidders strictly preferred their equilibrium bids over deviating up to b. Since bidders are indifferent to the uniform upward deviation up to b, it must be that bidders strictly prefer to deviate to b from bids $[0, b] \setminus X$! Of course, that cannot happen in a BCE.

Moreover, it turns out that there is a unique choice of $L(b|b', m)$ such that all pointwise upward incentive constraints hold with equality. Let B_m denote the support of $H(b|m)$. These supports are ordered so that if $b \in B_m$ and $b' \in B_{m'}$ with $m > m'$, then $b \geq b'$. Assuming that $L(x|y, m')$ has been defined for $m' < m$ and for $x \leq y$, with $L(x|x, m') = 1$, we inductively define

$$p^L(m)L(b|b',m)H(db'|m)$$
$$= \frac{1}{(m-b')^2}\left[\frac{1}{N-1}\sum_{m'\leq\tilde{m}} p^L(m')\left(\int_{x=0}^{b} H(x|m')dx + (m-b)H(b|m')\right)\right. \tag{26}$$
$$\left. + \sum_{m'<m}(m-m')p^L(m')\int_{x\in B_{m'}} L(b|x,m')H(dx|m')\right].$$

where $b \in B_{\tilde{m}}$ with $\tilde{m} \leq m$. $L(b|b',m)$ is identically 1 for $b \in B_{\tilde{m}}$ with $\tilde{m} > m$. The following result characterizes these losing bid distributions:

Lemma 6.

(i) The functions $L(b|b',m)$ defined by (26) are monotonically increasing in b, and satisfy $L(b|b,m) = 1$.

(ii) Moreover, if the marginal distribution of losing bids is $L(b|b',m)$, then bidders are almost surely indifferent between following a recommendation b and deviating upwards to any $b' > b$.

(iii) Finally, if losing bids are i.i.d.draws from $L(b|b',m)$, then bidders almost surely prefer the recommendation b to any $b' < b$.

The proof of this result appears in the Appendix. The proofs of (i) and (ii) are somewhat mechanical, with the latter essentially verifying that (26) is reverse-engineered from the assumption that the pointwise upward incentive constraints all hold with equality. The proof of (iii) is more subtle and, we are sorry to say, a bit mysterious. In the next section we will study the limit of this equilibrium as the distribution of values converges to a continuous distribution. The equilibrium simplifies dramatically in the limit, and we will use that simpler construction to provide better intuition for the result, or at the least a more transparent argument for why incentive constraints should be satisfied.

There is one last step before we state our main result. Through the discussion of minimum revenue, we have restricted attention to symmetric BCE, under the premise that this is without loss of generality. We now state this as a formal result.

Lemma 7. *If the distribution of values p is symmetric, then there exist symmetric BCE that minimize revenue and maximize total bidder surplus.*

Here is the basic idea behind this Lemma. BCE are a subset of the convex set of joint distributions over values and bids. For that reason, if one had an asymmetric BCE, it is always possible to "symmetrize" by (i) first drawing a permutation of the buyers' identities at random and (ii) drawing values and bids according to the original BCE and assigning them to the permuted identities. Note that this symmetrized BCE has the same distribution of values, since the original distribution was symmetric. Moreover, conditional on each permutation, the conditional distribution of bids and values is a BCE, so overall it must be a BCE as well. And finally, revenue and total bidder surplus for the symmetrized BCE are the expectation of those objectives conditional on the permutation, but conditional on the permutation, revenue and total bidder surplus are the same as under the original BCE.

Thus, the bounds from Proposition 5 are sharp and in particular must also be satisfied by asymmetric BCE. Since our constructed BCE attains the bounds, we have a complete welfare characterization:

Theorem 8 (Min revenue and max bidder surplus for unknown values)**.** *The solution to the unknown values relaxed problem, together with the losing bid distributions defined by* (26)*, constitute a BCE. This BCE simultaneously maximizes total bidder surplus and minimizes revenue over all unknown values BCE.*

4.1.4 Continuum Limit

The results from the previous section give a tight lower bound for revenue for any symmetric prior with finite support. The finiteness was convenient for our characterization of the relaxed program, but the discreteness of valuations necessitated that the minimum revenue BCE involve a lot of randomization of bids, in particular randomization of the winning bid recommendation given the losing value. This randomization is not entirely unexpected, given that similar behavior arises in prior work characterizing equilibria of first price auctions with discrete values, e.g., that of Maskin and Riley (1985) or Fang and Morris (2006). It should therefore also not surprise the reader that the analogous construction is substantially simpler when the distribution of values is continuous. In this section, we will explain what this limit looks like, with two objectives in mind: first, the limit equilibrium is extremely easy to state and verify, which to some extent may alleviate the mystery of why our discrete construction hangs together; and second, we can use the continuous construction to obtain deeper insight into the structure of information that attains minimum revenue, and also relate our results to the prior literature on first price auctions.

Before constructing our equilibrium, let us first consider how the solution to the relaxed program should look in the limit. This discussion is heuristic, and is not meant to be a formal derivation. Let us suppose there is a sequence $p^k(v)$ of discrete joint distributions of values that converge to some continuous weak-* limit $p(v)$ with support equal to the compact cube $[\underline{v}, \overline{v}]^N$. Let us write $p^L(m)$ for the distribution of the average losing value for the continuous limiting distribution. For example, in the case of two bidders with independent and identically distributed values, i.e., $p(v_1, v_2) = p(v_1)p(v_2)$, the average losing value is just the second-highest value, $m = \min(v_1, v_2)$, and

$$p^L(m) = 2p(m)(1 - p(m)).$$

We assume that p^L has compact support equal to $V = [\underline{v}, \overline{v}]$, and we write $P^L(m)$ for the corresponding cumulative distribution.

In the limit as the mass of any particular value of m goes to zero, the supports of the $H(v|m)$ will collapse to a singleton, and there will be a deterministic winning bid as a function of the average loser's value, which we denote by $\beta(m)$. Assume for now that this function is strictly increasing and differentiable. Then the continuum analogue of (21) is

$$(N-1)\int_{x=\underline{v}}^{\beta^{-1}(b)} (x - b)p^L(x)dx = \int_{x=\underline{v}}^{\beta^{-1}(b)} (b - \beta(x))p^L(x)dx.$$

Again, the left-hand side is the gain from winning when the average losing value ranges from $\underline{v}$ up to $\beta^{-1}(b)$, and the right-hand side is the loss from paying more when the average losing value is below $\beta^{-1}(b)$. We can substitute $b = \beta(m)$ into the limit of integration to rewrite this as

$$(N-1)\int_{x=\underline{v}}^{m} (x - \beta(m))p^L(x)dx = \int_{x=\underline{v}}^{m} (\beta(m) - \beta(x))p^L(x)dx. \tag{27}$$

This relationship must hold for all $m \in V$.

Thus, (27) pins down the winning bid function as the solution of a differential equation. In fact, there is a unique β that satisfies (27), which is given by

$$\beta(m) = \frac{N-1}{N}\frac{1}{(P^L(m))^{\frac{N-1}{N}}}\int_{x=\underline{v}}^{m} \frac{xp^L(x)}{(P^L(x))^{\frac{1}{N}}}dx. \tag{28}$$

This function is the continuum analogue of the $H(b|m)$ distributions from the discrete construction.

As in the discrete case, we can always construct an information structure and BNE in which the winner bids according to (28). Bidders will receive signals in $S_i = [\underline{v}, \overline{v}]$: if $v_i = \max v$, then bidder i receives the signal $s = m = \mu(v)$. Otherwise, bidders receive signals with independent draws from the conditional distribution $L(s|m)$ on $[\underline{v}, m]$:

$$L(s|m) = \left(\frac{P^L(s)}{P^L(m)}\right)^{\frac{1}{N}}. \tag{29}$$

This is in some sense the continuum analogue of (26). The total probability of getting the signal s, in ex-ante terms, is

$$\frac{p^L(s)}{N} + \frac{N-1}{N}\int_{x=s}^{\overline{v}} L(ds|m)p^L(m)dm = \frac{1}{N}\frac{p^L(s)}{(P^L(s))^{\frac{N-1}{N}}}.$$

But conditional on receiving the signal s, bidders think that s is the average losing value (and they have the high value, with others' signals being lower) with probability $(P^L(s))^{\frac{N-1}{N}}$, and with probability $1 - (P^L(s))^{\frac{N-1}{N}}$, they think that the true average losing value is higher than s, and is distributed according to the conditional distribution

$$\frac{(P^L(m))^{\frac{N-1}{N}} - (P^L(s))^{\frac{N-1}{N}}}{1 - (P^L(s))^{\frac{N-1}{N}}}.$$

In equilibrium, all bidders follow the monotonic pure strategy of bidding $\beta(s)$. Thus, in the event that the true m is greater than s, bidders believe that the highest bid is $\beta(m)$. Through integration by parts, we can deduce that the general form for minimum revenue is

$$\begin{aligned} R &= \int_{m=\underline{v}}^{\overline{v}} \beta(m)p^L(m)dm \\ &= \overline{v} + \int_{m=\underline{v}}^{\overline{v}} \left((N-1)P^L(m) - N(P^L(m))^{\frac{N-1}{N}}\right) dm. \end{aligned}$$

The following proposition asserts that these objects do in fact constitute an equilibrium:

Proposition 9 (Continuum limit)**.** *Suppose that the joint distribution has a density $p(v)$ with full support on $[\underline{v}, \overline{v}]^n$. Then there exists a BCE in which the buyer with the high value bids $\beta(\mu(v))$, where β is defined by* (28) *and is the solution to* (27)*, and the buyers with lower values bid $\beta(s)$, with the losing bidder's signals being independent draws from $L(s|\mu(v))$ defined by* (29)*.*

Thus, in the minimum revenue information structure, bidders receive one-dimensional signals that are informative about the average losing value m. Essentially, the signal could fall into one of two categories: one of the bidders, chosen at random, will receive a "precise" signal $s = m$, while the remaining $N-1$ bidders will receive "coarse" signals that are distributed between $[0, m]$ according to the losing signal distribution $L(s|m)$. Importantly, bidders are uncertain about whether they received the precise signal or the coarse signal. Moreover, this information structure supports a symmetric, monotonic, and pure strategy equilibrium in which everyone bids $\beta(s)$.

The reader may be curious about the connection to Milgrom and Weber (1982), who study a class of symmetric "affiliated" information structures that similarly admit a monotonic and pure strategy equilibrium. The affiliation property can be thought of as a form of positive correlation between the bidders' values and their one-dimensional signals. While affiliation is sufficient to support monotonic pure strategies, it is far from necessary, and indeed, regardless of the underlying distribution of values, the minimum revenue information structure will not be affiliated. This can be seen by considering the conditional distribution of signals given m and one of the bidder's signals s_i: for $s_i < m$, it must be that s_i was the coarse signal, so that some other bidder obtained a recommendation $s_j = m$, whereas if $s_i = m$, then others must have almost surely received coarse signals which are strictly below m. Thus, signals are "negatively correlated" given the true value of m.

Indeed, affiliation restricts the kind of inference that bidders can draw about their values upon losing the auction in a manner that strengthens the incentive to deviate upwards. In the event that some other bidder has won the auction, they must have received a higher signal, which means that losing the auction is a *positive* signal about one's own value. In turn, this means that the expected value conditional on winning is increasing as a bidder deviates upwards. On the other hand, winning when one should have lost in the minimum revenue information structure and equilibrium is, at the margin, a strong *negative* signal about one's value. For if a bidder wins with a bid of $\beta(s)$ after receiving the signal s, they believe that their value is distributed over the range $[s, \overline{v}]$. But when they deviate up to a bid $\beta(s+\epsilon)$, the incremental event on which they win is when their value was between s and $s+\epsilon$, so that this event pulls down the expectation of their value conditional on winning.

To conclude, let us apply these results to the uniform example with which we began the section. The formula for $\beta(s)$ reduces to what we computed for the uniform example at the beginning of the section, with $P^L(v) = 1-(1-v)^2$. The conditional distribution of the loser's signal, given the loser's value is v, is

$$L(s|v) = \sqrt{\frac{1-(1-s)^2}{1-(1-v)^2}}. \tag{30}$$

In the example, we first draw values (v_1, v_2) independently from the standard uniform distribution, with the lower of the two values being y. The bidder with the higher value is told to bid $\beta(s)$, where β is given by (8). For the bidder with the lower value, we draw an x according to (30) and tell the losing buyer to bid $\beta(x)$. As we have shown, this recommendation rule will satisfy both upward and downward incentive constraints and is therefore a BCE.

4.2 Known Values

We now turn our attention to minimum revenue and maximum bidder surplus when the buyers know their own valuations. We will pursue the same course that characterized revenue in the unknown values model, namely, (i) formulate and solve a relaxed program for revenue, and then (ii) extend this solution to a full BCE. In the known values relaxed program, we again only keep track of winning bid distributions and drop all incentive constraints except those corresponding to uniform upward deviations. In the unknown values case, there was a one-dimensional family of such constraints, indexed by the bid that the player deviates up to. With known values, however, there is a separate family of such constraints for each possible known value. This introduces a second-dimension to the problem, and for general known values models, the pattern of binding uniform upward incentive constraints can be quite complicated. We will, however, give a tight characterization of minimum revenue and maximum bidder surplus when there are only two possible valuations, high or low. We give examples of the complexities that arise with more than two values in Appendix C

Suppose that $K = 2$ and $V = \{v^1, v^2\}$ is the set of possible valuations. Bidders know whether or not they have the low-value or the high value. It turns out, however, that the strategic behavior of the low-valuation buyer is quite simple: When all buyers have low values, they will all bid the low value and tie. Otherwise, these buyers have to lose the auction in equilibrium, and they randomize so as to induce a high valuation buyer to win at a price weakly greater than v^1. Thus, we only need to determine the behavior of the high valuation buyer.

To characterize the high type's behavior, we will first study a relaxed program. It will turn out that in the revenue minimizing BCE, bidders will have somewhat complicated higher order beliefs about the number of high types. For example, bidders will sometimes make low bids because they think there is a high probability that they are the only bidder with a high value, and they are only facing low types. At other times, the high type will believe that they are likely to be bidding against another high type, who thinks that *they* are most likely facing only low types. Of course, this logic continues to higher orders, and in general, a high type may think that they are likely to be facing (and winning against) $k - 1$ other high types, who themselves think that they are facing $k - 2$ other high types, etc. For this reason, in formulating our relaxed program, it is necessary to keep track of the statistical relationship between winning bids and the number of bidders with high valuations.

For a profile of values v, we let

$$\zeta(v) = |\{i|v_i = v^2\}|$$

denote the number of high value buyers. The distribution p induces a distribution on $\{0, \dots, N\}$ given by

$$p^Z(z) = p(\zeta^{-1}(z)).$$

Let $H(b|z)$ denote the cumulative distribution of winning bids when there are z bidders with value v^2. Per the previous discussion, $H(\cdot|0)$ puts probability one on v^1.

The relaxed program is

$$\max \sum_{z=0}^{N} p^Z(z) \int_{b=0}^{v^2} H(b|z)db \tag{31}$$

subject to

$$H(b|z) \in [0,1] \text{ for all } b \in \mathbb{R}_+, z \in \{0, \dots, N\}; \tag{32}$$

$$(v^2 - b) \sum_{z=1}^{N} p^Z(z) \frac{z-1}{z} H(b|z) \leq \sum_{z=1}^{N} p^Z(z) \frac{1}{z} \int_{x=0}^{b} H(x|z)dx. \tag{33}$$

These constraints are almost entirely analogous to the feasibility and incentive constraints from the unknown values relaxed problem, with the exception that there is now only an incentive constraint for the high valuation type (who has a known value v^2). Note that when there are z high types, a bidder with a high valuation thinks that there is a $1/z$ chance that they are the winner and a $(z-1)/z$ chance that they will not win.

The analysis of the binary known values relaxed program proceeds in much the same manner as for unknown values. It must be that (33) binds whenever $H(b|z) < 1$ for some z, since otherwise we could push down the distribution of winning bids while further relaxing incentive constraints. Moreover, there is a straightforward analogue of the ordered supports property. In particular, $H(b|z)$ has a weight of $(z-1)/z$ on the left-hand side and $1/z$ on the right-hand side. $H(b|z)$ with lower z therefore provide more "bang for the buck" in relaxing incentive constraints. For example, when $z = 1$, $H(b|1)$ only appears on the right-hand side, and indeed this implies that $H(b|1)$ must put probability one on v^1. Thus, the correct ordered supports property is:

$$H(b|z) < 1 \implies H(b|z') = 0 \, \forall z' > z. \tag{34}$$

In other words, at an optimal solution, buyers should win with lower bids when there are fewer buyers with high valuations.

We summarize the solution of the binary known values relaxed program in the following proposition:

Proposition 10. *The solution to the binary known values relaxed problem is the unique* $H(b|z)$ *that satisfies* (33) *with equality whenever* $H(b|z) < 1$ *for some* z *and also satisfies* (34)*.*

The solution to the binary values relaxed program can be extended to a full BCE using similar techniques as with unknown values. We first draw the valuations according to p. If $z = 0$, the winning bidder will bid v^1. If $z > 1$, we select one of the high-valuation bidders uniformly to be the winning bidder and assign him or her a winning bid b from $H(b|z)$. Finally, we draw bids for the losing high-valuation buyers independently from the distribution

$$p^Z(z)\frac{z-1}{z}L(b|b',z)H(db'|z) = \frac{1}{(v^2-b')^2}\sum_{z'=1}^{\tilde{z}}\frac{p(z')}{z'}\int_{x=0}^{b'}(v^2-x)H(dx|z'), \tag{35}$$

where $b \in B_{\tilde{z}} = \operatorname{supp} H(\cdot|\tilde{z})$ and $b' \in B_z$ with $b' > b$. The following result, proven in the Appendix, asserts that these losing bid distributions, together with the winning bid distributions that solve the relaxed program, constitute a BCE:

Theorem 11 (Min revenue and max bidder surplus for binary known values)**.** *The solution to the binary known value relaxed problem, together with the losing bid distributions defined by* (35)*, constitute a BCE. This BCE simultaneously maximizes total bidder surplus and minimizes revenue over all binary known values BCE.*

We illustrate the binary known value result in the case where there are just two buyers, the possible valuations are 0 and 1, and values are i.i.d.draws with probability $p \in [0,1]$ of drawing a high value. In this case, $H(b|0)$ and $H(b|1)$ put probability 1 on $b = 0$, and $H(b|2)$ solves the following differential equation

$$(1-b)\frac{p}{2}H(b|2) = (1-p)b + \frac{p}{2}\int_{x=0}^{b}H(x|2)dx,$$

so

$$H(b|2) = \frac{1-p}{p}\frac{b(2-b)}{(1-b)^2}.$$

This distribution hits 1 at $\overline{b} = (1-\sqrt{1-p})/2$.

Note that the high type is indifferent to a strategy of always bidding $\overline{b}$, so the surplus of the high type is easily calculated as $1-\overline{b} = \sqrt{1-p}$, and total bidder surplus is therefore $2p\sqrt{1-p}$. The allocation is efficient, so total surplus is $1-(1-p)^2$ and the revenue is

$$1-(1-p)^2 - 2p\sqrt{1-p}.$$

For comparison, in the complete information benchmark, each high type bidder earns a surplus of 1 when facing the low type and a surplus of 0 when facing the high type, so that total bidder surplus is only $2p(1-p)$. Revenue under complete information is p^2. In a previous version of this paper, Bergemann, Brooks, and Morris (2013), we additionally construct simple information structures that give rise to this equilibrium. We also solve for the entire set of bidder surplus pairs that can obtain in a BCE.

The binary values example tells us a great deal about the qualitative differences between the unknown and known values cases. In both models, uniform upward deviations seem to be key for characterizing the revenue minimizing BCE. The information that is needed to evaluate those deviations is, however, quite different in the two cases. With unknown values, the probability of being the winner when the profile of values is some given v is independent of the profile, so that the only manner in which v affects the incentives to deviate uniformly upwards is through inference about the average losing value, $\mu(m)$.

Indeed, we could think about the profile v for a given upward deviation b as being associated with a *gain-loss ratio*, which is the extent to which winning when the profile is v tightens the left-hand side of the uniform upward incentive constraint for a given relaxation of the right hand side, i.e., $(N-1)(\mu(v)-b)$. Importantly, this gain-loss ratio associated with v is the same for all buyers, thus inducing a one-dimensional ordering over profiles. A general lesson from the unknown values analysis is that minimizing revenue requires a specific relationship between the distribution of winning bids and the gain-loss ratio. In particular, low winning bids should be associated with a low gain-loss ratio.

With known values, bidders know their value when they lose the auction, thus rendering $\mu(v)$ irrelevant for the evaluation of uniform upward deviations. On the other hand, the probability of winning versus losing depends very much on the distribution of values and on the bidder's own value: in an efficient equilibrium, having a value below the maximum means that one will lose for sure, and the more ties there are for the highest value, the less likely is a high valuation bidder to win the auction. For example, in the binary known values model, the gain loss-ratio for a high type is $(\zeta(v) - 1)(v^2 - b)$, and for a low type it is always $(v^1 - b)$ (since the low type always loses in equilibrium when the winning bid is above v^1). With binary values, the low type is strategically quite simple, so that it is only the high type's gain-loss ratio that matters and a simple ordering property that emerges for the supports of winning bids. In general, when different buyers have different gain-loss ratios, there is no obvious ordering property over the distribution of winning bids. In Appendix C, we analyze the relaxed program for three values and document the complications that arise with more than two values. We give a partial characterization of the solution to the relaxed program, though it turns out that in these cases where we solve the relaxed program, the bound it generates cannot possibly be tight.

5 Further Topics

5.1 Additional Welfare Outcomes

We have thus far focused on characterizing bounds on revenue and total bidder surplus. Maximizing and minimizing these outcomes correspond to characterizing the BCE that are maximal in four directions in welfare space. Of course, there are many other welfare objectives we might have considered. For example, what do BCE look like that maximize a single bidder's welfare, or that minimize the efficiency of the auction?

In this section, we will briefly consider the range of welfare outcomes in these and other directions. We will report some new analytic results, but also fully develop numerical simulation methods. Bayes correlated equilibria have a simple structure that lend themselves to computation: they are essentially joint distributions over values and bids that satisfy a family of linear incentive constraints, and welfare outcomes such as revenue and surplus are linear functions of that distribution. Maximizing a weighted sum of expected welfare outcomes over the set of BCE is therefore a linear program which can be solved efficiently.

We applied large-scale linear programming software to compute the BCE that maximize various welfare objectives for discretized examples. In particular, we studied a model in which there are two potential buyers who have 10 valuations and 50 possible bids. Values and bids are evenly spaced between 0 and 1, and the distribution of values is uniform. Figure 1 from the introduction depicts the pairs of revenue and total bidder surplus that can arise in a BCE, with the boundary of the unknown and known values welfare sets in blue and red, respectively.

Let us examine these sets in somewhat greater detail than in the introduction. The computed welfare sets have a number of prominent features that align with our theoretical results. The red set is contained within the blue set, which we know must be the case because any known value BCE is also an unknown values BCE. Also, welfare properties of the equilibria that we constructed are visible on the frontiers of the welfare sets. Minimum total bidder surplus is zero and is attained at points F, G, and H, and the unknown values BNE outcome of $(0, 0.5)$ (point B) is on the boundary of the unknown values surplus set. We also see the complete information BNE at point A on the boundary of both the known value and unknown values surplus sets. In this outcome, revenue is the expected second-highest value and total surplus is the expected highest value. As predicted, there are BCE that attain point F in which the outcome is socially efficient, but bidder surplus is zero and revenue is equal to the entire efficient surplus. Finally, minimum revenue for the unknown values model is attained in an efficient equilibrium at point D, as predicted by Theorem 8.

There are however new features which might not have been anticipated. An intriguing feature of the unknown values surplus set is southwestern frontier, in which the total surplus is below the efficient level. As we observed in the introduction, since there is no reserve price in the auction, the good is always allocated. A lower bound on total surplus is therefore the expected lowest valuation, which for this discretized example is $0.31\overline{6}$. A striking implication of Figure 1 is that there are equilibria in which this lower bound is attained exactly, meaning that the buyer with the lower valuation is always the one who receives the good. Moreover, minimum surplus is attained while both bidders receive zero surplus!

In fact, it is not hard to construct a BCE that accomplishes these objectives when there are two bidders and independent values. Let $P(v)$ denote the independent cumulative distribution on the range $[\underline{v}, \overline{v}]$. We construct a BCE in which each buyer i observes buyer j's valuation v_j, and bids

$$\beta(v_j) = \frac{\int_{x=\underline{v}}^{v_j} x\, dP(x)dx}{P(v)}, \tag{36}$$

where $\beta(\underline{v}) = \underline{v}$. In other words, each buyer bids the other buyer's expected value, conditional on it being below their true valuation. As long as the density is almost everywhere positive, this bidding function will be strictly increasing on the support of valuations. We claim that this is a BCE. Clearly, buyers always believe that their own value v_i is distributed according to P and that the other buyer bids $\beta(v_i)$. Conditional on a bid of $\beta(v)$, the buyer will win whenever $v_i \leq v$, so the expected valuation conditional on winning with a bid of b is the expectation of their value given $v_i \leq v$, which is precisely $\beta(v)$! Thus, all bids in the support of P result in an expected payoff of zero. Moreover, it is clear that since equilibrium bids are increasing in the *other* buyer's value, the winner of the auction will be the bidder with the lowest valuation, thus attaining point H. Similar arguments and constructions can be provided when values are discrete, while maintaining the assumptions of independence and two bidders.

Thus, we have characterized the three "corners" of the unknown values set, and by convexity, we can generate the both the western and northeastern flats. The remaining feature, hitherto unexplained, is the apparently smooth southwestern frontier that runs from the maximally inefficient equilibrium to the efficient revenue minimizing equilibrium. In fact, we know a great deal about the class of equilibria that generate this southwestern frontier. They are members of a class of "conditionally revenue minimizing" equilibria, which minimize revenue conditional on a fixed allocation of the good. As the allocation ranges from efficient to maximally inefficient, we move smoothly between points D and H. The southwestern frontier corresponds to a particular path of allocations for which we can give a partial characterization. These additional analytical results are in Appendix D.

The known values surplus set is depicted in red in Figure 1, and it is significantly smaller than the unknown values surplus set. At point C, revenue is maximized and bidders are held down to the lower bound from Section 4. In this example, minimum revenue for the known values model is approximately 0.11, while minimum revenue for the unknown values model is approximately 0.06. This difference corresponds to roughly 8 percent of the efficient surplus.

The extreme points of the surplus sets of Figure 1 are maximal for welfare objectives that are symmetric with respect to the bidders. For example, when we maximize revenue or total bidder surplus, the objective does not depend on which bidder bids which amount, but only on the average maximum bid. For symmetric objectives, it is without loss of generality to consider symmetric BCE (Lemma 7). However, even when the distribution of values is symmetric, asymmetries in information about values or in behavior could induce differences in welfare outcomes across buyers.

Figure 3 displays the sets of bidder surplus pairs (U_1, U_2) that can arise in our discretized uniform example. Again, blue is for unknown values and red is for known values. Points are labeled in correspondence with Figure 1. For example, point A in Figure 3 is the BNE outcome under known values, and corresponds to point A in Figure 1. Point FGH, at which bidder surpluses are minimized, corresponds to points F, G, and H in Figure 1, etc. The point marked E′ is the only labeled point on Figure 3 that does not have a direct counterpart in Figure 1: although the two are close together and are visually difficult to distinguish, minimum revenue and maximum bidder surplus need not coincide for the known values model.

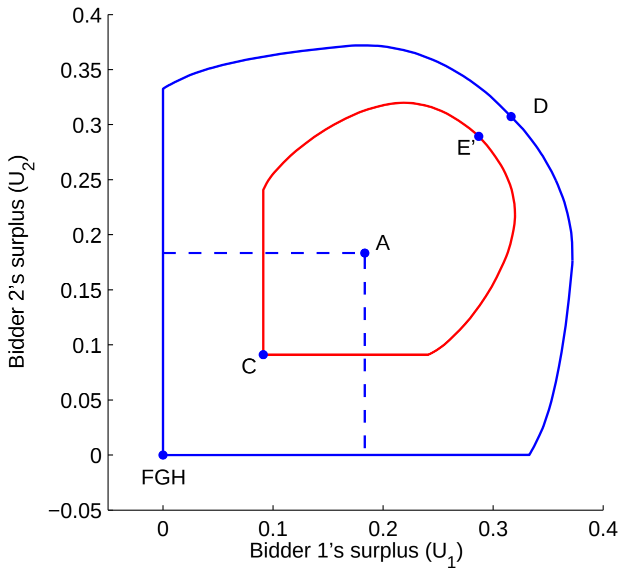

Figure 3: The set of bidder surplus pairs that can arise in a BCE. Computed for uniform distribution with grids of 10 valuations and 50 bids between 0 and 1.

5.2 Max/Min Ratios

The form that our results have taken thus far is to provide welfare bounds over all information structures and BNE that are consistent with a fixed prior distribution of values. But what is the broader message? Are the bounds wide or narrow? To help us get a sense of how much welfare can vary due to information that is less tied to a particular prior, we will employ a methodology that has been widely used in the algorithmic mechanism design literature, which is to study how welfare varies across information structures and BNE relative to some suitable benchmark. A natural candidate in our setting is complete information, in which there are no informational frictions whatsoever.

Thus, we can ask if there are bounds on how much revenue and bidder surplus can vary relative to their respective values in the complete information BNE. Let $\widehat{R} = \mathbb{E}[v^{(2)}]$ denote the expected second-highest value, and $\widehat{U} = \mathbb{E}[v^{(1)} - v^{(2)}]$ denote the residual surplus for bidders, which is the difference between the highest and second-highest values. Analogously, let $\overline{R}$, $\underline{R}$, $\overline{U}$, and $\underline{U}$ denote the highest and lowest revenues and highest and lowest total bidder surpluses, respectively, that occur in BCE. We will consider the ratios:

$$\underline{R}/\widehat{R}, \overline{R}/\widehat{R}, \underline{U}/\widehat{U}, \text{ and } \overline{U}/\widehat{U}.$$

Bounds on how far these ratios can differ from 1 would mean that these welfare objectives cannot be either too much larger or too much smaller than the complete information benchmark, regardless of the prior p.

It turns out, however, that there are no such bounds: there exist distributions of values such that these welfare ratios are arbitrarily small and arbitrarily large. Moreover, this remains true even if we restrict attention to known values information structures. We illustrate this with three examples.

First, we show that $\underline{R}/\widehat{R}$ and $\overline{U}/\widehat{U}$ can be arbitrarily small and large, respectively. Consider the binary values model with $V = \{0,1\}$, with p being the independent and symmetric probability of a high type. Revenue in the complete information benchmark is therefore p^2, and minimum revenue is $1-(1-p)^2 - 2p\sqrt{1-p}$. As p goes to zero, we can verify by L'Hôpital's rule that the ratio of minimum to benchmark revenue converges to

$$\lim_{p\downarrow 0} \frac{\frac{p}{\sqrt{1-p}} - 2\sqrt{1-p} + 2(1-p)}{p} = \lim_{p\downarrow 0}\left[\frac{2}{\sqrt{1-p}} - \frac{p}{2(1-p)^{\frac{3}{2}}} - 2\right] = 0.$$

With regard to $\widehat{U}/\overline{U}$, maximum bidder surplus is $2p\sqrt{1-p}$ and benchmark bidder surplus is $2p(1-p)$. Thus, the ratio $\overline{U}/\widehat{U}$ is $1/\sqrt{1-p}$, which blows up as $p \uparrow 1$.

Second, $\overline{R}/\widehat{R}$ can be arbitrarily large. Consider two bidders with values independently drawn from the cumulative distribution $P(v) = v^{\beta}$. The lower bound on bidder surplus is

$$\underline{U} = \int_{v=0}^{1} \max_{b}(v-b)P(b)p(v)dv.$$

With the distribution we have chosen, the optimal bid is $b^*(v) = \frac{\beta}{1+\beta}v$, so that lower-bound bidder surplus is $\frac{1}{1+2\beta}\left(\frac{\beta}{1+\beta}\right)^{\beta+1}$. The expected highest value is $\beta/(1+2\beta)$, so that maximum revenue is $\frac{\beta}{1+2\beta} - \frac{1}{1+2\beta}\left(\frac{\beta}{1+\beta}\right)^{\beta+1}$. The expected second-highest value is $\frac{\beta}{1+\beta}\frac{\beta}{1+2\beta}$, so the ratio of maximum to benchmark revenue is

$$\frac{1+\beta-\left(\frac{\beta}{1+\beta}\right)^{\beta}}{\beta},$$

which clearly blows up as $\beta \to 0$.

Finally, $\underline{U}/\widehat{U}$ can be arbitrarily close to zero. We will suppose that $v^{(1)} \sim U[0,1]$, and that conditional on the highest value, the second-highest value follows a truncated Pareto distribution on $[0, \gamma v^{(1)}]$, i.e.,

$$P(v^{(2)}|v^{(1)}) = \begin{cases} 1 & \text{if } v^{(2)} > \gamma v^{(1)}; \\ (1-\gamma)\,\frac{v^{(1)}}{v^{(1)}-v^{(2)}} & \text{if } 0 \leq v^{(2)} \leq \gamma v^{(1)}; \\ 0 & \text{otherwise.} \end{cases}$$

The expected highest value is 1/2, and the optimal bid is always $\gamma v^{(1)}$, so that the lower bound on bidder surplus is just $(1-\gamma)/2$. Benchmark revenue is a slightly more complicated calculation. The expected second-highest value, conditional on $v^{(1)}$, is

$$(1-\gamma)v^{(1)} \int_{x=0}^{\gamma v^{(1)}} \frac{x}{(v^{(1)}-x)^2} dx = (\gamma + (1-\gamma)\log(1-\gamma))v^{(1)}.$$

Hence, the ratio of the benchmark to lower bound bidder surplus is $1 - \log(1-\gamma)$, which goes to infinity as γ goes to one.

Thus, we find that the range of welfare outcomes can be very wide, depending on the specification of the prior. The extreme ratios are however obtained in extreme cases where either benchmark revenue or benchmark bidder surplus is going to zero. Another weakness of these ratios is that they are not invariant to translation of values: by adding a constant to all values, it is possible to make any of these ratios arbitrarily close to 1. In that sense, our bounds can also be arbitrarily narrow. It remains an open question whether there are natural classes of distributions or more intuitive benchmark comparisons for which welfare ratios are bounded.

5.3 Reserve Prices and Entry Fees

The first price auction is an important mechanism to study for many reasons, but at the end of the day it is just one of many possible mechanisms. Indeed, even in the classical setting when values are known and the distribution is symmetric, independent, and regular, the first price auction is generally only optimal if low valuation buyers are excluded from the auction using a minimum bid or a participation fee (Myerson, 1981). A general treatment of welfare bounds in mechanism design goes well beyond the scope of this paper, but we can make some progress by considering simple variants of the first price auction that include some device for the exclusion of low valuation buyers, i.e., entry fees and reserve prices.

We will explore these variants via a numerical example. Suppose that there are two buyers whose values are uniformly distributed on a grid between 0 and 1. Reserve prices and entry fees are modeled as follows: With a reservation price $r \geq 0$, buyers can submit

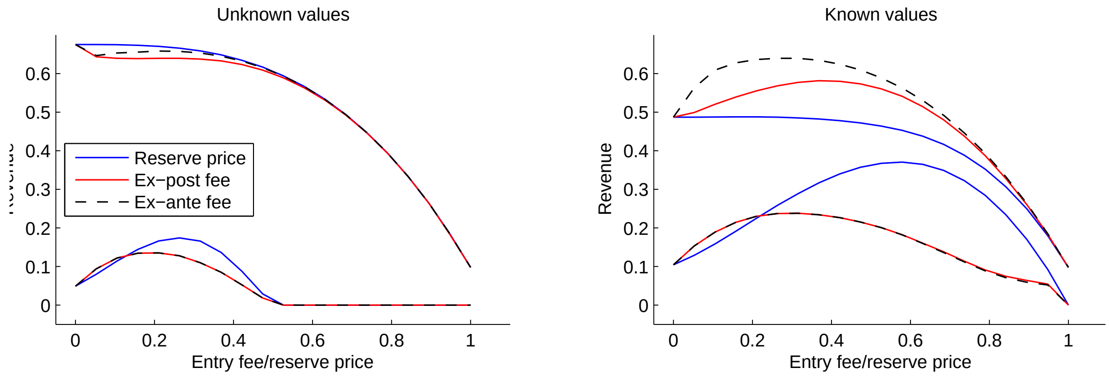

Figure 4: Comparison of reserve prices and entry fees. Computed for uniformly distribution with grids of 20 valuations and 20 bids between 0 and 1.

any bid they want, but if the high bid is less than the reservation price, the good will not be allocated and no transfers take place. Otherwise, the good is allocated to the buyer who bids the most, and the winning buyer pays their bid to the seller. With an entry fee, buyers first choose whether or not to enter the auction. Only entrants can bid, but an entrant must pay an additional fee of $e \geq 0$ regardless of whether or not the entrant wins the auction.

In the absense of an entry fee, we can interpret a bid of 0 (or any bid below the reserve price) as "not entering," so that the reserve price model effectively captures a model with a zero entry fee. With a positive entry fee, however, some modeling choices arise with regard to the timing of information. In particular, do buyers make their entry decisions before or after observing whatever signals will inform the equilibrium bid? Similarly, in the known values model, we may wish to assume that buyers learn their values before or after the entry decision. Neither alternative is obviously more compelling, so we will simulate both and compare.[13]

Again, we used linear programming software to numerically calculate maximum and minimum revenue for a range of reserve prices and entry fees. Figure 4 displays the output of these computations for a specification with 20 values and 20 bids, evenly spaced between 0 and 1. The left and right panels illustrate the revenue bounds for the unknown and known value models, respectively. The blue lines are bounds on revenue with reserve prices. The black lines bound revenue over all BCE with an "ex-ante" fee, which must be paid *before* the buyers acquire whatever signals inform their equilibrium bid, while the red lines bound revenue with an "ex-post" fee, which is paid *after* learning one's equilibrium bid.

[13] An additional timing question arises with entry fees in the known values model. In principle, bidders might learn their value before they learn their equilibrium bid, in which case we have to ask whether or not the fee is paid (i) before learning the value, (ii) after learning the value and before the bid, or (iii) after both learning the value in the bid. We will use (ii): buyers first learn their value, then pay the fee, then learn the bid.

Note that a BCE with an ex-post fee is also a BCE with an ex-ante fee, so that the black bounds are necessarily wider than the red bounds. In addition, since known values BCE are also unknown values BCE, the bounds on the left panel are necessarily wider than their counterparts on the right panel.

The simulation indicates some intriguing features of how reserve prices and entry fees could affect the auction. Consistent with the classical theory, it is very much possible for both reserve prices and entry fees to raise minimum revenue over all information structures. However, reserve prices are much more effective at boosting minimum revenue than are entry fees, at least for this example. In the known values model, $\max\min$ revenue over all reserve prices is approximately 0.3705, whereas $\max\min$ revenue over entry fees is 0.2379. With regard to entry fees, a surprising result is that ex-ante and ex-post entry fees result in virtually identical minimum revenue curves. We do not know of any theoretical justification for this phenomenon.

With regard to maximum revenue, if values are unknown, revenue might be equal to the efficient surplus. Thus, excluding low valuation buyers can only depress revenue. With known values, however, it appears that exclusion can increase maximum revenue. In addition, there is a stark ordering over maximum revenues. For each $x \in [0, 1]$, maximum revenue with an ex-ante fee of x is strictly higher than maximum revenue with an ex-post fee of x, which is strictly higher than maximum revenue with a reserve price of x. The weak ordering of ex-ante and ex-post fees is fairly transparent, but maximum revenue ranking of entry fees versus reserve prices is more subtle. In fact, we have found a relatively simple argument that an ex-ante fee is always going to generate weakly more revenue than a reserve prices, which we present in Appendix E. This conclusion corresponds with a result of Milgrom and Weber (1982) that entry fees induce greater revenue than reserve prices when signals and values are affiliated.

In sum, this example demonstrates that reserve prices and entry fees can have a large impact on revenue bounds. In future research, we hope to extend the analysis of welfare bounds to more general classes of mechanisms.

5.4 Continuum of Values

Though our formal results have been stated for the case of discrete values, we have frequently illustrated these results with examples in which there is a continuum of valuations. While discreteness has been useful for our inductive constructions, we do not regard it as essential, and we expect that our results would generalize to distributions of values characterized by Borel measures. However, the extension of the results to the continuum case seems to be a non-trivial technical exercise, for the simple reason that the set of BCE is not weak-* compact, and so sequences of BCE need not converge to BCE in the limit. For example, it may be that all along a sequence of BCE, bids are not equal with probability 1, but in the limit, bids are equal with probability 1, thus invoking the tie breaking rule. This is a well-studied phenomenon, and we refer the reader to Jackson and Swinkels (2005) and Jackson et al. (2002) for a more detailed discussion of the issue.

5.5 Extension to Asymmetric p

Throughout our analysis, we have maintained the assumption that the distribution p is symmetric. This has greatly simplified our arguments, especially for deriving solutions to the minimum revenue and maximum bidder surplus optimal control problems. We wish to pointout, however, that some of our results extend readily to models with asymmetric value distributions. Nothing in the proof of Theorem 2 relied on symmetry, and indeed, that result extends unchanged to the case where p is asymmetric, with different minimum surpluses for each bidder.

Symmetry is more important for our minimum revenue and maximum bidder surplus constructions, and there is no simple generalization to asymmetric priors. Our results do however generate bounds for asymmetric models: suppose that p is asymmetric and minimum expected revenue over all BCE is R. Then R is also minimum revenue over all BCE for the permuted distribution of values given by $p \circ \xi$ for some $\xi \in \Xi$, since all we have done is rearrange the identities of the bidders. Now consider the "symmetrized" distribution

$$\tilde{p} = \frac{1}{n!} \sum_{\xi \in \Xi} p \circ \xi.$$

Then clearly, there is a BCE for $\tilde{p}$ which is just the symmetrization of the BCE that yielded revenue of R for the original p, and this BCE will also yield expected revenue of R. As a result, R must be weakly larger than minimum revenue for $\tilde{p}$, for which we have tight characterizations.

6 Conclusion

The purpose of this paper has been to study the range of welfare outcomes that might obtain in a first price auction. In this exercise, we have sought to relax classical assumptions on the nature of the bidders' information that were made primarily for the purpose of tractability. For general specifications of information, in which values can be arbitrarily correlated and signals can be multidimensional, the structure of Bayes Nash equilibrium could be quite complicated. By focusing on those information structures that generate extreme welfare outcomes, we have maintained tractability while generating new insights into the range of welfare outcomes that might occur. In particular, we have shown that while there is a wide range of behavior that could occur, there are non-trivial limits on what can happen to revenue, bidder surplus, and total surplus. The information structures that generate extreme outcomes give us further insight into what kinds of information might be good or bad for the seller and buyers. In particular, revenue is lower when buyers receive partial information about whether or not they have the high value that induces them to bid less, opening the door for other partially informed buyers to win with lower bids as well. Revenue is higher when buyers receive precise information about whether or not they have the high value, but partial information about the losing buyer's value. Social surplus can be harmed when there is precise information about others' values and in the absense of precise information about the buyer's own value.

Many important questions remain. In a computational exercise, we showed that the introduction of devices for excluding low valuation buyers can raise both minimum and maximum revenue over all information structures. While these results are suggestive, clearly they represent a very limited exploration of welfare bounds across the universe of mechanisms. A rich and open question is how to design mechanisms in the face of large uncertainty about the beliefs held by agents. A takeaway from our work is that the design of equilibria can be complemented by the design of information. By simultaneously constructing knowledge and behavior, we obtain bounds on the outcomes of interest as well as gain insight into how information interacts with the given game form. It is our hope that this methodology will find use for a wide range of problems within mechanism design.

References

Āzacis, H. and P. Vida (2015): "Collusive communication schemes in a first-price auction," *Economic Theory*, 58, 125–160.

Bergemann, D., B. Brooks, and S. Morris (2015a): "The Limits of Price Discrimination," *The American Economic Review*, 105, 921–957.

Bergemann, D., B. A. Brooks, and S. Morris (2013): "Extremal Information Structures in the First Price Auction," *Cowles Foundation Discussion Paper*.

——— (2015b): "Bayes Correlated Equilibrium in Infinite Games," *Discussion Paper, Yale University and the University of Chicago and Princeton University*.

Bergemann, D. and S. Morris (2013): "Robust predictions in games with incomplete information," *Econometrica*, 81, 1251–1308.

——— (2015): "Bayes correlated equilibrium and the comparison of information structures in games," Tech. rep., Cowles Foundation for Research in Economics, Yale University.

Clarke, F. (2013): *Functional analysis, calculus of variations and optimal control*, vol. 264, Springer Science & Business Media.

Cohn, D. L. (1980): *Measure theory*, vol. 1993, Springer.

Fang, H. and S. Morris (2006): "Multidimensional private value auctions," *Journal of Economic Theory*, 126, 1–30.

Forges, F. (1993): "Five legitimate definitions of correlated equilibrium in games with incomplete information," *Theory and decision*, 35, 277–310.

Jackson, M. O., L. K. Simon, J. M. Swinkels, and W. R. Zame (2002): "Communication and equilibrium in discontinuous games of incomplete information," *Econometrica*, 70, 1711–1740.

Jackson, M. O. and J. M. Swinkels (2005): "Existence of equilibrium in single and double private value auctions1," *Econometrica*, 73, 93–139.

Kaplan, T. R. and S. Zamir (2014): *Advances in Auctions*, Elsevier.

Kim, J. and Y.-K. Che (2004): "Asymmetric information about rivals' types in standard auctions," *Games and Economic Behavior*, 46, 383–397.

Landsberger, M., J. Rubinstein, E. Wolfstetter, and S. Zamir (2001): "First–price auctions when the ranking of valuations is common knowledge," *Review of Economic Design*, 6, 461–480.

Maskin, E. S. and J. G. Riley (1985): "Auction theory with private values," *The American Economic Review*, 150–155.

Milgrom, P. R. and R. J. Weber (1982): "A theory of auctions and competitive bidding," *Econometrica: Journal of the Econometric Society*, 1089–1122.

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of operations research*, 6, 58–73.

A Omitted Proofs

Proof of Theorem 2. We first construct the BCE that attains the maximum revenue and the minimum bidder surplus. The BCE is constructed as follows. A bidder i with the highest valuation v_i, and hence $v^{(1)} = v_i$ is informed that he is the winning bidder and receives a bid recommendation $b_i \in \left\{v^1, ..., v^{(1)}\right\}$. The probability distribution of his bid recommendation is constructed in detail in the subsequent paragraph. Every losing bidder is informed that he is a losing bidder and is asked to bid his value v_i. This is, every losing bidder with the exception of the bidder with the second highest valuation $v^{(2)}$ when the bid recommendation of the winner is $b_i = v^{(2)}$. In this case, the bidder j with $v_j = v^{(2)}$ is asked to chose his bid b_j from the cumulative distribution

$$F_j\left(b_j \left| v^{(1)}, v^{(2)}\right.\right) = \frac{v^{(1)} - v^{(2)}}{v^{(1)} - b_j}$$

on a small interval $\left[v^{(2)} - \epsilon, v^{(2)}\right]$ for some ϵ sufficiently small. With this distribution the winning bidder receives the object with probability one in the symmetric tie-breaking rule even though he just bids the value of the second highest valuation.

Now we construct the bid distribution of the winning bidder with $v_i = v^{(1)}$. We will draw b_i according to the following procedure based on the conditional distribution of the second highest valuation $v^{(2)}$ given the highest valuation $v^{(1)}$. Let $p_i^{(2)}(v^{(2)}|v^{(1)})$ denote the conditional probability of the second-highest value $v^{(2)}$ given that bidder i has the highest value, that is $v^{(1)} = v_i$. In particular, this is:

$$p_i^{(2)}(w'|w) = \frac{\sum_{\left\{v \in V^N | v_i = w = v^{(1)}, v^{(2)} = w'\right\}} p(v)}{\sum_{\left\{v \in V^N | v_i = w = v^{(1)}\right\}} p(v)},$$

and for the corresponding cumulative distribution we write

$$P_i^{(2)}(w'|w) = \sum_{w'' \leq w'} p_i^{(2)}(w''|w)$$

(These objects are only defined when there exists a $v \in V^N$ such that $v_i = v^{(1)} = v$).

Let $k_i^*(v_i) \in \{1, ..., K\}$ be the index of the valuation such that $b_i^*(v_i) = v^{k_i^*(v_i)}$. We now construct the distribution of valuations and (almost bids) that the winning bidder faces when he is asked to make a bid $b_i = v^l$. Let α^l be defined for all $v^{k_i^*(v_i)} \leq v^l \leq v_i$ by:

$$(v_i - v^l)\left(\alpha^l P_i^{(2)}(v^{l-1}|v_i) + p_i^{(2)}(v^l|v_i)\right) = (v_i - b_i^*(v_i))\alpha^l P_i^{(2)}(b_i^*(v_i)|v_i), \tag{37}$$

and rearranging:

$$\alpha^l = \frac{(v_i - v^l)p_i^{(2)}(v^l|v_i)}{(v_i - b_i^*(v_i))P_i^{(2)}(b_i^*(v_i)|v_i) - (v_i - v^l)P_i^{(2)}(v^{l-1}|v_i)}. \tag{38}$$

By the construction of (4), it follows that $\alpha^l = 1$ if $k_i^*(v_i) = l$, and that $0 < \alpha^l < 1$ if $v^{k_i^*(v_i)} < v^l < v_i$, and if $p_i^{(2)}(v^l|v_i) = 0$, then α^l is defined to be 0.)

Now based on the weights α^l that oversample v^l relative to its true frequency in the conditional distribution of $v^{(2)}$, the bid $b_i = v^l$ is recommended to bidder i with probability

$$x^l = \prod_{l'=l+1}^{k} (1 - \alpha^{l'}) \tag{39}$$

when $v^{(2)} = v^l$, and with probability

$$y^l = \alpha^l x^l \tag{40}$$

when $v^{(2)} < v^l$.

We claim that these rules define a BCE. Note that the bidder j with $v_j = v^{(2)}$ only receives recommendation $b_j = v^l$ when $v^{(2)} \leq v^l$, so the only way to deviate upwards and win the auction when one would lose by following the recommendation is to bid more than $v^{(2)}$, which would result in negative surplus. Thus, it is sufficient to check that no bidder would like to deviate downwards. To assess the value of such deviations, first observe that the conditional distribution of $v^{(2)}$ conditional on $b_i = v^l$ has a particular form: $v^{(2)} = v^l$ with probability

$$\frac{p_i^{(2)}(v^l|v_i)}{p_i^{(2)}(v^l|v_i) + \alpha^l P_i^{(2)}(v^{l-1}|v_i)},$$

and $v^{(2)} = v^{l'}$ for $l' < l$ with probability

$$\frac{\alpha^l p_i^{(2)}(v^{l'}|v_i)}{p_i^{(2)}(v^l|v_i) + \alpha^l P_i^{(2)}(v^{l-1}|v_i)}.$$

Thus, by construction bidder i is indifferent between a bid of v^l and $b_i^*(v_i)$, and moreover the latter bid is superior to any bid other bid $v^{l'}$ with $l' < l$ by the definition of $b_i^*(v_i)$ and the fact that probabilities of winning for $v^{l'}$ with $l' < l$ are all proportional to $P_i^{(2)}(v^{l'}|v_i)$. Moreover, by choosing ϵ sufficiently small, we can be sure that bids in V are superior to bids not in V.

Finally, this construction ensures that bidder i is always indifferent between following the recommendation b_i and bidding $b_i^*(v_i)$. Thus, the surplus in equilibrium must be equal to the surplus from following the latter strategy, which is precisely $\underline{U}_i(v_i)$. Moreover, a bidder with $v_i = v^{(1)}$ always wins the auction, so that the outcome is efficient. $\square$

Proof of Theorem 3. We construct the BCE as follows. If $v^{(1)} = v^{(2)}$, then all high bidders bid $v^{(1)}$, and losing bidders can bid anything less than $v^{(1)}$.

Otherwise, if $v^{(1)} > v^{(2)}$, the equilibrium first draws an x in $[0,1]$ according to the distribution $F(x) = x^k$. The high value bidder is then told to bid $\overline{b} = xv^{(1)} + (1-x)v^{(2)}$, and losing bidders are told to bid $yv^{(1)} + (1-y)v^{(2)}$ with $y \in [0,x]$, where $y \sim \frac{G(y)}{G(x)} = \left(\frac{y}{1-y}\frac{1-x}{x}\right)^k$ and $G(x) = \left(\frac{x}{1-x}\right)^k$.

We claim that this is a BCE for all k. Clearly, there is always positive conditional probability that a given bid is both a winning bid and a losing bid. In particular, conditional on the profile of values, a recommendation in the range $[v^{(2)}, b]$ is a winning bid with probability

$$\frac{1}{N}G\left(\frac{b - v^{(2)}}{v^{(1)} - v^{(2)}}\right),$$

and a losing bid is recommended in the range $[v^{(2)}, b]$ with probability

$$\frac{N-1}{N}G\left(\frac{b - v^{(2)}}{v^{(1)} - v^{(2)}}\right)\left(1 + \int_{x=\frac{b-v^{(2)}}{v^{(1)}-v^{(2)}}}^{1} \frac{g(x)}{G(x)}dx\right).$$

Both of these expressions are differentiable and strictly increasing on the range $b \in [v^{(2)}, v^{(1)}]$, and are zero when $b = v^{(2)}$, and thus bids in any open subinterval of $[v^{(2)}, v^{(1)}]$ arise with positive probability conditional on both winning and losing. Thus, conditional on any bid, it is possible to have value $v^{(1)}$, so that the strategy of bidding b is undominated. However, if b is a losing recommendation, $b > v_i$, so that it is never profitable to deviate to a higher bid.

In addition, conditional on it being a winning bid $b = xv^{(1)} + (1-x)v^{(2)}$, and conditional on the highest and second highest values, the benefit of deviating downward to some $b'^{(1)} + (1-y)v^{(2)}$ with $y \geq 0$ is proportional to

$$(v^{(1)} - v^{(2)})(1-y)\frac{G^{N-1}(y)}{G^{N-1}(x)} \leq (v^{(1)} - v^{(2)})(1-y)\frac{G(y)}{G(x)}$$

since $\frac{G(y)}{G(x)} < 1$, and moreover the left-hand side and right-hand side coincide at $y = x$. Thus, it is sufficient to show that $(1-y)G(y)$ is increasing in y. This function is $\frac{y^k}{(1-y)^{k-1}}$, which is obviously increasing in y since the numerator is increasing and the denominator is decreasing.

Finally, as k goes to ∞, the expected value of x converges to 1. For if we write $H(x)$ for any CDF that is strictly less than 1 with probability 1, then

$$\begin{aligned}\int_{x=0}^{1} x dH^k(x) &= xH^k(x)|_{x=0}^{1} - \int_{x=0}^{1} H^k(x)dx \\ &= 1 - \int_{x=0}^{1} H^k(x)dx \\ &\geq 1 - zH^k(z) - (1-z)\end{aligned}$$

for any $z \in (0,1)$ (this just comes from the observation that $H(x) \leq H(z)$ for $x \in [0,z]$ and $H(x) \leq 1$ for all $x \in [z,1]$). Since $H^k(z) < 1$, by choosing z sufficiently close to 1 and k sufficiently large, we can make this number arbitrarily close to 1. Thus, in the limit as k goes to ∞, it must be that $\mathbb{E}\left[\overline{b}\right] = \mathbb{E}[x]\sum_v p(v)v^{(1)} + (1-\mathbb{E}[x])\sum_v p(v)v^{(2)}$ goes to $\sum_v p(v)v^{(1)} = \overline{S}$. Since $S(F) = R(F) + U(F) \leq \overline{S}$, it must be that $U(F)$ goes to 0. □

Proof of Lemma 4. The proof essentially follows the derivation of the relaxed problem in Section 4.1.2. Let

$$H(b|v_1,\ldots,v_n) = \int_{\left\{x\in\mathbb{R}_+^N | x_i \leq b \,\forall i\right\}} q_1(x)F(dx|v_1,\ldots,v_N). \tag{41}$$

Thus, our convention is that it is bidder 1 who wins the auction. Since the BCE is symmetric, H so defined is invariant to permutations of $(v_2,\ldots,v_N)$. Thus, the objective of minimizing revenue is equivalent to maximizing

$$\sum_{v\in V^N} p(v)\int_{b=0}^{v^K} H(b|v)db, \tag{42}$$

which is the many-bidder analogue of (11). Recall that Ξ denotes the set of permutations of $\{1,\ldots,N\}$, and we associate a permutation $\xi \in \Xi$ with a mapping from V^N to V^N, where $\xi(v)$ is the permuted profile of valuations in which $\xi_i(v) = v_{\xi(i)}$. The feasibility constraint that is analogous to (10) requires that

$$\sum_{\xi\in\Xi} H(b|\xi(v)) \in [0,1] \,\forall v \in V^N. \tag{43}$$

Finally, the incentive constraint that is analogous to (12) says that

$$(N-1)\sum_{v\in V^N} p(v)H(b|v_1,v_{-1})\left(\frac{1}{N-1}\sum_{i\neq 1} v_i - b\right) \leq \sum_{v\in V^N} p(v)\int_{x=0}^{b} H(x|v)dx. \qquad (44)$$

This expression is close to (12), with a notable difference being the sum over bidders other than i on the left hand side. To derive this equation, let us consider uniform upward deviation $\sigma_i^b(x) = \max\{x, b\}$. Then using (3), we can write the IC constraint for these deviations as

$$\sum_{v\in V^N} p(v)\int_{\left\{x\in\mathbb{R}_+^N | x_i\leq b\right\}} \left[(v_i - x_i)q_i(x) - (v_i - b)q_i(b, x_{-i})\right] F(dx|v) \geq 0.$$

Since $q_i(b, x_{-i}) \leq 1$, then clearly this constraint implies that

$$\sum_{v\in V^N} p(v)\left[\int_{\left\{x\in\mathbb{R}_+^N | x_i\leq b\right\}} (v_i - x_i)q_i(x)F(dx|v) - (v_i - b)F(b\mathbf{1}|v)\right] \geq 0.$$

Now, from (41) and the assumption of symmetry, it must be that

$$H(db|v_i, v_{-i}) = \int_{\left\{x\in\mathbb{R}_+^N | x_i=b\right\}} q_i(x)F(dx|v).$$

Moreover, $F(b\mathbf{1}|v)$ is the total probability that *some* buyer wins with a bid less than b when values are v, which is $\sum_{i=1}^N H(b|v_i, v_{-i})$. Thus, we can rewrite this as

$$\sum_{v\in V^N} p(v)\left[\int_{x=0}^{b} (b-x)H(db|v_i, v_{-i}) - (v_i - b)\sum_{j\neq i} H(b|v_j, v_{-j})\right] \geq 0.$$

So now, integrating by parts, we conclude that

$$\sum_{v\in V^N} p(v)\left[\int_{x=0}^{b} H(x|v_i, v_{-i})dx - (v_i - b)\sum_{j\neq i} H(b|v_j, v_{-j})\right] \geq 0.$$

Note that by symmetry, $\sum_{v\in V^N} H(b|v_i, v_{-i}) = \sum_{v\in V^N} H(b|v_1, v_{-1})$ for all i. Also, note that

$$\sum_{v\in V^N} p(v)(v_i - b)\sum_{j\neq i} H(b|v_j, v_{-j}) = \sum_{v\in V^N} p(v)H(b|v_i, v_{-i})\sum_{j\neq i}(v_j - b).$$

By pulling out a factor $N-1$, we obtain (44).

Thus, an intermediate relaxed program is to maximize (42) subject to (43) and (44). It is without loss of generality to consider efficient solutions, in which

$$H(b|v_1, v_{-1}) = 0 \text{ if } v_1 < \max v. \tag{45}$$

To see this, we will transform the relaxed program one more time. Recall that Ξ is the set of permutations, which we have identified with mappings from V^N to V^N. Thus, we can let V^N/Ξ denote the set of equivalence classes modulo permutation, i.e., $[v] \in V^N/\Xi$ denotes

$$\{\xi(v) \, | \xi \in \Xi\}.$$

Moreover, $p([v])$ is the probability of $[v]$ that is induced by $p(v)$. From symmetry, we know that $H(b|v_i, v_{-i})$ is invariant to permutations of the v_{-i}. Thus, the distribution only depends on the set of valuations and the valuation of the winner. We can write $H(b|k, [v])$ for this permutation-invariant winning bid distribution when the profile is in $[v]$ and the winner's value is v^k. We also let

$$H(b|[v]) = \sum_{k=1}^{K} c(k, [v]) H(b|k, [v])$$

denote the distribution of winning bids given $[v]$, where

$$c(k, [v]) = |\{i|v_i = v^k\}|$$

is the number of values in $[v]$ with value v^k, which is obviously invariant to the choice of representative. We finally denote by $k^*([v])$ the largest k such that $c(k, [v]) > 0$ and $c^*([v]) = c(k^*([v]), [v])$.

With this notation in hand, we can rewrite the relaxed program as

$$\max \sum_{[v] \in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k, [v]) \int_{x=0}^{v^K} H(x|k, [v]) dx, \tag{46}$$

subject to

$$\sum_{k=1}^{K} c(k,[v])H(b|k,[v]) \in [0,1] \forall [v] \in V^N/\Xi; \tag{47}$$

$$\begin{aligned}&\sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v])(v^k - b)(H(b|[v]) - H(b|k,[v])) \\ &\quad \leq \sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v]) \int_{x=0}^{b} H(x|k,[v])dx.\end{aligned} \tag{48}$$

Now, an inefficient solution corresponds to one where $H(b|k,[v]) > 0$ even though $k < k^*([v])$. We can perturb the solution to

$$\tilde{H}(b|k,[v]) = \begin{cases} \frac{1}{c^*([v])} \sum_{k=1}^{K} c(k,[v])H(b|k,[v]) \text{ if } k = k^*([v]); \\ 0 \text{ otherwise.} \end{cases}$$

Clearly, this solution will still satisfy (43) and result in the same objective (42), since

$$\begin{aligned}\sum_{k=1}^{K} c(k,[v])\tilde{H}(k,[v]) &= c^*([v])\tilde{H}(k^*([v]),[v]) \\ &= \sum_{k=1}^{K} c(k,[v])H(k,[v]).\end{aligned}$$

Moreover, the incentive constraint will be relaxed, since

$$\begin{aligned}&\sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v])(v^k - b)(\tilde{H}(b|[v]) - \tilde{H}(b|k,[v])) \\ &= \sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v])(v^k - b)(H(b|[v]) - H(b|k,[v])) \\ &\quad + \sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v])(v^k - b)\big(\tilde{H}(b|[v]) - H(b|[v]) - \tilde{H}(b|k,[v]) + H(b|k,[v])\big).\end{aligned}$$

But since $H(b|[v]) = \tilde{H}(b|[v])$, the last line is equivalent to

$$\begin{aligned}
&\sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v])(v^k - b)\big(H(b|k,[v]) - \tilde{H}(b|k,[v])\big) \\
&= \sum_{[v]\in V^N/\Xi} p([v]) \left(\sum_{k=1}^{K} c(k,[v])(v^k - b) H(b|k,[v]) - (\max v - b) c^*([v]) \tilde{H}(b|k^*[v],[v]) \right) \\
&= \sum_{[v]\in V^N/\Xi} p([v]) \left(\sum_{k=1}^{K} c(k,[v])(v^k - b) H(b|k,[v]) - (\max v - b) \sum_{k=1}^{K} c(k,[v]) H(b|k,[v]) \right) \\
&= \sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v])(v^k - \max v) H(b|k,[v]) \leq 0.
\end{aligned}$$

Thus, the right-hand side of the IC constraint has weakly decreased, though the right-hand side is the same, which proves that incentive compatibility is still satisfied.

We can therefore restrict attention to efficient solutions to the relaxed program. The constraint (48) then simplifies to

$$\begin{aligned}
&\sum_{[v]\in V^N/\Xi} p([v]) \left(\sum_{k=1}^{k^*(v)} c(k,[v])(v^k - b) - (\max v - b) \right) c^*([v]) H(k^*([v]),[v]) \\
&= \sum_{[v]\in V^N/\Xi} p([v])(N-1) \left(\frac{1}{N-1} \left(\sum_{k=1}^{k^*(v)} c(k,[v]) v^k - \max v \right) - b \right) c^*([v]) H(k^*([v]),[v]) \\
&= \sum_{[v]\in V^N/\Xi} p([v])(N-1) \left(\mu(v) - b\right) c^*([v]) H(k^*([v]),[v]) \\
&\qquad \leq \sum_{[v]\in V^N/\Xi} p([v]) \sum_{k=1}^{K} c(k,[v]) \int_{x=0}^{b} H(x|k,[v]) dx.
\end{aligned}$$

Using the functions μ and p^L defined by (17) and (18), we can define

$$H(b|m) = \frac{1}{p^L(m)} \sum_{\{[v]\in V^N/\Xi | \mu(v)=m\}} p([v]) c^*([v]) H(b|k^*([v]),[v]) \tag{49}$$

as the distribution of winning bids conditional on the average losing value. Substituting this expression into (46) yields , (48) becomes (21), and (47) becomes (20). □

Proof of Proposition 5. First, we observe that a solution to the unknown values relaxed program must exist. The set of bounded and measurable functions from $\left[0, v^K\right]$ to $[0,1]$ is compact. Moreover, for each $b \in \mathbb{R}_+$, (21) is a linear restriction that is continuous in H in the weak-* topology, so that the set of distributions satisfying (21) is closed. Therefore, the feasible set for the relaxed program is compact, and the objective is weak-* continuous, so a solution exists.

Next, let us show that $H(b|m) = 0$ almost surely for all $b < \underline{m} \triangleq \min M$. Let us define the function

$$H(b) = \sum_{m \in M} p^L(m) H(b|m).$$

Then (21) implies that

$$(\underline{m} - b)\frac{N-1}{N} H(b) \leq \frac{1}{N}\int_{x=0}^{b} H(x)dx.$$

Clearly, $H(b|m) > 0$ for some v^L if and only if $H(b) > 0$ as well, so it is sufficient to prove that $H(b) = 0$ almost surely for $b < \underline{m}$. If $H(b) > 0$ for some non-null set of $b < \underline{m}$, then there exists a $b < \underline{m}$ such that $\int_{x=0}^{b} H(x)dx > 0$. Let $\underline{b}$ be the infimum of such b (here we are using the assumption that B is bounded below). Then it must be that $H(x) = 0$ for $x < b$. Let $z(\epsilon) > 0$ be the supremum of $H(b)$ over the range $[\underline{b}, \underline{b} + \epsilon]$ for some small ϵ. Then it must be that

$$(\underline{m} - \underline{b} - \epsilon)\frac{N-1}{N} z(\epsilon) \leq \frac{1}{N}\epsilon z(\epsilon),$$

which implies that $(\underline{m} - \underline{b})\frac{N-1}{N} \leq \epsilon$. But this has to hold for arbitrarily small ϵ. Thus, $\underline{b} \geq \underline{m}$.

As a result, it must be that

$$H(b|\underline{m}) = \begin{cases} 1 \text{ if } b \geq \underline{m}; \\ 0 \text{ otherwise.} \end{cases}$$

The reason is that the weight on $H(b|\underline{m})$ in (21) is non-positive, so that increasing $H(b|\underline{m})$ always weakly decreases the left-hand side and increases the right-hand side. Incidentally, this argument also implies that $H(b|m) = 1$ for $b \geq m$, for all $m \in M$.

Now, consider a solution to the relaxed program such that there exists a non-null set $X \subset \mathbb{R}$ such that for all $b \in X$, $H(b|m) < 1$ for some $m \in M$ and (21) is slack. Let

$$G(b) = \frac{1}{N} \sum_{m \in M} p^L(m) \left(\int_{x=0}^{b} H(x|m)dx - (N-1)(m-b)H(b|m) \right)$$

denote the slack in the constraint at b, so that $G(b) > 0$ on X. Then we can define an alternative solution:

$$\tilde{H}(b|m) = H(b|m) + \begin{cases} 0 \text{ if } b \notin X \text{ or if } b \geq m; \\ \min\left\{1 - H(b|m), \frac{G(b)}{(N-1)(m-b)}\right\} \text{ otherwise.} \end{cases}$$

Thus, it must be that $\tilde{H}(b|m) \geq H(b|m)$, and for a non-null set of $b \in X$, it must be that $\tilde{H}(b|m) > H(b|m)$. Hence, we have that

$$\begin{aligned} \sum_{m \in M} p^L(v^L) \frac{N-1}{N}(m-b)\tilde{H}(b|m) &\leq \sum_{m \in M} p^L(m) \frac{1}{N} \left[(n-1)(m-b)H(b|m) + G(b)\right] \\ &= \sum_{m \in M} p^L(m) \frac{1}{N} H(x|m)dx \\ &\leq \sum_{m \in M} p^L(m) \frac{1}{N} \tilde{H}(x|m)dx, \end{aligned}$$

so that $\tilde{H}(b|m)$ is feasible. However, since $\tilde{H}(b|m) > H(b|m)$ on a non-null set, we must have that (19) is higher with $\tilde{H}$ than with H, so the objective has improved. Thus, if (21) is slack on a non-null set such that $H(b|m) < 1$ for some $m \in M$, then (19) must be strictly below its optimal value.

Now, let us argue that (22) must be satisfied at the optimum. Suppose it is not the case, and let X be a non-null set such that $H(b|m) > 0$ while $H(b|m') < 1$ for some $m' > m$. Let

$$\phi(b) = \min\left\{p^L(m')\frac{1 - H(b|m')}{2}, p^L(m)H(b|m)\right\}.$$

Note that $\phi(b) > 0$ on x. We can then define an alternative solution

$$\tilde{H}(b|m'') = \begin{cases} H(b|m'') \text{ if } m'' \notin \{m, m'\}; \\ H(b|m') + \frac{1}{p^L(m')}\phi(b) \text{ if } m'' = m'; \\ H(b|m) - \frac{1}{p^L(m)}\phi(b) \text{ if } m'' = m. \end{cases}$$

From the definition of ϕ, we know that $\tilde{H}(b|m'') \in [0,1]$ for all $b \in \mathbb{R}_+$ and for all $m'' \in M$. In addition, whenever $\phi(b) > 0$,

$$\begin{aligned}
\sum_{m'' \in M} p^L(m'') \frac{N-1}{N}(m''-b)\tilde{H}(b|m'') &= \sum_{m'' \in M} p^L(m'') \frac{N-1}{N}(m''-b)H(b|m'') \\
&\quad + \frac{N-1}{N}\phi(b)(m'-m) \\
&< \sum_{m'' \in M} p^L(m'') \frac{N-1}{N}(m''-b)H(b|m'') \\
&\leq \sum_{m'' \in M} p^L(m'') \frac{1}{N} \int_{x=0}^{b} H(x|m'')dx \\
&= \sum_{m'' \in M} p^L(m'') \frac{1}{N} \int_{x=0}^{b} \tilde{H}(x|m'')dx,
\end{aligned}$$

and

$$\sum_{m'' \in M} p^L(m'')H(b|m'') = \sum_{m'' \in M} p^L(m'')\tilde{H}(b|m'')$$

for all b. Thus, we conclude that (21) is slack on a non-null set even though $\tilde{H}(b|m'') < 1$ for some m'', and therefore (19) must be strictly below the optimum at $\tilde{H}$. But (19) is the same at $\tilde{H}$ and at H. Thus, H cannot be the optimal solution either. We conclude that (22) must hold almost surely.

We are essentially done. $H(b|m)$ is now inductively pinned down by (21) and (22). In particular, suppose that we have defined $H(b|m')$ for all $m' < m \in M$. Let $\widehat{b} = \sup\{b | H(b|m') < 1 \text{ for some } m' < m\}$. Then it must be that $H(\widehat{b}|m') < 1$ for all $m' < m$ and $x \leq \widehat{b}$, and thus

$$\begin{aligned}
&(N-1)(m-b)H(b|m) - \int_{x=0}^{b} H(x|m)dx \\
&= \frac{1}{p^L(m)} \left[\sum_{m' < m} p^L(m') \left(\int_{x=0}^{\widehat{b}} H(x|m')dx - (N-1)(m'-b) \right) + (b - \widehat{b}) \right] \\
&= C_1 + C_2 b.
\end{aligned}$$

This is a first-order linear ordinary differential equation. The homogenous solutions would be of the form $(m-b)^{\frac{1}{N-1}}$, and using this, one can derive the non-homogenous solution

$$\int_{x=-\infty}^{b} H(x|m)dx = C_3(m-b)^{-\frac{1}{N-1}} - C_1 - C_2 \frac{(N-1)m+b}{N},$$

and thus,

$$H(b|m) = \frac{C_3}{N-1}(m-b)^{-\frac{N-2}{N-1}} - \frac{C_2}{N}.$$

The constant C_3 has to be chosen so that $H(\widehat{b}|m) = 0$, so

$$C_3 = C_2 \frac{N-1}{N}(m - \widehat{b})^{\frac{N-2}{N-1}}.$$

Our final form for $H(b|m)$ is

$$H(b|v) = \frac{1}{N}C_2\left(\left(\frac{m-\widehat{b}}{m-b}\right)^{1+\frac{1}{N-1}} - 1\right). \tag{50}$$

Strictly speaking, $H(b|m)$ is zero for $b < \widehat{b}$, given by (50) for $b \in [\widehat{b}, \bar{b}]$ where $\bar{b}$ is the point at which $H(b|m)$ hits 1, and then is 1 for $b > \bar{b}$. Clearly, as $b \uparrow m$, the right-hand side of (50) blows up, so that $H(b|m)$ must hit 1 before b hits m. Now we can inductively continue the solution for the next higher $m \in M$, and since M has are only finitely many elements, this process eventually terminates and we have defined the solution. We note that the functions $H(b|m)$ so defined are monotonically increasing in b, so that they are in fact CDFs. $\square$

Proof of Lemma 6. If $m = \underline{m} = \min M$, then we define $L(b|b', m)$ to randomize over an interval$[\underline{m} - \epsilon, \underline{m}]$ so as to support bidding at $\underline{m}$ by a type with expected valuation $v^W(\underline{m})$. In particular, we can have

$$L(b|\underline{m}, \underline{m}) = \begin{cases} 1 \text{ if } b \geq \underline{m}; \\ \frac{v^W(\underline{m}) - m}{v^W(\underline{m}) - b} \text{ if } b \in [\underline{m} - \epsilon, \underline{m}]; \\ 0 \text{ otherwise.} \end{cases}$$

Now, inductively suppose that $L(b|b', m')$ has been defined for $m' < m$, and satisfies the properties in the lemma. The derivative of the first term in (26) is

$$\sum_{m' \leq \tilde{m}} p^L(m')(m-b)H(db|m')$$

which is necessarily non-negative, since $b \leq m$ and $H(db|m')$ is positive from Proposition 5. And since we have inductively assumed that $L(b|x, m)$ is increasing for all $m' < m$, it must be that $L(b|b', m)$ must be increasing as well.

Now, let us argue that $L(b|b,m) = 1$. Since $L(b|b,m') = 1$ for all $m' < m$ and since $b > x$ for all $x \in B_{m'}$ with $m' < m$, (26) can be rewritten as

$$\begin{aligned}
&p^L(m)L(b|b,m)H(db|m) \\
&= \frac{1}{(m-b)^2}\Bigg[\frac{1}{N-1}\sum_{m'\leq m} p^L(m')\left(\int_{x=0}^{b} H(x|m')dx + (m-b)H(b|m')\right) \\
&\qquad + \sum_{m'<m}(m-m')p^L(m')\int_{x\in B_{m'}} H(dx|m')\Bigg]
\end{aligned}$$

which further rearranges to

$$\begin{aligned}
&p^L(m)(m-b)^2 L(b|b,m)H(db|m) \\
&= \frac{1}{N-1}\sum_{m'\leq m} p^L(m')\left(\int_{x=0}^{b} H(x|m')dx + (m-b)H(b|m')\right) \\
&\quad + \sum_{m'<m}(m-m')p^L(m')\int_{x\in B_{m'}} H(dx|m')
\end{aligned}$$

Since (21) holds with equality and the $H(b|m)$ are almost everywhere differentiable, it must be that

$$\sum_{m'\leq m} p^L(m')\frac{N-1}{N}(m'-b)H(db|m') = \sum_{m'\leq m} p^L(m')H(b|m'). \tag{51}$$

But $H(db|m') = 0$ for all $m' \neq m$, so in fact

$$p^L(m)\frac{N-1}{N}(m-b)H(db|m) = \sum_{m'\leq m} p^L(m')H(b|m').$$

We can substitute this and (21) into (26) to rewrite it as

$$\begin{aligned}
&L(b|b,m)(m-b)\frac{N}{N-1}\sum_{m'\leq m} p^L(m')H(b|m') \\
&= \frac{1}{N-1}\sum_{m'\leq m} p^L(m')\left((N-1)(m'-b)H(b|m') + (m-b)H(b|m')\right) \\
&\quad + \sum_{m'<m}(m-m')p^L(m')H(b|m')
\end{aligned}$$

or

$$\left(L(b|b,m)\frac{N}{N-1} - \frac{1}{N-1}\right)(m-b)\sum_{m'\leq m} p^L(m')H(b|m')$$
$$= \sum_{m'\leq m} p^L(m')(m'-b)H(b|m') + \sum_{m'<m}(m-m')p^L(m')H(b|m')$$
$$= \sum_{m'\leq m} p^L(m')(m-b)H(b|m')$$

which implies that $L(b|b,m) = 1$.

Now let us show (ii). (25) clearly holds as an equality if $b' = b$. We will show that the derivative of the left-hand side is zero, so that it will hold as an equality for all $b' > b$ as well. Now, if $b' \in B_m$, then $H(db'|m')$ is zero except for $m' = m$. In addition, if $b \in B_{\tilde{m}}$, then $H(db|m) = 0$ unless $m = \tilde{m}$. Thus, we can decompose the left-hand side into

$$(N-1)p^L(m)(m-b')\int_{x=-\infty}^{b'} L(db|x,m)H(dx|m)$$
$$+ (N-1)\sum_{m'<m}(m'-b')p^L(m')\int_{x=0}^{b'} L(db|x,m')H(dx|m') - (b'-b)p^L(\tilde{m})H(db|\tilde{m})$$

Now, it must be the case that $b' < m$, so we can divide this equation by $(N-1)(m-b')$ to obtain

$$p^L(m)\int_{x=-\infty}^{b'} L(db|x,m)H(dx|m)$$
$$+ \sum_{m'<m}\frac{m'-b'}{m-b'}p^L(m')\int_{x=0}^{b'} L(db|x,m')H(dx|m') - \frac{b'-b}{m-b'}p^L(\tilde{m})H(db|\tilde{m})$$

The incentive constraint is satisfied if and only if this expression is non-positive. Differentiating with respect to b', we obtain

$$p^L(m)L(db|b',m)H(db'|m)$$
$$- \frac{1}{(m-b')^2}\left[\sum_{m'<m}(m-m')p^L(m')\int_{x=0}^{b'} L(db|x,m')H(dx|m') - (m-b)p^L(\tilde{m})H(db|\tilde{m})\right]$$

which we can conclude is zero from differentiating (26).

Now consider the payoff to a bidder who is told to bid $b' \in B_m$ and deviates down to $b < b'$. Let $v^W(m)$ be the expected value of a buyer who wins when the average losing value is m, i.e.,

$$v^W(m) = \frac{1}{p^L(m)} \sum_{\{v \in V^N | \mu(v)=m\}} p(v) \max v.$$

Then the payoff of a buyer who deviates down is

$$\frac{1}{N} p^L(m)(v^W(m) - b)L(b|b', m)^{N-1} H(db'|m),$$

where the $N-1$ exponent comes from the fact that there are $N-1$ losing bidders who receive independent draws from this distribution. This expression is increasing in b if and only if

$$(v^W(m) - b)(N-1)L(db|b', m) - L(b|b', m) \geq 0$$

for all $b \in [0, b']$. Using (26), a sufficient condition for this to be the case is that the following functions are weakly increasing:

$$(v^W(m) - b)\left(\sum_{m' \leq \tilde{m}} p^L(m) \int_{x=0}^{b} H(x|m')dx\right)^{N-1};$$
$$(v^W(m) - b)L(b|x, m')^{N-1} \forall b \leq x \leq b', m' < m.$$

Again, if we inductively suppose that $(v^W(m') - b)L(b|x, m')^{N-1}$ is increasing for $m' < m$, then since $v^W(m) > v^W(m')$, we have that $(v^W(m) - b)L(b|x, m')^{N-1}$ is increasing as well. Let us then argue that the first of these two functions is increasing. Again, a sufficient condition is that

$$(v^W(m) - b)(N-1) \sum_{m' \leq \tilde{m}} p^L(m')H(b|m') - \sum_{m' \leq \tilde{v}} p^L(m') \int_{x=0}^{b} H(x|m')dx \geq 0.$$

But since $v^W(m) \geq m$, this is implied by (21). This verifies that downward deviations are not profitable, and therefore we have constructed a BCE. $\square$

Proof of Lemma 7. Let $g(b, v)$ be exchangeable in $b \in \mathbb{R}^N_+$ and in $v \in V^N$, i.e., $g(b, v) = g(\xi(b), \xi(v))$ for all $\xi \in \Xi$. This property is satisfied by revenue, bidder surplus, and total surplus. For example, in the case of revenue, $g(b, v) = \sum_{i=1}^n q_i(b)b_i$, which is exchangeable since $q_{\xi(i)}(\xi(b)) = q_i(b)$. Consider a BCE given by the family of conditional bid distributions

$\{F(db|v)\}_{v\in V^N}$ which attains an expected objective of

$$x = \sum_{v\in V^N} p(v) \int_{b\in\mathbb{R}_+^N} g(b,v)F(db|v).$$

Then we claim that there is an exchangeable BCE that also attains x, which is defined by

$$\tilde{F}(X|v) = \frac{1}{N!}\sum_{\xi\in\Xi} F(\xi(X)|\xi(v))$$

where $X \subseteq \mathbb{R}_+^N$ is any measurable set of bid profiles (note that the function ξ is measurable). To see this, note that

$$\begin{aligned}
\sum_{v\in V^N} p(v) \int_{b\in\mathbb{R}_+^N} g(b,v)\tilde{F}(db|v) &= \sum_{v\in V^N} p(v) \int_{b\in\mathbb{R}_+^N} g(b,v)\frac{1}{N!}\sum_{\xi\in\Xi} F(d\xi(b)|\xi(v)) \\
&= \frac{1}{N!}\sum_{\xi\in\Xi}\sum_{v\in V^N} p(\xi(v)) \int_{b\in\mathbb{R}_+^N} g(\xi(b),\xi(v))F(d\xi(b)|\xi(v)) \\
&= \frac{1}{N!}\sum_{\xi\in\Xi}\sum_{v\in V^N} p(v) \int_{b\in\mathbb{R}_+^N} g(b,v)F(db|v) \\
&= \frac{1}{N!}\sum_{\xi\in\Xi} x = x
\end{aligned}$$

The first line is just exchanging the order of the summation and integration. The second line follows from the fact that p and g are exchangeable. The third line comes from the fact that ξ is a bijection, so that we can equivalently sum over V^N and compose with ξ, or sum over $\xi(V^N)$ without the composition, which is still just V^N. $\square$

Proof of Proposition 9. The derivative of (27) with respect to m is

$$\beta'(m) = \frac{N-1}{N}(m-\beta(m))\frac{p^L(m)}{P^L(m)}. \tag{52}$$

It is easy enough to verify that $\beta(m)$ defined by (28) satisfies this differential equation, with the initial condition $\beta(\underline{v}) = \underline{v}$. For L'Hospital's rule implies that

$$\lim_{m\downarrow\underline{v}} \beta(m) = \frac{N-1}{N}\lim_{m\downarrow\underline{v}} \frac{\frac{mp^L(m)}{(P^L(m))^{\frac{1}{N}}}}{\frac{N-1}{N}\frac{p^L(m)}{(P^L(m))^{\frac{1}{N}}}} = \underline{v}.$$

Now, let us calculate the gains from deviating after a signal of s. If such a bidder were to deviate downwards to $\beta(s')$ with $s' < s$, they win only if (1) the deviator had the high value, since otherwise some other bidder's signal was greater than s, and (2) all others' signals are less than s', which occurs with probability $L(s'|s)$. Since others bids are independent draws, this occurs with probability $(L(s'|s))^{N-1}$. Thus, the surplus from deviating downwards is proportional to

$$(v^W(s) - \beta(s'))(L(s'|s))^{N-1},$$

where $v^W(s) \geq s$ is the expected high valuation given that the average loser's valuation is s. Thus, a sufficient condition for bidders to not want to deviate downwards is that

$$(v^W(s) - \beta(s'))(N-1)(L(s'|s))^{N-2}L(ds'|s) - \beta'(s')(L(s'|s))^{N-1} \geq 0.$$

A sufficient condition for this to be the case is that

$$(s - \beta(s'))(N-1)L(ds'|s) - \beta'(s')L(s'|s) \geq 0.$$

Plugging in the definition of L and the formula for $\beta'(s')$, we obtain

$$\frac{N-1}{N}\frac{1}{(P^L(s))^{\frac{1}{N}}}\frac{p^L(s')}{(P^L(s'))^{\frac{N-1}{N}}}(s-s') \geq 0,$$

thus verifying that bidders will not want to deviate downwards.

Now let us consider upward deviations. Bidders think that their signal s is the average losing value m with likelihood $p^L(s)/N$. The upward deviator therefore loses surplus proportional to

$$(\beta(s') - \beta(s))\frac{p^L(s)}{n}$$

from deviating up to $\beta(s')$. On the other hand, the likelihood of the signal $s < m$ is $L(ds|m)\frac{N-1}{N}p^L(m)$, and conditional on winning when the average losing value is m, the upward deviator expects to gain $m - \beta(s')$. Integrated over all $m < s'$, this is

$$\int_{m=s}^{s'}(m - \beta(s'))L(ds|m)\frac{N-1}{N}p^L(m)dm = \frac{N-1}{N}\frac{1}{N}\frac{p^L(s)}{(P^L(s))^{\frac{N-1}{N}}}\int_{m=s}^{s'}(m-\beta(s'))\frac{p^L(m)}{(P^L(m))^{\frac{1}{N}}}dm.$$

Thus, the total change in surplus is proportional to

$$\frac{N-1}{N}\frac{1}{(P^L(s))^{\frac{N-1}{N}}}\int_{m=s}^{s'}(m-\beta(s'))\frac{p^L(m)}{(P^L(m))^{\frac{1}{N}}}dm-(\beta(s')-\beta(s)),$$

which is clearly zero when $s'=s$. The derivative with respect to s' is

$$\frac{N-1}{N}\frac{1}{(P^L(s))^{\frac{N-1}{N}}}(s'-\beta(s'))\frac{p^L(s')}{(P^L(s'))^{\frac{1}{N}}}-\beta'(s')\frac{N-1}{N}\frac{1}{(P^L(s))^{\frac{N-1}{N}}}\int_{m=s}^{s'}\frac{p^L(m)}{(P^L(m))^{\frac{1}{N}}}dm-\beta'(s').$$

Again, using the differential equation for $\beta'(s')$, we can rewrite this quantity as

$$\beta'(s')\left[\left(\frac{P^L(s')}{P^L(s)}\right)^{\frac{N-1}{N}}-\frac{N-1}{N}\frac{1}{(P^L(s))^{\frac{N-1}{N}}}\int_{m=s}^{s'}\frac{p^L(m)}{(P^L(m))^{\frac{1}{N}}}dm-1\right].$$

By solving out the integral, we obtain

$$\beta'(s')\left[\left(\frac{P^L(s')}{P^L(s)}\right)^{\frac{N-1}{N}}-\frac{1}{(P^L(s))^{\frac{N-1}{N}}}\left((P^L(s'))^{\frac{N-1}{N}}-(P^L(s))^{\frac{N-1}{N}}\right)-1\right]=0.$$

Thus, bidders are indifferent to deviating up to any bid in the support of winning bids. □

Proof of Proposition 10. The proof is similar to that of Proposition 5. The objective is weak-* continuous and the feasible set is weak-* compact, by analogous arguments, so an optimum exists. We first argue that no buyer can bid below v^1. Then, we argue that (33) must bind whenever $H(b|z)<1$ for some z, or else there is another solution that strictly improves the objective and lowers revenue. Finally, we show that the ordered supports property (34) must be satisfied, or else it is possible to find an alternative solution for which revenue is the same but also for which (33) is slack even though $H(b|z)<1$ for some z.

We can write

$$H(b)=\sum_{z=0}^{N}p^Z(z)H(b|z)$$

for the ex-ante cumulative distribution of the winning bid. Because $v^2>v^1$, (33) implies that $H(b)$ satisfies the following inequality:

$$(v^1-b)H(b)\le\int_{x=0}^{b}H(x)dx.$$

Since the support of bids is bounded below, there is an infimum b such that $H(b) > 0$, which we denote by $\underline{b}$. Note that for all $b \in [\underline{b}, \underline{b}+\epsilon]$, $H(b) \leq H(\underline{b}+\epsilon)$, so

$$(v^1 - b)H(\underline{b}+\epsilon) \leq \epsilon H(\underline{b}+\epsilon)$$

which implies that $v^1 - b \leq \epsilon$ for all $\epsilon > 0$. Thus, $b \geq v^1$, and hence $\underline{b} \geq v^1$.

Together with the assumption of weakly undominated strategies, this implies that $H(b|0)$ is a mass point on v^1. In addition, it must be that $H(b|1)$ is also a mass point on v^1. The reason is that the weight on $H(b|1)$ on the left-hand side of (33) is zero, so increasing $H(b|1)$ only increase the right-hand side of (33), as well as increasing (31).

Now we claim that if (33) is not satisfied almost surely when $H(b|z) < 1$ for some z, then it is possible to strictly improve the objective and decrease revenue. The proof is as in the unknown values case. Let

$$G(b) = \frac{1}{\sum_{z=1}^N p^Z(z)} \sum_{z=1}^N p^Z(z) \left[\frac{1}{z} \int_{x=0}^b H(x|z)dx - (v^2 - b)\frac{z-1}{z} H(b|z) \right]$$

denote the slack in the constraint at b. Suppose there is a non-null set X for which $G(b) > 0$ and $H(b|z) < 1$ for some z. Then we can define the alternative solution

$$\tilde{H}(b|z) = H(b|z) + \begin{cases} 0 \text{ if } b \notin x; \\ \min\left\{1 - H(b|z), \frac{G(b)}{v^2-b}\frac{z}{z-1}\right\} \text{ otherwise} \end{cases}$$

Note that $\tilde{H}(b|z) > H(b|z)$ on a non-null set. Thus, the objective (31) must be larger with $\tilde{H}$ than under H. In addition,

$$\begin{aligned}
(v^2 - b)\sum_{z=1}^N p^Z(z)\frac{z-1}{z}\tilde{H}(b|z) &\leq (v^2 - b)\sum_{z=1}^N p^Z(z)\frac{z-1}{z}H(b|z) + \sum_{z=1}^N p^Z(z)G(b) \\
&= \sum_{z=1}^N p^Z(z)\frac{1}{z}\int_{x=0}^b H(x|z)dx \\
&\leq \sum_{z=1}^N p^Z(z)\frac{1}{z}\int_{x=0}^b \tilde{H}(x|z)dx,
\end{aligned}$$

so that the new solution is feasible as well.

Now let us show that (34) must be satisfied almost surely. Suppose that X is some non-null set on which $H(b|z) < 1$ but $H(b|z') > 0$ for some $z' > z$. Let

$$\phi(b) = \min\left\{p^Z(z)\frac{1-H(b|z)}{2}, p^Z(z')H(b|z')\right\}.$$

As before, $\phi(b)$ must be strictly positive on X. Now define the perturbed solution

$$\tilde{H}(b|z'') = \begin{cases} H(b|z'') \text{ if } z'' \notin \{z, z'\}; \\ H(b|z) + \frac{1}{p^Z(z)}\phi(b) \text{ if } z'' = z; \\ H(b|z') - \frac{1}{p^Z(z')}\phi(b) \text{ if } z'' = z'. \end{cases}$$

We claim that $\tilde{H}$ is feasible and induces weakly higher objective. Observe that the right-hand side of (33) has increased, since

$$\begin{aligned}
&\sum_{z''=0}^{N} \frac{p^Z(z'')}{z''} \int_{x=0}^{b} (\tilde{H}(x|z'') - H(x|z))dx \\
&= \frac{p^Z(z)}{z} \int_{x=0}^{b} (\tilde{H}(x|z) - H(x|z))dx + \frac{p^Z(z')}{z'} \int_{x=0}^{b} (\tilde{H}(x|z') - H(x|z'))dx \\
&= \left(\frac{1}{z} - \frac{1}{z'}\right) \int_{x=0}^{b} \phi(x)dx > 0.
\end{aligned}$$

On the other hand, the change in the left-hand side is

$$\begin{aligned}
&(v^2 - b) \sum_{z''=1}^{N} p^Z(z'')\frac{z''-1}{z''}(\tilde{H}(b|z'') - H(b|z'')) \\
&= (v^2 - b)\left[p^Z(z)\frac{z-1}{z}\left(\tilde{H}(b|z) - H(b|z)\right) + p^Z(z')\frac{z'-1}{z'}\left(\tilde{H}(b|z') - H(b|z')\right)\right] \\
&= (v^2 - b)\phi(b)\left(\frac{z-1}{z} - \frac{z'-1}{z'}\right) < 0.
\end{aligned}$$

Since (33) was weakly satisfied under H, it must be strictly satisfied for a positive measure of b such that $\phi(b) > 0$, and moreover it is clear that $H(b|z) < 1$ for all such b. Thus, it is possible to further perturb the solution as we did above in a manner that strictly decreases expected revenue.

Having established that $H(b|0)$ and $H(b|1)$ are mass points on 0, it must be that there is an interval $[v^1, b^2]$ for which (33) is solved with $H(b|1) = 1$ and with $H(b|z) = 0$ for all

$z > 2$. In particular,

$$(v^2 - b)\frac{p^Z(2)}{2}H(b|2) = \int_{x=v^1}^{b}\left[p^Z(1) + \frac{p^Z(2)}{2}H(x|2)\right]dx.$$

The solution of this differential equation is

$$H(b|2) = \frac{2p^Z(1)}{p^Z(2)}\frac{v^2(b - v^1) - \frac{b^2}{2}}{(v^2 - b)^2},$$

which explodes as $b \uparrow v^2$. Thus, it must hit 1/2 before b reaches v^2, and at that point (b^2) we change over to solving (33) for equality with $H(b|3)$.

Inductively, there will be ranges $B_z = [b^{z-1}, b^z]$ over which (33) is solved as an equality with $H(b|z') = 1$ for $z' < z$ and $H(b|z') = 0$ for $z' > z$. The differential equation is:

$$\begin{aligned}(v^2 - b)&\left[\sum_{z'=1}^{z-1} p^Z(z')\frac{z'-1}{z'} + p^Z(z)\frac{z-1}{z}H(b|z)\right] \\ &= \sum_{z'=1}^{z-1}\frac{p^Z(z')}{z'}\left[\int_{x=0}^{b^{z'}} H(x|z')dx + (b - b^{z'})\right] \\ &+ \frac{p^Z(z)}{z}\int_{x=b^{z-1}}^{b} H(x|z)dx.\end{aligned}$$

All of the terms corresponding to $z' < z$ can be treated as constants, so that this differential equation can be rewritten as

$$(v^2 - b)\left[C_1^z + \frac{z-1}{z}H(b|z)\right] = C_2^z + \frac{1}{z}\int_{x=b^{z-1}}^{b} H(x|z)dx,$$

which rearranges to

$$(v^2 - b)\frac{z-1}{z}H(b|z) - \frac{1}{z}\int_{x=b^{z-1}}^{b} H(x|z)dx = C_2^z - C_1^z v^2 + C_1^z b.$$

The solution to the differential equation is of the form

$$\int_{x=b^{z-1}}^{b} H(x|z)dx = C_3^z\left(v^2 - b\right)^{-\frac{1}{z-1}} - C_1^z b + C_4^z.$$

Thus,

$$H(b|z) = C_3^z\frac{1}{z-1}(v^2 - b)^{-\frac{z}{z-1}} - C_1^z.$$

The coefficient C_3^z is chosen so that $H(b^{z-1}|z) = 0$, so

$$C_3^m = (z-1)C_1^z(v^2 - b^{z-1})^{\frac{z}{z-1}}.$$

Clearly, $H(b|z)$ is blowing up as $b \uparrow v^2$, so given that $b^{z-1} < v^2$, we must have $H(b|z)$ hit 1 at $b^z < v^2$ as well. This completes the construction of the solution to the relaxed program. $\square$

Proof of Theorem 11. Recall that the losing bid distributions are defined by

$$p^Z(z)\frac{z-1}{z}L(b|b',z)H(db'|z) = \frac{1}{(v^2-b')^2}\sum_{z'=1}^{\tilde{z}}\frac{p^Z(z')}{z'}\int_{x=0}^{b}(v^2-x)H(dx|z').$$

Clearly, these functions are monotonically increasing in b and are zero at $b < v^1$, and if $b = b'$ and $z = \tilde{z}$, then this formula reduces through integration by parts to

$$\begin{aligned} p^Z(z)\frac{z-1}{z}L(b|b,z)H(db|z) &= \frac{1}{(v^2-b)^2}\sum_{z'=1}^{z}\frac{p^Z(z')}{z'}\int_{x=0}^{b}(v^2-x)H(dx|z') \\ &= \frac{1}{(v^2-b)^2}\sum_{z'=1}^{z}\frac{p^Z(z')}{z'}\left[\int_{x=0}^{b}H(x|z')dx + (v^2-b)H(b|z')\right]. \end{aligned}$$

Substituting in (33), this further reduces to

$$p^Z(z)\frac{z-1}{z}L(b|b,m)H(db|m) = \frac{1}{v^2-b}\sum_{z'=1}^{z}p^Z(z')H(b|z').$$

Since (33) must hold as an equality, we can differentiate both sides to obtain

$$(v^2-b)p^Z(z)\frac{z-1}{z}H(db|z) = \sum_{z'=1}^{z}p^Z(z')H(b|z'),$$

where z is the almost-surely unique z' such that $b \in [b^{z'-1}, b^{z'}]$. Thus, we conclude that $L(b|b,m)$ must be 1.

Now let us verify that incentive constraints are satisfied. We will use the same approach as before, showing that incentive compatibility is met almost surely with respect to the conditional distributions of values and bids given a bidder's own bid. In particular, a downward deviation from the recommended bid $b' \in [b^{z-1}, b^z]$ to some $b < b'$ is suboptimal if

$$(v^2-b)L(b|b',z)^{z-1} \tag{53}$$

is increasing in b. $L(b|b', z)$ is continuous and almost everywhere differentiable, so a sufficient condition for (53) to be increasing is that

$$(v^1 - b)(z-1)L(db|b', z) \geq L(b|b', z)$$

wherever $L(b|b', z)$ is differentiable. Up to a constant that does not depend on b,

$$L(b|b', z) \propto \sum_{z'=1}^{\tilde{z}} \frac{p^Z(z')}{z'} \int_{x=0}^{b} (v^2 - x) H(dx|z'),$$

where $b \in B_{\tilde{z}}$, so that the increasing condition is equivalent to

$$(v^2 - b)^2 \frac{z-1}{\tilde{z}} p^Z(\tilde{z}) H(db|\tilde{z}) \geq \sum_{z'=1}^{\tilde{z}} \frac{p^Z(z')}{z'} \int_{x=0}^{b} (v^2 - x) H(dx|z').$$

Since $z \geq \tilde{z}$, the left-hand side is at least

$$(v^2 - b)^2 \frac{\tilde{z}-1}{\tilde{z}} p^Z(\tilde{z}) H(db|\tilde{z}).$$

Again, from the differential form of (33), this quantity can be rewritten as

$$\begin{aligned}
(v^2 - b) \sum_{z'=1}^{\tilde{z}} p^Z(z') H(b|z') &= \sum_{z'=1}^{\tilde{z}} p^Z(z') \left[(v^2 - b) \frac{z'-1}{z'} H(b|z') - (v^2 - b) \frac{1}{z'} (v^2 - b) H(b|z') \right] \\
&= \sum_{z'=1}^{\tilde{z}} \frac{p^Z(z')}{z'} \left[\int_{x=0}^{b} H(x|z') dx + (v^2 - b) H(b|z') \right] \\
&= \sum_{z'=1}^{\tilde{z}} \frac{p^Z(z')}{z'} \int_{x=0}^{b} (v^2 - x) H(dx|z'),
\end{aligned}$$

where the second line follows from (33). Thus, downward deviations are not attractive.

Now, let us see that it is not optimal to deviate upwards from a recommendation of $b \in B_z$ to a bid of $b' > b$. With probability $\frac{p^Z(z)}{z} H(db|z)$, b is a winning recommendation and the upward deviator is losing surplus from paying more when they would have won anyway. With probability $\sum_{z'=z}^{N} p^Z(z') \frac{z'-1}{z'} \int_{x=b}^{b'} L(db|x, z') H(dx|z')$, the bid b is a losing recommendation and the winner was told to bid some $x \in [b, b']$. The incentive constraint is therefore that

$$(v^2 - b') \sum_{z'=z}^{N} p^Z(z') \frac{z'-1}{z'} \int_{x=b}^{b'} L(db|x, z') H(dx|z') \leq (b' - b) \frac{p^Z(z)}{z} H(db|z).$$

We will argue that the two sides are always equal, or rather that

$$\sum_{z'=z}^{n} p^Z(z')\frac{z'-1}{z'}\int_{x=b}^{b'} L(db|x,z')H(dx|z') = \frac{b'-b}{v^2-b'}\frac{p^Z(z)}{z}H(db|z).$$

Clearly the two are equal (and zero) at $b' = b$, and we will argue that the derivatives of both sides with respect to b' are always equal. This requires that

$$p^Z(\tilde{z})\frac{\tilde{z}-1}{\tilde{z}}L(db|b',\tilde{z})H(db'|\tilde{z}) = \frac{1}{(v^2-b')^2}\frac{p^Z(z)}{z}H(db|z).$$

where $b' \in B_{\tilde{z}}$. But of course, this is precisely the derivative of (35) with respect to b, so we are done.

Finally, it is obvious that the winning bid distribution induced by this equilibrium is equal to the solution to the binary known values relaxed program, and the latter generates a lower bound on revenue over all binary known value BCE. Thus, the BCE constructed above attains the lower bound on revenue. The allocation is also efficient, and so the BCE attains an upper bound on total bidder surplus as well. □

B Examples

B.1 Raising Revenue

Consider a small variation on the complete information benchmark. Bidders receive signals $t_i = (v_i, x_i)$ where $x_i \in [0,1]$ is determined according to the following procedure. Let us first consider how information might induce an outcome in which revenue is higher than the benchmark and bidder surplus is lower. We start close to complete information, and suppose that the buyers receive signals that indicate which of them has the higher valuation. In particular, we assume that the buyers' signals contain the maximum value.

In addition, let us suppose that rather than learning the losing value, the winning bidder receives a noisy signal. Fix $\epsilon \in [0,1/2]$, and let us suppose that the bidders observe signals $t_i = (v_i, v^{(1)}, \max\{\epsilon v^{(1)}, v^{(2)}\})$. Thus, if $v^{(2)} > \epsilon v^{(1)}$, the high value buyer learns the lowest value, but otherwise all he learns is that the lowest value is less than ϵ of his own value.

We can construct a simple equilibrium in this case where if $v^{(2)} > \epsilon v^{(1)}$, the high value buyer bids $v^{(2)}$, and the low value buyer randomizes over bids below his value, say according to the cumulative distribution function $F(b|v^{(1)}, v^{(2)}) = \frac{v^{(1)}-v^{(2)}}{v^{(1)}-b}$ on $[0, v^{(2)}]$, and the high value buyer wins with a bid of $v^{(2)}$. On the other hand, if $v^{(2)} < \epsilon v^{(1)}$, the low value buyer simply bids his value and the winner bids $v^{(2)}$. This is incentive compatible because the winner thinks that the loser's bid is uniformly distributed on $[0, \epsilon v^{(1)}]$, so the surplus from a bid b is $(v^{(1)} - b)\frac{b}{\epsilon v^{(1)}}$, which is increasing as long as $b < v^{(1)}/2$.

With this information and in this equilibrium, the cumulative distribution of winning bids first-order stochastically dominates that under complete information, and so revenue must be strictly higher and bidder surplus strictly lower. Intuitively, under the complete information outcome, when $v^{(2)} = \epsilon v^{(1)}$, the high bidder strictly prefers to bid $v^{(2)}$ over any smaller bid. This bid would still be optimal even if there was a modest probability of $v^{(2)}$ and the losing bid being less than $\epsilon v^{(2)}$. Thus, by creating partial information about the losing bid and losing bidder's value, the winner could be induced to bid more than in the benchmark.

B.2 Lowering Revenue

The equilibrium is normally as with the complete information, with $x_i = v_j$, and bidder i bids x_i if $x_i < v_i$ and randomizes over $[0, v_i]$ if $x_i > v_i$ in order to support bidder j bidding v_i. However, there is a small probability that the following happens: If $v_i > v_j$, then with probability $\epsilon v_j(v_i - v_j)$, bidder i is given the signal $x_i = \alpha(v_i, v_j)$, which is strictly increasing in v_i and v_j and for fixed v_j, has range $[2v_j/3, v_j]$, and bidder j is given the signal $x_j = x_i/2$. In particular, we use the function:

$$\alpha(v_i, v_j) = \frac{v_j}{3}\left(2 + \frac{v_i - v_j}{1 - v_j}\right),$$

which linearly interpolates bids between $[2v_j/3, v_j]$ and assigns then to winning bidders v_i in order, for $v_i \in [v_j, 1]$.

In equilibrium, when $x_i > v_i$, bidder i randomizes over the interval $[\alpha(v_j, v_i), v_i]$ so as to make bidder j indifferent to bidding v_i, and otherwise bidders bid their signal: $b_i = x_i$.

Let us show that this is an equilibrium. If $x_i > v_i$, buyer i is indeed willing to randomize as intended, since the "small probability" event has not happened and $b_j = v_i$.

If $x_i < v_i$, then there are three possibilities:

(a) $x_i = v_j < v_i$.

(b) $x_i = \alpha(v_i, v_j)$ for some $v_j < v_i$.

(c) $x_i = \alpha(v_j, v_i)/2$ for some $v_j > v_i$.

Note that event (c) could only occur if $x_i \in [v_i/3, v_i/2]$. Let us assess the conditional probabilities of these three events. Since v_j is uniformly distributed, event (a) simply occurs with ex-ante probability $1 - \epsilon x_i(v_i - x_i)$. Event (b) occurs with ex-ante probability

$\epsilon\gamma(v_i, x_i)(v_i - \gamma(v_i, x_i))$ where $\gamma(v_i, x_i)$ is the solution to $\alpha(\gamma, v_i) = x_i$. Event (c) occurs with ex-ante probability $\epsilon v_i(\xi(v_i, x_i) - v_i)$, where $\xi(v_i, x_i)$ solves $\alpha(v_i, \xi) = 2x_i$.

Now, to verify that incentive constraints are satisfied, first observe that the surplus from following the recommendation is

$$(v_i - x_i)(\underbrace{1 - \epsilon x_i(v_i - x_i)}_{\text{(a)}} + \underbrace{\epsilon\gamma(v_i, x_i)(v_i - \gamma(v_i, x_i))}_{\text{(b)}} + \underbrace{0}_{\text{(c)}}).$$

Note that the bidder gets no surplus from event (c) because by following the equilibrium strategy, player j will win with a bid of $2x_i$. By deviating to a bid $b = x_i/2$ (the most attractive downward deviation), the bidder's surplus will be approximately

$$(v_i - x_i/2)\epsilon\gamma(v_i, x_i)(v_i - \gamma(v_i, x_i)),$$

which is less than the surplus from $b = x_i$ as long as

$$\epsilon \leq \frac{v_i - x_i}{\gamma(v_i, x_i)(v_i - \gamma(v_i, x_i))/2 + v_i - x_i} \frac{1}{x_i}.$$

Observing that $x_i \leq \gamma(v_i, x_i) \leq v_i$, it must be that $\gamma(v_i - \gamma) \leq v_i(v_i - x_i)$, so that the right-hand side is at least

$$\frac{1}{v_i/2 + v_i - x_i} \frac{1}{x_i}$$

which is at least $4/5$, so for any $\epsilon < 4/5$, a downward deviation is not attractive.

Similarly, deviating up to a bid slightly larger than $2x_i$ would allow the bidder to win on event (c), yielding a surplus of

$$(v_i - 2x_i)(1 - \epsilon x_i(v_i - x_i) + \epsilon\gamma(v_i, x_i)(v_i - \gamma(v_i, x_i)) + \epsilon v_i(\xi(v_i, x_i) - v_i)),$$

which is less than the surplus from bidding x_i if

$$\epsilon \leq \frac{x_i}{(v_i - 2x_i)v_i(\xi(v_i, x_i) - v_i) + x_i^2(v_i - x_i) - \gamma(v_i, x_i)(v_i - \gamma(v_i, x_i))}.$$

But since $x_i \geq v_i/3$, the right-hand side must be at least

$$\frac{1}{3} \frac{1}{(v_i - 2x_i)(\xi(v_i, x_i) - v_i) + x_i^2/3},$$

which is itself at least 1/2. Thus, choosing $\epsilon \leq 1/2$, both upward and downward incentive constraints will be satisfied.

In the end, it is incentive compatible to follow the "recommendation" x_i. Moreover, this equilibrium induces winning bids where $x_i < v^{(2)}$ with positive probability, so that revenue must be strictly lower. The outcome is, however, still efficient, so bidder surplus must have risen relative to the complete information case.

B.3 An Example with Uniformly Distributed Values

We now illustrate Theorem 2, both the result and the construction of the information structure with an example of two bidders with independent standard uniformly distributed values in $[0,1]$, thus $F_i(v_i) = v_i$ for $i = 1, 2$. The lower bound on bidder surplus for a bidder with valuation v_i when the competing bidder always bids his value is

$$\underline{u}_i(v_i) = \max_{b \in [0, v_i]} (v_i - b)b = \frac{v_i^2}{4}$$

as the optimal bidding strategy is $b_i^*(v_i) = v_i/2$. Thus, the ex-ante lower bound on bidder surplus is

$$\underline{U}_i = \int_{v=0}^{1} \underline{u}_i(v_i) dv_i = \int_{v=0}^{1} \frac{v_i^2}{4} dv_i = \frac{1}{12}.$$

The sum of the worst case bidder surplus is therefore 1/6, and since the efficient surplus is 2/3, the maximum revenue for the seller is $2/3 - 1/6 = 1/2$. This can be contrasted with the revenue under the conventional BNE – in which each bidder knows his value and maintains the common prior as his belief about his competitor – which is 1/3.

We can explicitly construct a BCE that attains the outcome $(\underline{U}, \overline{R}) = (1/6, 1/2)$. When values are (v_i, v_j), with $v_i > v_j$, the bidders receive recommendations in $[v_i/2, v_i] \times \{v_j\}$. We can easily adapt the construction of Theorem 2 to the case of continuous random variables. The indifference condition (6) identified the weight α on the *conditional distribution* of the second highest valuation conditional on it being below the bid b_i of the winner and the second highest valuation being equal to the bid b. In the case of the continuous random variable the indifference condition (6) simply becomes

$$(v_i - b_i)\left(\alpha \Pr\left(v^{(2)} < b_i \left| v^{(1)}\right.\right) + \left(1 - \alpha \Pr\left(v^{(2)} < b_i \left| v^{(1)}\right.\right)\right)\right) = (v_i - b_i^*(v_i))\, \alpha \Pr\left(v^{(2)} \leq b_i^*(v_i) \left| v^{(1)}\right.\right). \tag{54}$$

$$\tag{55}$$

With two bidders and the uniform distribution this condition can be written as

$$(v_i - b_i)\left(\alpha \frac{b_i}{v_i} + \left(1 - \alpha \frac{b_i}{v_i}\right)\right) = \left(v_i - \frac{1}{2} v_i\right) \alpha \frac{1}{2}$$

and yields

$$\alpha = \frac{4\,(v_i - b_i)}{v_i}.$$

With this information, we can immediately construct the conditional distribution of the loser's value given a bid recommendation b_i to a winner with valuation v_i, or:

$$G_{v_i,b_i}(v_j) = \begin{cases} \frac{4(v_i - b_i)}{v_i^2} v_j, & \text{if} \quad 0 \leq v_j \leq b_i; \\ 1, & \text{if} \quad b_i \leq v_j. \end{cases} \tag{56}$$

Note that there is a conditional mass point of size

$$1 - \frac{4(v_i - v_j)v_j}{v_i^2}$$

on $v_j = b_i$, given the recommendation b_i.We also define the distribution G_{v_i,v_i} as the pointwise limit of G_{v_i,b_i} as $b_i \uparrow v_i$, which puts probability one on $v_j = v_i$.

Next, we construct the distribution of bid recommendations b_i, given the loser's value v_j, so that the conditional distribution of the loser's value given a recommendation of b_i has the shape of G_{v_i,b_i}. As a result, it will always be optimal to bid b_i, and the bidder will be indifferent between bidding b_i and bidding $v_i/2$, which is the best response when others are bidding their values.

Let us write $F(b_i|v_i)$ for the marginal distribution of winning bid recommendations when the winner's value is v_i, which will have a continuously differentiable density. In order for the probabilities of the losing values v_j to add up, it must be that the prior density of valuations is equal to the expected interim density. For $v_j \in [v_i/2, v_i]$, this requirement can be written as

$$\int_{b_i = v_j}^{v_i} \frac{4(v_i - b_i)}{v_i^2} f(b_i|v_i) db_i + f(v_j|v_i)\left(1 - \frac{4(v_i - v_j)v_j}{v_i^2}\right) = \frac{1}{v_i},$$

Differentiating this expression with respect to v_j, we conclude that

$$-\frac{4(v_i - v_j)}{v_i^2} f(v_j|v_i) + \left(1 - \frac{4(v_i - v_j)v_j}{v_i^2}\right) f'(v_j|v_i) - f(v_j|v_i)\frac{4(v_i - 2v_j)}{v^2} = 0.$$

Finally, this rearranges to the differential equation

$$\frac{f'(v_j|v_i)}{f(v_j|v_i)} = \frac{4(v_i - 2v_j) + 4(v_i - v_j)}{(v_i - 2v_j)^3},$$

which admits simple solutions of the form

$$f(v_j|v_i) = \frac{C(v_i)}{(2v_j - v_i)^3} \exp\left(-\frac{v_i}{2v_j - v_i}\right).$$

The constant of integration has to be chosen so that the density of bid recommendations integrates to one:

$$1 = \int_{b_i = v_i/2}^{v_i} f(b_i|v_i) db_i = \frac{C(v_i)}{v_i^2 e} \Leftrightarrow C(v_i) = v_i^2 e.$$

The resulting conditional distribution function

$$F(b_i|v_i) = \frac{b_i}{2b_i - v_i} \exp\left(\frac{2v_i - 2b_i}{v_i - 2b_i}\right)$$

increases continuously from 0 to 1 as b_i increases from $v_i/2$ to v_i.

To complete the construction of the BCE, we simply have the losing bidder bid his or her value if the winning bid is strictly larger than the loser's value, and if the winner is told to bid the loser's value v_j, then the losing bidder randomizes over an interval, say $[v_j/2, v_j]$ according to the cumulative distribution $G(b_j|v_i, v_j) = (v_i - v_j)/(v_i - b_j)$, where v_i is the winner's value.

C Beyond Binary Values

Suppose that $V = \{v^1, v^2, v^3\}$ that $n = 2$, and that the distribution of values is independent. To condense notation, we write $H_{kk'}(b)$ for the cumulative distribution of winning bids when v^k wins against $v^{k'}$, and p_k will denote the independent symmetric prior $p(v^k)$. For example, $H_{23}(b)$ denotes the cumulative distribution of winning bids when the v^2 type wins against v^3 type. We restrict attention to efficient solutions in which $H_{23}(b) = 0$ for all b. As with binary values, whenever some bidder wins against v^1, they win at a bid of v^1. The relaxed program in this case is:

$$\max \sum_{k,k' \in \{1,2,3\}} p_k p_{k'} \int_{x=0}^{v^k} H_{kk'}(x) dx \tag{57}$$

subject to

$$H_{kk'}(b) + H_{kk'}(b) \in [0,1] \forall b \in \mathbb{R}, k, k' \in \{1,2,3\}; \tag{58}$$

and

$$(v^2 - b)\left(p_2 H_{22}(b) + p_3 H_{32}(b)\right) \leq \int_{x=v^1}^{b} \left(p_1 + p_2 H_{22}(x)\right) dx; \tag{59.1}$$

$$(v^3 - b) p_3 H_{33}(b) \leq \int_{x=v^1}^{b} \left(p_1 + p_2 H_{32}(x) + p_3 H_{33}(x)\right) dx. \tag{59.2}$$

Our usual arguments imply that at a solution to the relaxed program, (59.1) must be binding whenever either $H_{22}(b) < 1/2$ or $H_{32}(b) < 1$, and (59.2) must bind whenever $H_{33}(b) < 1/2$. If not, we could increase one of these controls while maintaining feasibility and pushing down the winning bid distribution. Note that this pins down the path for $H_{33}(b)$ given the path of $H_{32}(b)$. However, while we know that (59.1) is binding, we do not know the order in which $H_{22}(b)$ and $H_{32}(b)$ should rise. Let us consider the tradeoffs. At a given b, there is a certain amount of slack in the right-hand side of (59.1) which could be allocated either to $H_{22}(b)$ or to $H_{32}(b)$. If that space is allocated to $H_{22}(b)$, then the slack on the right-hand side of (59.1) increases more quickly, thereby allowing both $H_{33}(b)$ and $H_{32}(b)$ to increase faster. Alternatively, if the slack is allocated to $H_{32}(b)$, then the right-hand side of (59.2) increases faster so that $H_{33}(b)$ rises faster. In economic terms, the more v^2 loses to v^2 with low bids, the faster the distribution of bids that win against v^2 will rise. The more v^2 loses to v^3 at low bids, the faster the distribution bids that win against v^3 will rise.

The resolution of this tradeoff turns out to depend very much on the parameters of the model. The reason is that benefits of relaxing each of the incentive constraints only accrue while that constraint binds. Thus, if one of the two constraints is binding for a wider range of bids, then there is more benefit to relaxing that inequality. For example, if v^3 is much larger than v^2, then it will take much longer for $H_{33}(b)$ to hit its upper bound than $p_2 H_{22}(b) + p_3 H_{32}(b)$. By increasing H_{32} faster early on, there are large gains in reducing H_{33}. On the other hand, if the parameters are such that $H_{33}(b)$ hits its upper bound much faster than $p_2 H_{22}(b) + p_3 H_{32}(b)$, for example when p_3 is very small relative to p_2, then the opposite intuition holds, that it is better to have v^2 win against v^2 with low bids.

For this latter case, we can provide an exact characterization of the solution to the relaxed program. In particular, if p_2 is sufficiently large relative to p_3, then the optimal solution should satisfy the following ordered supports property:

$$H_{22}(b) < 1/2 \implies H_{32}(b) = 0. \tag{60}$$

We formalize this in the following proposition:

Proposition 12. *If p_2/p_3 is sufficiently large, the solution to the efficiency-constrained known value relaxed program is the unique $H_{kk'}(b)$ that satisfies* (59.1) *and* (59.2) *with equality whenever* (58) *is slack and also satisfies* (60).

Proof of Proposition 12. Arguments analogous to those for the unknown and binary values cases can be used to demonstrate the following facts:

In any solution to the relaxed problem, no buyer wins with a bid less than v^1, and all buyers win with bids of v^1 when the second-highest value is v^1.

The incentive constraint (59.k) should bind whenever $H_{kk'}(b) < \frac{1+\mathbb{I}_{k\neq k'}}{2}$ for some k'.

The remaining piece to characterize the equilibrium is the ordered supports property (60). Define T_2 to be the infimum of b such that

$$p_2 H_{22}(x) + p_3 H_{32}(x) = \frac{p_2}{2} + p_3$$

for all $x \geq b$. Analogously define T_3 to be the infimum b such that $H_{33}(x) = 1/2$ for all $x \geq b$.

Now, suppose that there exists a non-null set of b for which (60) is violated and $b > T_3$. In other words, (59.2) is slack, (59.1) binds, and $H_{32}(b) > 0$ while $H_{22}(b) < 1/2$. In that case, we claim there is a simple perturbation that improves the objective: Let $\tilde{H}_{kk'}(b)$ be defined by

$$H_{kk'}(b) = H_{kk'}(b) + \begin{cases} 0 \text{ if } b < T_3 \text{ or } (k,k') = (3,3); \\ \frac{1}{p_k}\phi(b) \text{ otherwise,} \end{cases}$$

where

$$\phi(b) = \min\left\{ p_3 H_{32}(b), p_2 \frac{1}{2}\left(\frac{1}{2} - H_{22}(b)\right)\right\}.$$

Then clearly the left-hand side of (59.1) has stayed the same while the right-hand side has increased, so that (59.1) is slack. The right-hand side of (59.2) has decreased, but

$$
\begin{aligned}
&\int_{x=v^1}^{b} (p_1 + p_2 H_{22}(x) + p_3 H_{32}(x))\, dx \\
&\geq \int_{x=v^1}^{T_3} (p_1 + p_2 H_{22}(x) + p_3 H_{32}(x))\, dx \\
&= (v^3 - T_3)p_3 = (v^3 - T_3)p_3 H_{33}(T_3) \\
&\geq (v^3 - b)p_3 H_{33}(b)
\end{aligned}
$$

for $b > T_3$. Thus, this solution is clearly feasible. But (59.1) is slack for a non-null set of b for which $H_{33}(b) < 1/2$, so that there is another perturbation that strictly increases revenue. Thus, for $b > T_3$, any optimal solution must satisfy (60).

Now we will use the assumption that p_2 is much larger than p_3 to argue that $T_2 > T_3$, for *any* feasible solution to the relaxed problem. Let $\alpha = p_2/p_3$. A lower bound on T_2 is obviously achieved by the solution that satisfies (60), and in that solution, T_2 is at least as large as the infimum of b for which $H_{22}(x) = 1/2$ for all $x \geq b$, which we will denote by T_{22}. On the other hand, an upper bound on T_3 is achieved by a solution where $H_{32}(b) = 0$ for all b. Thus, T_2 is at least the first time that $H_{22}(b)$ hits $1/2$ when $H_{22}(b)$ solves

$$
\begin{aligned}
(v^2 - b)H_{22}(b) &= \int_{x=v^1}^{b} \left(\frac{p_1}{p_2} + H_{22}(x) \right) dx \\
&\leq \left(\frac{p_1}{p_2} + \frac{1}{2} \right) (b - v^1),
\end{aligned}
$$

and T_3 is no more than the first time that $H_{33}(b)$ hits $1/2$ where $H_{33}(b)$ solves

$$
\begin{aligned}
(v^3 - b)H_{33}(b) &= \int_{x=v^1}^{b} \left(\frac{p_1}{p_3} + H_{33}(x) \right) dx \\
&= \int_{x=v^1}^{b} \left(\alpha \frac{p_1}{p_2} + H_{33}(x) \right) dx \\
&\geq \alpha \frac{p_1}{p_2} (b - v^1).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
H_{22}(b) &\leq \left(\frac{p_1}{p_2} + \frac{1}{2} \right) \frac{b - v^1}{v^2 - b} \\
H_{33}(b) &\geq \alpha \frac{p_1}{p_2} \frac{b - v^1}{v^3 - b}.
\end{aligned}
$$

Clearly, if α is sufficiently large, then $H_{33}(b)$ will have to hit $1/2$ before $H_{22}(b)$.

Now, let us argue that (60) must be satisfied for $b < T_3$, when $T_2 > T_3$. Suppose that there is a non-null set $X \subseteq [v^1, T_3]$ on which (60) is violated. We will construct a new, perturbed solution

$$\tilde{H}_{kk'}(b) = H_{kk'}(b) + \psi_{kk'}(b).$$

This perturbed solution will have $T_3 < \tilde{T}_3 < T_2$, and will satisfy $\psi_{kk'}(b) = 0$ for $b > \tilde{T}_3$. Moreover, we will perturb the solution in such a way that $\psi_{21}(b) \leq 0$, $\psi_{22}(b) \leq 0$, and $\psi_{21}(b) \geq 0$, but over the supports of these perturbations, (59) will be satisfied as an equality. In particular, write

$$\Psi_{kk'}(b) = \int_{x=v^1}^{b} \psi_{kk'}(x)dx.$$

Note that at the solution H, (59) holds with equality over the range $[v^1, T_3]$. For some $\tilde{T}_3$ just slightly larger than T_3, let $\widehat{H}_{33}(b)$ be the path that continues to solve (59.2) as an equality on the range $[T_3, \tilde{T}_3]$, and is equal to 1 for $b > \tilde{T}_3$. Note that $\widehat{H}_{33}(\tilde{T}_3) > 1/2$ is infeasible for the original problem, which imposes $\widehat{H}_{33}(b) \in [0, 1/2]$, though it results in a higher objective (since $\widehat{H} \geq H$).

The perturbed solution $\tilde{H}$ will satisfy (59) as an equality as well, also over the range $[v^1, \tilde{T}_3]$. Thus, differencing the incentive constraints across the perturbed solution and the alternative solution, we conclude that ψ has to satisfy

$$\begin{aligned}
&(v^2 - b)(p_2\psi_{22}(b) + p_3\psi_{32}(b)) = p_2\Psi_{22}(b) \\
&(v^3 - b)p_3\psi_{33}(b) = p_2\Psi_{32}(b) + p_3\Psi_{33}(b).
\end{aligned}$$

Integrating these equations by parts, we conclude that

$$\begin{aligned}
&(v^2 - b)(p_2\Psi_{22}(b) + p_3\Psi_{32}(b)) = -p_3 \int_{x=v^1}^{b} \Psi_{32}(x)dx \\
&(v^3 - b)p_3\Psi_{33}(b) = p_2 \int_{x=v^1}^{b} \Psi_{32}(x)dx
\end{aligned}$$

and so finally

$$p_2\Psi_{22}(b) = -p_3\left(\Psi_{32}(b) + \frac{1}{v^2-b}\int_{x=v^1}^{b}\Psi_{32}(x)dx\right) \tag{61.1}$$

$$p_3\Psi_{33}(b) = p_2\frac{1}{v^3-b}\int_{x=v^1}^{b}\Psi_{32}(x)dx. \tag{61.2}$$

Thus, once we have fixed the perturbation $\psi_{32}(b)$, it is possible to back out the perturbations $\Psi_{22}(b)$ and $\Psi_{33}(b)$ that maintain (59) with equality. In addition, the above formulae imply that

$$p_2^2\Psi_{22}(b) + p_2p_3\Psi_{32}(b) + p_3^2\Psi_{33}(b) = p_2p_3\left(\frac{1}{v_3-b} - \frac{1}{v_2-b}\right)\int_{x=v^1}^{b}\Psi_{32}(x)dx. \tag{62}$$

The left-hand side of this equality, evaluated at $b=\tilde{T}_3$, is the change in the objective function at the perturbed solution relative to the alternative solution (since $\psi_{kk'}(b)=0$ for $b>\tilde{T}_3$). Thus, if we can construct $\psi_{32}(b)$ so that the perturbed solution is feasible for the relaxed problem and such that $\int_{x=v^1}^{\tilde{T}_3}\Psi_{32}(x)dx<0$, we will have demonstrated a perturbation that improves the objective.

We need to derive some bounds on how large perturbations might be that are induced in this manner. If $|\psi_{32}(b)|<\kappa$, then $|\Psi_{32}(b)|\leq(b-v^1)\kappa$, and $|\int_{x=v^1}^{b}\Psi_{32}(b)|\leq\frac{(b-v^1)^2}{2}\kappa$. Differentiating (61), we conclude that

$$p_2\psi_{22}(b) = -p_3\left(\psi_{32}(b) + \frac{1}{v^2-b}\left(\Psi_{32}(b) + \frac{1}{v^2-b}\int_{x=v^1}^{b}\Psi_{32}(x)dx\right)\right)$$

$$p_3\psi_{33}(b) = p_2\frac{1}{v^3-b}\left(\Psi_{32}(b) + \frac{1}{v^3-b}\int_{x=v^1}^{b}\Psi_{32}(x)dx\right)$$

Thus, if $|\psi_{32}(b)|<\kappa$, we conclude that

$$p_2|\psi_{22}(b)| \leq p_3\left(1 + \frac{1}{v^2-b}\frac{(b-v^1)^2}{2} + \frac{b-v^1}{(v^2-b)^2}\right)\kappa \tag{63.1}$$

$$p_3|\psi_{33}(b)| \leq p_2\frac{1}{v^3-b}\left((b-v^1) + \frac{(b-v^1)^2}{2(v^3-b)}\right)\kappa. \tag{63.2}$$

Our perturbation will only involve $b\in[v^1,T_2]$, with $T_1<v^1$, so that by choosing a perturbation $\psi_{32}(b)$ that is uniformly small, we can be guaranteed that the induced $\psi_{22}(b)$ and $\psi_{33}(b)$ will be uniformly small as well. This is necessary in order to ensure that $\tilde{H}$ will be feasible.

Now, let κ small, and let

$$X_\kappa = \left\{ b \in X \,\middle|\, H_{32}(b) \geq \kappa, H_{22}(b) \leq \frac{1}{2} - \kappa, H_{33}(b) \geq \kappa \right\}.$$

For κ sufficiently small, X_κ must be non-null as well. We will define ψ to be the perturbation generated by

$$\psi_{32}(b) = -\epsilon \mathbb{I}_{b \in X_\kappa}$$

for some κ and ϵ. Note that this perturbation induces $\psi_{22}(b) \geq 0$ and $\psi_{33}(b) \leq 0$. Since $T_2 < v^2$, there is an ϵ small enough such that $\tilde{H}_{22}(b) \leq 1/2$, $\tilde{H}_{33}(b) \geq 0$, and $\tilde{H}_{32}(b) \geq 0$. Moreover, since $\psi_{33}(b)$ is continuous in ϵ, there will be an ϵ for which $\tilde{H}_{33}(b)$ hits $1/2$ at $\tilde{T}_3$, as long as $\tilde{T}_3$ is sufficiently close to T_3. In addition, by construction of the perturbations, $\tilde{H}_{kk'}(b)$ will satisfy (59) for $b \in [0, \tilde{T}_3]$. In addition, the right-hand side of (59.1) for $b > \tilde{T}_3$ is

$$\int_{x=v^1}^{b} (p_1 + p_2 H_{22}(x))\, dx + p_2 \Psi_{22}(\tilde{T}_3) + p_3 \Psi_{32}(\tilde{T}_3).$$

But since $\int_{x=v^1}^{\tilde{T}_3} \Psi_{32}(x)dx < 0$, as we argued above, (62) implies that $p_2\Psi_{22}(\tilde{T}_3) + p_3\Psi_{32}(\tilde{T}_3) > 0$. Thus, the right-hand side is larger than under the original solution H, and the paths that $H_{22}(b)$ and $H_{32}(b)$ followed for $b > \tilde{T}_3$ will still satisfy (59). $\square$

Alas, while we can characterize the efficient solution to the relaxed program for certain parameters, we have little hope of constructing BCE that will attain the bounds. This is at issue particularly for range of parameter values in Proposition 12, in which $H_{33}(b)$ hits its upper bound before $H_{22}(b)$ and $H_{32}(b)$. In this case, there are bids in the support of $H_{32}(b)$ that are above the supports of $H_{22}(b)$ and $H_{33}(b)$. Bids in this range are only made by the type v^3 against v^2, in which case the v^2 bidder surely loses. This leads to the following quandary: if v^2 is recommended to bid in this range, then he knows he will lose for sure and would want to increase his bid so as to sometimes win. If v^2 is never told to bid in this range, then v^3 could shade down to highest bid in the support of H_{22}, and continue to win with a lower bid. Thus, it is impossible to satisfy both upward and downward incentive constraints while attaining the solution to the relaxed program.

Corollary 13. *There exist distributions of values such that the welfare bounds from the trinary known value relaxed program are not tight.*

Proof of Corollary 13. Take as given the $H_{kk'}(b)$ as constructed in the proof of Proposition 12, for the case in which α is large. Thus, $T_2 > T_3$, so that $X = \operatorname{supp} H_{32}(b) \setminus [v^1, T_3]$

has positive measure under H_{32}. Note that H is the almost surely unique solution to the relaxed problem, so that any other feasible solution has strictly lower revenue. As a result, if there were BCE that attained revenue close to that in the solution of the relaxed problem, the cumulative distributions of winning bids would have to converge to the solution to the relaxed problem as well.

Consider such a sequence of BCE with marginal winning bid distributions $H^l_{kk'}(\cdot)$ that converge to $H_{kk'}(\cdot)$ in the weak-* topology. We will argue that for such sequence, there exists an l large enough such that $H^l_{kk'}$ cannot be induced by a BCE. Let L^l denote the probability that the v^1 type loses with a bid in X, let W^l_2 denote the probability that the v^2 type wins with a bid in X, and let W^l_3 denote the probability that the v^3 type wins with a bid in X.

Claim: L^l goes to 0 as l goes to ∞. Since the H distributions are absolutely continuous (above v^1), the probability that the v^2 type wins with a bid in X under the lth BCE must go to zero (since this is true under H), so W^l_2 goes to 0. Thus, if L^l goes to $L > 0$, then the deviation which calls for always bidding T_2 when told to bid in X results in an asymptotic gain in surplus of at least

$$(v^2 - T_2)L^l - (T_2 - \min X)W^l_2 \to (v^2 - T_2)L > 0,$$

since the winning bid recommendation in x must be at least $\min X$. For large enough l, this deviation would be profitable.

On the other hand, there is positive probability in the limit that type v^3 wins against v^2 with a bid in X, i.e., W^l_3 goes to $W_3 > 0$. Consider the deviation in which the v^3 type bids $\min X$ whenever told to win with a bid above X. Let $\widehat{b}^l$ denote the average winning bid made by the v^3 type, conditional on winning with a bid in X. By weak-* convergence, $\widehat{b}^l$ converges to some $\widehat{b} > \min X$. The change in surplus from this deviation is

$$(\widehat{b}^l - \min X)(W^l_3 - L^l) - (v^3 - \min X)L^l$$

which will be strictly positive for sufficiently large l.

Thus, it is impossible to have a sequence of BCE whose winning bid distributions converge weakly to the solution to the relaxed problem, and therefore, infimum revenue over all BCE must be bounded away from the solution to the relaxed problem. □

In the end, we conclude that the approach that worked in the unknown values and binary known value cases will not yield a tight characterization of minimum revenue and maximum bidder surplus for general known values models. Nonetheless, the BCE solution concept is highly tractable from a computational perspective. We will take this approach to better understand the known values model in the next section.

D The southwestern frontier with unknown values

Let us explore the southwestern frontier of the set revenue and bidder surpluses that can arise in unknown values BCE, under the assumption that there are two bidders and independent and symmetrically distributed values. Our starting point is a partial equilibrium analysis where we fix the allocation of the good and obtain a lower bound on revenue that is consistent with that allocation. We will then give a general characterization of the allocations that minimize weighted sums of revenue and total surplus. While this characterization does not uniquely pin down the optimal allocation, it does narrow the focus to a relatively small class of candidate solutions. For allocations within this class, we construct BCE that implement the allocation as well as the conditionally revenue minimizing winning bid function. Finally, we will use the calculus of variations to give an even tighter characterization of the optimal allocations for the case of two bidders with independent and standard uniform valuations.

D.1 Preliminaries

Let $q_1(v_1, v_2)$ be the (symmetric) probability that buyer 1 receives the good in equilibrium, and let

$$P^L(v) = 2\int_{v_1=\underline{v}}^{v}\int_{v_2=\underline{v}}^{\overline{v}} (1 - q_1(v_1, v_2))\, p(v_1)p(v_2)dv_2dv_1$$

denote the distribution of the losing value. Note that this function pins down the surplus generated by the auction:

$$\begin{aligned} TS &= 2\int_{v_1=\underline{v}}^{\overline{v}} v_1 \int_{v_2=\underline{v}}^{\overline{v}} q_1(v_1, v_2)p(v_1)p(v_2)dv_2dv_1 \\ &= 2\int_{v_1=\underline{v}}^{\overline{v}} v_1 p(v_1)dv_1 - \overline{v + \int_{v=\underline{v}}^{\overline{v}} (P^L(v_1))\, dv_1}. \end{aligned}$$

There are two special allocations: $\overline{q}$, which is the maximally efficient allocation, and $\underline{q}$, the maximally inefficient allocation. These are defined by

$$\begin{aligned} \overline{q}_1(v_1, v_2) &= \mathbb{I}_{v_1 \geq v_2} \\ \underline{q}_1(v_1, v_2) &= \mathbb{I}_{v_1 \leq v_2}. \end{aligned}$$

Since the independent distribution P is non-atomic, the event that $v_1 = v_2$ has probability zero, so it is irrelevant to welfare (and strategic concerns, as we shall see) what happens as far as allocation on this event.

By analogous arguments as those we employed in Section 4, we can deduce that a lower bound on revenue is given by the winner bidding a deterministic function of the losing value $\beta(v)$ that satisfies (28). Moreover, revenue in this equilibrium is

$$R = \int_{v=\underline{v}}^{\overline{v}} \beta(v) p^L(v) dv = \overline{v} + \int_{v=\underline{v}}^{\overline{v}} \left(P^L(v) - 2\sqrt{P^L(v)} \right) dv.$$

Thus, minimizing a weighted sum $\lambda TS + R$ (with $\lambda > 0$) is equivalent to minimizing

$$\int_{v=\underline{v}}^{\overline{v}} \left((1+\lambda) P^L(v) - 2\sqrt{P^L(v)} \right) dv. \tag{64}$$

We will characterize the minimum of this objective over all allocations $q(v_1, v_2)$. We note for future reference that (64) is a strictly convex function of P^L.

D.2 General characterization

Let us consider the first variation of $P^L(v)$ in the direction $\Xi(v)$, i.e., perturbing to a distribution of the losing value

$$P^L_\epsilon(v) = P^L(v) + \epsilon \Xi(v).$$

The derivative of (64) evaluated at P^ϵ with respect to ϵ, evaluated at $\epsilon = 0$, is

$$\int_{v=\underline{v}}^{\overline{v}} \left(1 + \lambda - \frac{1}{\sqrt{P^L(v)}} \right) \Xi(v) dv = \int_{v=\underline{v}}^{\overline{v}} \mu(v) \Xi(v) dv.$$

Note that as $v \to \underline{v}$, $P^L(v) \to 0$, so that $\mu(v) \to -\infty$. On the other hand, as $v \to \overline{v}$, $P^L(v) \to 1$, so that $\mu(v) \to \lambda > 0$. Moreover, as $P^L(v)$ is strictly increasing, $\mu(v)$ is strictly increasing as well, so that there is a cutoff $\widehat{v}$ such that $\mu(v) < 0$ for $v < \widehat{v}$ and $\mu(v) > 0$ for $v > \widehat{v}$. Thus, the objective decreases if the variation $\Xi(v)$ is positive for $v < \widehat{v}$ and if $\Xi(v)$ is negative for $v > \widehat{v}$. We will refer to this cutoff as the *inflexion point*, i.e., the solution to

$$P^L(\widehat{v}) = \frac{1}{(1+\lambda)^2}.$$

This observation will be immensely useful in characterizing the optimum. For starters, suppose that the optimum is associated with the inflexion point $\widehat{v}$. Then we claim that for profiles such that $v_1 \geq \widehat{v}$ and $v_2 \geq \widehat{v}$, $q_1(v_1, v_2) = \underline{q}_1(v_1, v_2)$ if $v_i < v_j$, and for profiles such that $v_1 \leq \widehat{v}$ and $v_2 < \widehat{v}$, then $q_1(v_1, v_2) = \overline{q}_1(v_1, v_2)$. Why? If this were not the case, then we could change the allocation to the one that allocates the good efficiently if $v_1 \leq \widehat{v}$ and $v_2 \leq \widehat{v}$. This will clearly not change $P^L(\widehat{v})$ (or for $P^L(v)$ with $v > \widehat{v}$), so the inflexion point remains the same. However, we will have weakly increased $P^L(v)$ for all $v \leq \widehat{v}$, since

$$
\begin{aligned}
P^L(v) = 2 &\int_{v_1=\underline{v}}^{v} \int_{v_2=v}^{\overline{v}} \left(1 - q_1(v_1, v_2)\right) p(v_1) p(v_2) dv_2 dv_1 \\
&+ 2 \int_{v_1=\underline{v}}^{v} \int_{v_2=\underline{v}}^{v} \left(1 - q_1(v_1, v_2)\right) p(v_1) p(v_2) dv_2 dv_1 \\
\leq 2 &\int_{v_1=\underline{v}}^{v} \int_{v_2=v}^{\overline{v}} \left(1 - q_1(v_1, v_2)\right) p(v_1) p(v_2) dv_2 dv_1 \\
&+ \int_{v_1=\underline{v}}^{v} \int_{v_2=\underline{v}}^{v} p(v_1) p(v_2) dv_2 dv_1
\end{aligned}
$$

which is what obtains with an efficient allocation on the region $[\underline{v}, \widehat{v}]^2$. By a similar argument, we can conclude that inflexion to the inefficient allocation on the region $[\widehat{v}, \overline{v}]^2$ weakly reduces $P^L(v)$ for $v > \widehat{v}$ without changing $P^L(v)$ for $v \leq \widehat{v}$.

Let us say that a set X is *strictly less than* X' if there exists some (v_1, v_2) such that $w_1 < v_1$ and $w_2 < v_2$ for all $(w_1, w_2) \in X$, and $w_1 > v_1$ and $w_2 > v_2$ for all $(w_1, w_2) \in X'$. Now, suppose there are non-null sets X and X' in $[\underline{v}, \widehat{v}] \times [\widehat{v}, \overline{v}]$ such that X is less than X'. We claim that at the optimum, it cannot be $q_1(v_1, v_2) > 0$ for $(v_1, v_2) \in X$ and that $q_1(v_1, v_2) < 1$ for $(v_1, v_2) \in X'$. If this were true, we could change the allocation rule to

$$
\tilde{q}_1(v) = \begin{cases} (1 - \epsilon) q_1(v) \text{ if } v \in X \\ (1 - \epsilon') q_1(v) + \epsilon' \text{ if } v \in X' \\ q_1(v) \text{ otherwise} \end{cases}
$$

for suitably chosen small $\epsilon > 0$ and $\epsilon' > 0$. In particular, let $\lambda(X)$ denote the Lebesgue measure, and let

$$
\begin{aligned}
L &= \int_{x \in X} q_1(x) \lambda(dx) \\
L' &= \int_{x \in X'} q_1(x) \lambda(dx).
\end{aligned}
$$

We want to reduce q_1 on X and increase q_1 on X' so that $L + L'$ remains the same. Note that

$$\tilde{L} = \int_{x\in X} \tilde{q}_1(x)\lambda(dx) = (1-\epsilon)L$$
$$\tilde{L}' = \int_{x\in X'} \tilde{q}_1(x)\lambda(dx) = (1-\epsilon')L' + \epsilon'\lambda(X').$$

Thus, we need $\epsilon L = \epsilon'(\lambda(X') - L')$, so we let $\epsilon' = \min\{1, (\lambda(X') - L')/L\}$ and then define ϵ by the preceding equation. The net effect of this perturbation is that $P^L(\widehat{v})$ has remained the same, but $P^L(v)$ has weakly increased at $v < \widehat{v}$ and weakly decreased at $v > \widehat{v}$.

An implication of this result is that at an optimum, the set of points for which $q_1(v) \in (0,1)$ must have Lebesgue measure zero. Let us call the set X, and write $\widehat{X}$ for the subset of X with Lebesgue density equal to 1 Cohn (1980, Corollary 6.2.6). Now, suppose there is some $v, v' \in X$ with $v_1 < v'_1$ and $v_2 < v'_2$. Then for $\epsilon < \min\{v'_1 - v_1, v'_2 - v_2\}/2$, we have that $B_\epsilon(v)$ is strictly less than $B_\epsilon(v') = \emptyset$, so that $Y = X \cap B_\epsilon(v)$ and $Y' = X \cap B_\epsilon(v)$ (i) both have positive measure, (ii) Y is strictly less than Y', and (iii) $q_1(v) \in (0,1)$ for all $v \in Y$ and $v \in Y'$, which is impossible. Thus, we conclude that there are no two $v, v' \in \widehat{X}$ that can be ordered in this manner. But now we can write $\widehat{X}_1$ for the projection of $\widehat{X}$ onto its first coordinate, and write $f : \widehat{X}_1 \to \mathbb{R}$ for the function $f(x_1) = \sup\{x_2 | (x_1, x_2) \in \widehat{X}\}$. Thus, $\widehat{X}$ is contained in the union of the graph of f, together with vertical line segments that contain the discontinuities, of which there are countably many. Both the graph and the countably many vertical line segments have Lebesgue measure zero, so $\widehat{X}$ has Lebesgue measure zero. But $\lambda(\widehat{X}) = \lambda(X)$, so we are done.

Thus, $q_1(v) \in \{0,1\}$ almost surely, and we can always change $q_1(v)$ so that this property holds everywhere without changing P^L. Let us write $Z(v) = q_1(v)v_2 + (1 - q_1(v))v_1$, i.e., $Z(v)$ is the valuation of the loser. We additionally write $M(v) = \{x | z(x) \le v\}$, so that

$$P^L(v) = P^2(M(v)),$$

where we identify P^2 as the (independent and symmetric) prior measure on $[\underline{v}, \overline{v}]^2$. Note that because of our previous discussion, we know that $v \in M(\widehat{v})$ almost surely implies that $Z(v) = \min v$ and $v \notin M(\widehat{v})$ almost surely implies that $Z(v) = \max v$. Moreover, it must be that the set $M(\widehat{v})$ is "downward closed" in the sense that for almost all $v \in M(\widehat{v})$,

$$P^2\left(M(\widehat{v}) \cap ([\underline{v}, v_1] \times [\underline{v}, v_2])\right) = P^2\left([\underline{v}, v_1] \times [\underline{v}, v_2]\right). \tag{65}$$

We can then modify q_1 so that this condition holds exactly, by setting $Z(v) = \min v$ if and only if (65) holds.

Thus, the allocation and the distribution of the losing value are pinned down by the choice of the downward closed set $M(\widehat{v})$, which is the region on which the allocation is efficient. Because it is downward closed, this set is the hypograph of a monotonically decreasing boundary curve $\phi : [\underline{v}, \overline{v}] \to [\underline{v}, \overline{v}]$, where

$$\phi(v_1) = \max\{v_2 | (v_1, v_2) \in M(\widehat{v})\}.$$

It turns out that $P^L(v)$ can be compactly expressed as a function of ϕ. Note that $\phi(\underline{v}) > \underline{v}$, and we note that $\widehat{v} = \inf\{v | \phi(v) \leq v\}$. For $v \in [\underline{v}, \widehat{v}]$, we have that

$$P^L(v) = 2\int_{x=\underline{v}}^{v} (P(\phi(x)) - P(x))p(x)dx.$$

Conversely, for $v > \widehat{v}$,

$$P^L(v) = P^L(\phi^{-1}(v)) + (P(v) - P(\phi^{-1}(v)))^2,$$

where

$$\phi^{-1}(v) = \sup\{v' \in [\underline{v}, \widehat{v}] | \phi(v') \geq v\}$$

in cases where $\phi(v)$ has flats. When ϕ is invertible, we can more simply write

$$P^L(\phi(v)) = P^L(v) + (P(\phi(v)) - P(v))^2.$$

This concludes our general characterization of the optimal allocation. Optimal allocations always have a monotonically decreasing boundary function; the allocation is efficient below the boundary, and the allocation is inefficient above the boundary. The optimal boundary function turns out to be the solution to a rather messy variational problem. We will subsequently have more to say in for the case of the uniform distribution, for which the variational problem turns out to be quite tractable. However, we will now show that regardless of what the monotonically decreasing boundary is (optimal or otherwise), we can construct BCE which implement the corresponding allocation and conditionally minimize revenue, as we shall now see.

D.3 Extension to an equilibrium

Let us suppose that we have a monotonically decreasing boundary function $\phi(v)$ so that the allocation is efficient whenever $\min v \leq \widehat{v}$ and $\max v \leq \phi(\min v)$, and otherwise the allocation is inefficient. This boundary function induces a distribution of the losing value $P^L(v)$, and through that a function β that determines the winning bid for each possible losing value.

We now construct a BCE in which the winning bid coincides with β. The construction will closely follow the pattern that we have previously established: Since bidders are indifferent to all uniform upward deviations, they must also almost surely be indifferent to pointwise upward deviations, which pins down the marginal distribution of losing signals as

$$L(s|v) = \frac{\sqrt{P^L(s)}}{\sqrt{P^L(v)}}$$

where s is the loser's signal and v is the loser's valuation. This has to be the marginal distribution of the losing bidder's signal conditional on the winner's signal (and hence the loser's value) in order for pointwise upward incentive constraints to be satisfied. However, there is some flexibility as to how loser's signals are correlated with the winner's value. From the allocation, we know that if the winner's signal v is less than $\widehat{v}$, the loser's value is the lowest value, and the winner's value is distributed according to a truncated prior on $[v, \phi(v)]$. But if v is larger than $\widehat{v}$, then it must be that v is the highest of the two values, so that the winner's value is distributed according to a truncated prior on the region $[\phi^{-1}(v), v]$. In either case, we will make the joint distribution of the loser's signal comonotonic with the winner's value. In other words, given the winner's signal, the percentile of the winner's value is perfectly correlated with the percentile of the loser's signal. This corresponds with our construction of the maximally inefficient equilibrium in Section 5.1, but contrasts with our construction of the revenue minimizing equilibrium in Section 4. In the latter case, we made the loser's signal independent of the winner's value conditional on the winner's signal.

Claim: regardless of which case we are in, the distribution of the winner's value must (weakly) first-order stochastic dominate the distribution of the loser's signal. This is obvious in the first case since the support of the winner's value is above the support of the loser's signal, which is $[\underline{v}, v]$. In the second case, the winner's value v' is distributed on $[\phi^{-1}(v), v]$ according to the cumulative distribution

$$G(v'|v) = \frac{P(v') - P(\phi^{-1}(v))}{P(v) - P(\phi^{-1}(v))},$$

while the loser's value is distributed according to $L(s|v)$ on $[\underline{v}, v]$. We will show the even stronger result that

$$\frac{\sqrt{P^L(s)}}{\sqrt{P^L(v)}} \geq \frac{P(s)}{P(v)}$$

so we just need to prove that $\sqrt{P^L(s)}/P(s)$ is decreasing in s. The derivative of this expression is

$$\frac{\frac{p^L(s)}{2\sqrt{P^L(s)}}P(s) - p(s)\sqrt{P^L(s)}}{(P(s))^2} = \frac{\frac{p^L(s)}{2}P(s) - p(s)P^L(s)}{\sqrt{P^L(s)}(P(s))^2}.$$

For $s < \widehat{v}$, this evaluates to

$$\frac{(P(\phi(s)) - P(s))P(s) - P^L(s)}{\sqrt{P^L(s)}(P(s))^2}p(s)$$

so it is sufficient to show that $P(\phi(s)) - P(s) \leq P^L(s)$. But this has to be the case since

$$P^L(s) = 2\int_{x=\underline{v}}^{s}(P(\phi(x)) - P(x))p(x)dx \geq 2P(s)P(\phi(s)) - P(s)),$$

since ϕ is weakly decreasing. For $s > \widehat{v}$, we know that

$$P^L(s) = P^L(\widehat{v}) + 2\int_{x=\widehat{v}}^{s}(P(s) - P(\phi^{-1}(s)))p(x)dx.$$

Thus, $p^L(s) = 2(P(s) - P(\phi^{-1}(s)))p(s)$, so that condition for $\sqrt{P^L(s)}/P(s)$ to be decreasing simplifies to

$$(P(s) - P(\phi^{-1}(s)))P(s) - P^L(s) \leq 0$$

which is obviously true because $P^L(s) \geq (P(s))^2$.

Thus, the distribution $G(s|v)$ stochastically dominates $L(s|v)$. Thus, if we write $v^W(s)$ for the winner's value as a function of the loser's signal according to the comonotonic joint distribution, then $v^W(s) \geq s$ for all s. Let us use this to verify that downward constraints are satisfied. Suppose that a bidder is told to bid $\beta(v)$ and considers deviating downwards to $\beta(s)$. That buyer's expected payoff is proportional to

$$\int_{x=\underline{v}}^{s} (v^W(x) - \beta(s))L(dx|v).$$

Thus, the derivative of this payoff with respect to s is

$$(v^W(s) - \beta(s))\frac{p^L(s)}{2\sqrt{P^L(s)}\sqrt{P^L(v)}} - \beta'(s)\frac{\sqrt{P^L(s)}}{\sqrt{P^L(v)}}.$$

Substituting in (28), we conclude that the derivative is

$$(v^W(s) - s)\frac{p^L(s)}{2\sqrt{P^L(s)}\sqrt{P^L(v)}},$$

which we know must be non-negative. Thus, the deviator's payoff is increasing in s for $s < v$, and the bid of $\beta(v)$ is rational.

D.4 Uniform example

In general, what we can say is that the equilibrium that minimizes $\lambda TS + R$ for $\lambda > 0$ corresponds to a conditionally revenue minimizing equilibrium for some allocation given by the boundary function ϕ. The precise ϕ that minimizes the objective turns out to be the solution of a rather messy problem in the calculus of variations. For the case of the uniform distribution, however, this variational problem has an elegant solution. In particular, let us suppose that $P(v) = v$ and $[\underline{v}, \overline{v}] = [0, 1]$, so that for $v \leq \widehat{v}$,

$$P^L(v) = 2\int_{x=0}^{v} (\phi(x) - x)dx,$$

and for $v \geq \widehat{v}$,

$$P^L(v) = P^L(\phi^{-1}(v)) + (v - \phi^{-1}(v))^2.$$

Let us consider a region $[\underline{w}, \overline{w}]$ on which ϕ is strictly decreasing. In that case, we can write $P^L(\phi(v))$ for $v \in [\underline{v}, \overline{v}]$ as

$$P^L(\phi(v)) = P^L(v) + (\phi(v) - v)^2.$$

Thus, the contribution to the objective for $v \in [\underline{w}, \overline{w}]$ and $v \in [\phi(\overline{w}), \phi(\underline{w})]$ is

$$\begin{aligned}
&\int_{v=\underline{w}}^{\overline{w}} ((1+\lambda)P^L(v) - 2\sqrt{P^L(v)})dv \\
&\quad - (1+\lambda)\int_{v=\underline{w}}^{\overline{w}} (P^L(v) + (\phi(v) - v)^2)\phi'(v)dv \\
&\quad + 2\int_{v=\underline{w}}^{\overline{w}} \sqrt{P^L(v) + (\phi(v) - v)^2}\phi'(v)dv.
\end{aligned} \tag{66}$$

This objective can be simplified even further. Note also that

$$\begin{aligned}
p^L(v) &= 2(\phi(v) - v); \\
p^{L,\prime}(v) &= 2(\phi'(v) - 1).
\end{aligned}$$

Thus, (66) can be further simplified as

$$\begin{aligned}
J(P^L) = &\int_{v=\underline{w}}^{\overline{w}} ((1+\lambda)P^L(v) - 2\sqrt{P^L(v)})dv \\
&- (1+\lambda)\int_{v=\underline{w}}^{\overline{w}} \left(P^L(v) + \left(\frac{p^L(v)}{2}\right)^2\right)\left(1 + \frac{p^{L,\prime}(v)}{2}\right)dv \\
&+ 2\int_{v=\underline{w}}^{\overline{w}} \sqrt{P^L(v) + \left(\frac{p^L(v)}{2}\right)^2}\left(1 + \frac{p^{L,\prime}(v)}{2}\right)dv,
\end{aligned} \tag{67}$$

thereby replacing any reference to the boundary function ϕ with references to derivatives of the function $P^L(v)$. We will use the calculus of variations to identify necessary conditions that must be satisfied by a solution P^L that is smooth on the range $[\underline{w}, \overline{w}]$. Let us compute the first variation of J. In particular, let us consider a candidate solution

$$P^L_\epsilon(v) = P^L(v) + \epsilon\Xi(v)$$

where $\Xi(\underline{w}) = \Xi(\overline{w}) = \xi(\underline{w}) = \xi(\overline{w}) = 0$. This variation corresponds to a smooth variation of ϕ such that $\phi(\underline{w})$ and $\phi(\overline{w})$ are unchanged. The first variation is

$$\begin{aligned}
\frac{d}{d\epsilon} J(P^L_\epsilon)\bigg|_{\epsilon=0} = & \int_{v=\underline{w}}^{\overline{w}} \left(1 + \lambda - \frac{1}{P^L(v)}\right) \Xi(v) dv \\
& - (1+\lambda) \int_{v=\underline{w}}^{\overline{w}} \left(\Xi(v) + \frac{p^L(v)}{2}\xi(v)\right) \left(\frac{1}{2}p^{L,\prime}(v) + 1\right) dv \\
& - (1+\lambda) \int_{v=\underline{w}}^{\overline{w}} \left(P^L(v) + \left(\frac{p^L(v)}{2}\right)^2\right) \frac{1}{2}\xi'(v) dv \\
& + \int_{v=\underline{w}}^{\overline{w}} \frac{\Xi(v) + \frac{p^L(v)}{2}\xi(v)}{\sqrt{P^L(v) + \left(\frac{p^L(v)}{2}\right)^2}} \left(\frac{1}{2}p^{L,\prime}(v) + 1\right) dv \\
& + 2 \int_{v=\underline{w}}^{\overline{w}} \sqrt{P^L(v) + \left(\frac{p^L(v)}{2}\right)^2} \frac{1}{2}\xi'(v) dv.
\end{aligned}$$

Through repeated application of integration by parts, this expression reduces to

$$\begin{aligned}
\frac{d}{d\epsilon} J(P^L_\epsilon)\bigg|_{\epsilon=0} = & \int_{v=\underline{w}}^{\overline{w}} \left(1 + \lambda - \frac{1}{\sqrt{P^L(v)}}\right) \Xi(v) dv \\
& - (1+\lambda) \int_{v=\underline{w}}^{\overline{w}} \left(\frac{1}{2}p^{L,\prime}(v) + 1\right) \Xi(v) dv \\
& + \int_{v=\underline{w}}^{\overline{w}} \frac{1}{\sqrt{P^L(v) + \frac{1}{4}(p^L(v))^2}} \left(\frac{1}{2}p^{L,\prime}(v) + 1\right) \Xi(v) dv.
\end{aligned}$$

If this variation is zero for all Ξ, it must be that

$$1 + \lambda - \frac{1}{P^L(v)} - \left(1 + \lambda - \frac{1}{\sqrt{P^L(v) + \frac{1}{4}\left(p^L(v)\right)^2}}\right) \left(\frac{1}{2}p^{L,\prime}(v) + 1\right) = 0$$

for all v. This is typically called the Euler-Lagrange equation (Clarke, 2013, Chapter 14.1). This reduces to the second-order non-linear ordinary differential equation

$$p^{L,\prime}(v) = 2\left(\frac{1 + \lambda - \frac{1}{\sqrt{P^L(v)}}}{1 + \lambda - \frac{1}{\sqrt{P^L(v) + \frac{1}{4}(p^L(v))^2}}} - 1\right). \tag{68}$$

We note that the Lagrangean $J(P^L)$ is strictly convex in $(P^L, p^L, p^{L,\prime})$, so that a solution to (68) must be a global minimum among smooth solutions (Clarke, 2013, Theorem 15.9). We also note that it is possible to back out the boundary function from $p^L(v)$ using

$$\phi(v) = \frac{1}{2}p^L(v) + v.$$

The differential equation (68) must be satisfied by the induced allocation on any stretch $[\underline{w}, \overline{w}]$ on which ϕ is decreasing, i.e., $\phi'(v) \leq 0$. This is true if and only if $p^{L,\prime}(v) \leq -2$. There are global feasibility constraints on $p^L(v)$: $p^L(v) \geq 0$ for all v, and $p^L(v) \leq 2(1-v)$.

Finally, we have not yet determined the correct inflexion point $\widehat{v}$ corresponding to a given λ, which must satisfy $P^L(\widehat{v}) = 1/(1+\lambda)^2$. Now, if $p^L(\widehat{v}) > 0$, then $p^{L,\prime}(v) \to -2$, which corresponds to a boundary with $\phi'(\widehat{v}) = 0$, i.e., the boundary function is locally flat close to the cutoff. On the other hand, if $p^L(\widehat{v}) = 0$, let us use L'Hôpital's rule to compute the limit of (68) as $v \uparrow \widehat{v}$. This must be

$$\begin{aligned} p^L(\widehat{v}) &= -2 + 2\lim_{v\uparrow\widehat{v}} \frac{\frac{p^L(v)}{(P^L(v))^{3/2}}}{\frac{p^L(v)(1+1/2p^{L,\prime}(v))}{(P^L(v)+(p^L(v))^2/4)^{3/2}}} \\ &= -2 + 2\frac{1}{1+1/2p^{L,\prime}(\widehat{v})}. \end{aligned}$$

Thus, $p^L(\widehat{v})$ must be a solution to the polynomial equation $x^2 + 4x = 0$ which has a single solution less than -2, and we conclude that $p^L(\widehat{v}) = -4$. This corresponds to $\phi'(\widehat{v}) = -1$.

It remains an open question, analytically speaking, whether or not $p^L(\widehat{v}) > 0$ is optimal. However, numerical simulations have indicated to us that it is the latter structure, in which $p^L(\widehat{v}) = 0$, which minimizes the objective, and thus we will focus on "smooth" solutions where $p^L(\widehat{v}) = 0$. There are two remaining possibilities: It could be that the constraint that $\phi(v) \leq 1$ (which corresponds to the constraint that $p^L(v) \leq 2(1-v)$) is initially binding, and there is a flat where $\phi(v) = 1$, after which $\phi(v)$ decreases smoothly to $\phi(\widehat{v}) = \widehat{v}$ according to (68). In the second alternative, $\phi(\underline{v}) < 1$ initially, and the path of P^L is everywhere pinned down by (68). Note that (68) blows up as $v \to 0$, since $P^L(v) \to 0$ so that $p^{L,\prime}(v) \to -\infty$.

In practice, we solved examples by guessing the correct $\widehat{v}$ for a given λ and sending out the solution according to the terminal conditions that $p^{L,\prime}(\widehat{v}) = -1$ and $p^L(\widehat{v}) = 0$. If the path of $p^L(v)$ ever goes above the $2(1-v)$ boundary, we truncate it at that point, and look for solutions such that $P^L(\widehat{v}) = 1/(1+\lambda)^2$. Figure 5 displays the results of this computation for $\lambda = 1.1$ and $\lambda = 2.0$. The blue line is the optimal distribution of the losing value, with the red line being its density. The purple line is function ϕ, which is the upper boundary of the region $M(\widehat{v})$ on which the allocation is efficient. The reader can now see why we have referred to $\widehat{v}$ as the inflexion point, since $p^L(\widehat{v}) = 0$.

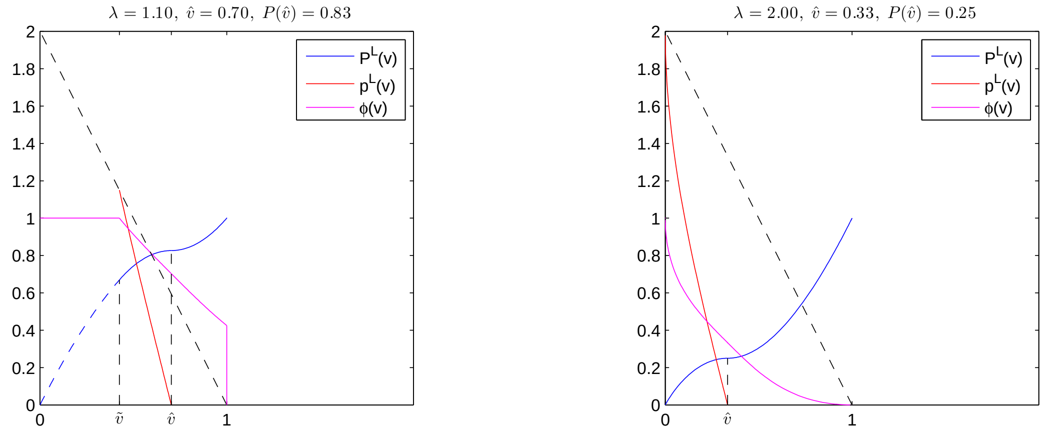

Figure 5: Two examples of the optimal boundary and the corresponding cumulative distribution of losing valuations.

In the left panel, λ is relatively low and the direction is somewhat close to minimizing revenue. As a result, the allocation is close to efficient, which is reflected in the boundary function $\phi(v)$ being quite high. In fact, ϕ is stuck at its upper bound for low to moderate valuations. In the right panel, λ is somewhat higher and the allocation is now closer to always inefficient. Indeed, the boundary function ϕ never hits its upper bound and asymptotes at zero. Based on the preceding analysis, we have also constructed more precise computations of the set of possible welfare outcomes for two bidders and independent standard uniform distributions. The results of this computation are presented in Figure 6.

We close by commenting on how this example might generalize. For more general independent distributions, the corresponding differential equation to (68) is much more complicated, and essentially only implicitly defines $p^{L,\prime}(v)$ as a function of $P^L(v)$ and $p^L(v)$. We suspect, however, that the optimal boundary function will have a smooth structure for those cases. Beyond two players and independence, there is still a well defined lower bound on revenue conditional on the allocation, but it is no longer clear that such allocations can be implemented subject to downward incentive constraints.The known values surplus set is depicted in red in Figure 1, and it is significantly smaller than the unknown values surplus set. At point G, revenue is maximized and bidders are held down to the lower bound from Section 4. In this example, minimum revenue for the known values model is approximately 0.11, while minimum revenue for the unknown values model is approximately 0.06. This difference corresponds to roughly 8 percent of the efficient surplus.

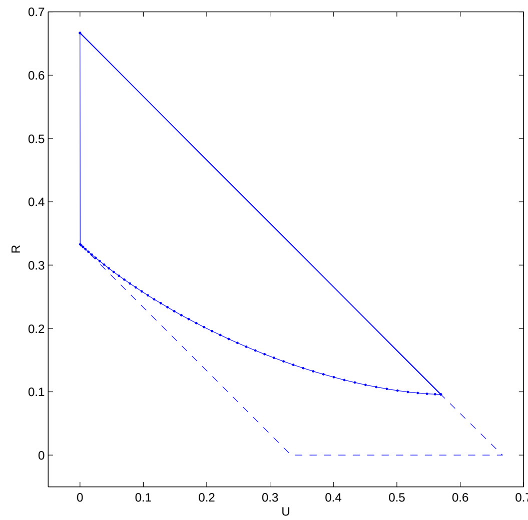

Figure 6: The set of welfare outcomes with two bidders and independent standard uniform valuations.

E Ex-Ante Entry Fee versus Reserve Prices

Our characterization from Theorem 2 can explain why ex-ante entry fees will dominate reserve prices in terms of maximum revenue when values are known. Suppose that there are n buyers whose values are i.i.d. draws from a cumulative distribution $P(v)$ on $[\underline{v}, \overline{v}]$. An entry fee e and reserve price r together induce an *exclusion level* $\underline{x}(r, e)$, which is the lowest-value that would enter the auction in a revenue maximizing equilibrium. It is still the case that, conditional on entering, buyers will not bid more than their values, and indeed, it is possible to construct an equilibrium via similar methods in which bidder surplus is exactly what bidders could attain by best responding when others bid their values. The difference is that when a buyer does not enter, that effectively imputes a bid for that buyer which is equal to the reserve price. Moreover, the allocation is efficient conditional on the set of types that enter, so that the lowest entering type will only win when no other buyer enters, and in ex-ante terms, the cutoff type must receive a surplus of zero. This means that the exclusion level $\underline{x}(r, e)$ is implicitly defined by

$$e = (\underline{x} - r)P^{N-1}(\underline{x}). \tag{69}$$

It is straightforward to verify that $\underline{x}$ is increasing in both of its arguments and if P is strictly increasing, then $\underline{x}$ will be strictly increasing in each argument. (We will assume P is strictly increasing for the remainder of the heuristic argument.)

Now, let us consider two distinct pairs (r, e) and (r', e') that induce the same exclusion level x. Then it is without loss of generality to assume that $r < r'$ and $e > e'$. We will argue that revenue is higher under (r, e) by arguing that bidder surplus is uniformly lower for all types that enter under (r, e) than under (r', e'). Let $b^*(v)$ be the solution to

$$\max_b (v - b)P^{N-1}(b) = \underline{u}(v).$$

Thus, $\underline{u}(v)$ is the interim lower bound bidder surplus when others bid their values and all types enter. It is straightforward to show that $b^*(v)$ is (weakly) increasing and $\underline{u}(v)$ is strictly increasing. Thus, among buyers who enter, the optimal lower bound bidding strategy will involve another cutoff $\widehat{x}(r, e) \geq \underline{x}(r, e)$, where buyers with values between $\underline{x}(r, e)$ and $\widehat{x}(r, e)$ will bid r and buyers with values above $\widehat{x}(r, e)$ will bid $b^*(v)$, and at the cutoff,

$$(\widehat{x} - r)P^{N-1}(\underline{x}(r, e)) = (\widehat{x} - b^*(\widehat{x}))P^{N-1}(b^*(\widehat{x})).$$

The cutoff $\widehat{x}$ must also be increasing in (r, e). Thus, inducing a given exclusion level with a lower reserve price and a higher fee tends to increase bidding at the reserve price. Specifically, if $r < r'$ and $\underline{x}(r, e) = \underline{x}(r', e') = \underline{x}$, then $\widehat{x}(r, e) \geq \widehat{x}(r', e')$. This is intuitive, because the probability of winning at the reserve price is fixed but the cost of winning at the reserve price goes down, then clearly bidding at the reserve price must become more attractive to all types.

We can now compare bidder surplus across the two exclusion mechanisms. If v is between $\underline{x}$ and $\widehat{x}(r', e')$, then interim bidder surplus must be the same using either reserve and fee pair, because these types bid the reserve price and obtain surplus

$$(v - r)P^{N-1}(\underline{x}) - e = (v - v')P^{N-1}(\underline{x}) = (v - r')P^{N-1}(\underline{x}) - e'.$$

This follows from the entry condition (69). If v is greater than $\widehat{x}(r, e)$, lower bound surplus is necessarily lower with the lower reserve price, since the reserve price does not distort bidding behavior (leaving interim surplus the same) but the entry fee is higher. Finally, for valuations that are between $\widehat{x}(r', e')$ and $\widehat{x}(r, e)$, the difference in surplus is

$$u(v) - (v - r)P^{N-1}(\underline{x}) - e \geq (v - r)P^{N-1}(\underline{x}) - e - (v - r')P^{N-1}(\underline{x}) + e' = 0,$$

so that these buyers attain higher surplus under (r', e') than under (r, e).

Intuitively, conditional on the exclusion level, a lower reserve price induces greater distortion in bidding behavior upon entry away from the unconditional optimum b^*. Both distortion and higher entry fees tend to decrease bidder surplus, relative to the no exclusion case. Thus, by setting a lower reserve and a higher fee, the seller simultaneously induces more distortion and extracts more rents from the buyers' whose behavior is not distorted, which must be decreasing the lower bound bidder surplus.

We note that this conclusion corresponds with a result of Milgrom and Weber (1982) that entry fees induce greater revenue than reserve prices when signals and values are affiliated. Their argument also involves comparison of two different pairs of reserve prices and fees that induce the same exclusion level, and to some extent, a similar logic may be underlying the two results.