cemmap

centre for microdata methods and practice

Identification of the distribution of valuations in an incomplete model of English auctions

Andrew Chesher
Adam Rosen

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP30/15

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

Identification of the Distribution of Valuations in an Incomplete Model of English Auctions*

Andrew Chesher and Adam M. Rosen
CeMMAP and UCL

June 26, 2015

Abstract

An incomplete model of English auctions with symmetric independent private values, similar to the one studied in Haile and Tamer (2003), is shown to fall in the class of Generalized Instrumental Variable Models introduced in Chesher and Rosen (2014). A characterization of the sharp identified set for the distribution of valuations is thereby obtained and shown to refine the bounds available until now.

Keywords: English auctions, partial identification, sharp set identification, generalized instrumental variable models.

1 Introduction

The path breaking paper Haile and Tamer (2003) (HT) develops bounds on the common distribution of valuations in an incomplete model of an open outcry English ascending auction in a symmetric independent private values (IPV) setting.

One innovation in the paper was the use of an incomplete model based on weak plausible restrictions on bidder behavior, namely that a bidder never bids more than her valuation and never allows an opponent to win at a price she is willing to beat. An advantage of an incomplete model is that it does not require specification of the mechanism relating bids to valuations. Results obtained using the incomplete model are robust to misspecification of such a mechanism. The incomplete model may be a better basis for empirical work than the button auction model of Milgrom and Weber (1982) sometimes used to approximate the process delivering bids in an English open outcry auction.

*We gratefully acknowledge financial support from the UK Economic and Social Research Council through a grant (RES-589-28-0001) to the ESRC Centre for Microdata Methods and Practice, and from the European Research Council (ERC) grant ERC-2012-StG-312474. We thank Phil Haile for comments. We have benefited from feedback given by seminar participants at Chicago, Yale, and Cornell, as well as participants at a 2015 Seoul National University and CeMMAP conference on Advances in Microeconometrics. Any errors are our own.

On the down side the incomplete model is partially, not point, identifying for the primitive of interest, namely the common conditional probability distribution of valuations given auction characteristics. HT derive bounds on this distribution and shows how to use these to make inferences about the distribution and about interesting features that are functionals of the distribution such as the optimal reserve price.

The question of the sharpness of those bounds was left open in HT. In this paper we resolve this question. We consider a slightly simplified version of the model in HT and show that the model falls in the class of Generalized Instrumental Variable (GIV) models introduced in Chesher and Rosen (2014), (CR). We obtain a characterization of the sharp identified set for the auction model by applying the general characterization for GIV models given in CR. We show that there are bounds additional to those given in HT and in numerical calculations demonstrate that they can be binding.

The characterization of the sharp identified set of valuation distributions comprises a dense system of infinitely many inequalities restricting not just the value of the distribution function via pointwise bounds on its level but also restricting its shape as it passes between the pointwise bounds.

Partial identification has been usefully applied to address other issues in auction models since HT. Tang (2011) and Armstrong (2013) both study first-price sealed bid auctions. Tang (2011) assumes equilibrium behavior but allows for a general affiliated values model that nests private and common value models. Without parametric distributional assumptions model primitives are generally partially identified, and he derives bounds on seller revenue under counterfactual reserve prices and auction format. Armstrong (2013) studies a model in which bidders play equilibrium strategies but have symmetric independent private values *conditional* on unobservable heterogeneity, and derives bounds on the mean of the bid and valuation distribution, and other functionals thereof. Aradillas-Lopez, Gandhi, and Quint (2013) study second price auctions that allow for correlated private values. Theorem 4 of Athey and Haile (2002) previously showed non-identification of the valuation distribution in such models, even if bidder behavior follows the button auction model equilibrium. Aradillas-Lopez, Gandhi, and Quint (2013) impose a slight relaxation of the button auction equilibrium, assuming that transaction prices are determined by the second highest bidder valuation. They combine restrictions on the joint distribution of the number of bidders and the valuation distribution with variation in the number of bidders to bound seller profit and bidder surplus.

The restrictions of the HT model we study are set out in Section 2. In Section 3 GIV models are introduced and the auction model is placed in the GIV context. The sharp identified set for the auction model is characterized in Section 4 and the inequalities that feature in this characterization are explored in Section 5. A numerical example is presented in Section 6 and calculation of approximations to identified sets of parameters in a parametric model is given in Section 7. Section

8 concludes.

2 Model

We study open outcry English ascending auctions with a finite number of bidders, M, which may vary from auction to auction. The model is a slight simplification of the model studied in HT in that there is no reserve price and the minimum bid increment is zero. These conditions simplify the exposition and are easily relaxed.[1]

The final bid made by each bidder is observed. Each auction is associated with a value, z, of an observable variable Z. In some applications the number of bidders, M, could be an element of Z.

The mth largest value in an M element list of numbers $x = (x_1, \ldots, x_M)$ or random variables $X = (X_1, \ldots, X_M)$ will be denoted by $x_{m:M}$ and $X_{m:M}$, respectively. For example, $x_{M:M} = \max(x)$.

Restriction 1. In an auction with M bidders, the final bids and valuations are realizations of random vectors $B = (B_1, \ldots, B_M)$ and $V = (V_1, \ldots, V_M)$ such that for all $m = 1, ..., M$, $B_m \leq V_m$ almost surely.

Restriction 2. In every auction the second highest valuation, $V_{M-1:M}$, is no larger than the highest final bid, $B_{M:M}$. That is, $V_{M-1:M} \leq B_{M:M}$ almost surely.

Restriction 3. There are independent private values conditional on auction characteristics $Z = z$ such that the valuations of bidders are identically and independently continuously distributed with conditional distribution function given $Z = z$ denoted by $F_z(\cdot)$.

Restriction 4. Conditional on auction characteristics $Z = z$, bids $(B_1, \ldots, B_M)$ are exchangeable.

Auctions are characterized by a bid vector B, a vector of valuations V, the number of bidders M, and auction characteristics Z. B, V, M, Z are presumed to be realized on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with sigma algebra $\mathcal{A}$ endowed with the Borel sets on Ω. Valuations V are not observed. Observations of (B, Z, M) across auctions render the joint distribution of these variables identified. The goal of our identification analysis is to determine what this joint distribution reveals about $F_z(\cdot)$. In the remainder of the paper, inequalities involving random variables, such as those in Restrictions 1 and 2 and those stated in Lemma 1 below are to be understood to mean these inequalities hold $\mathbb{P}$ almost surely.

Restrictions 1 and 2 are the HT restrictions on bidder behavior. Restriction 3 focuses our analysis to the independent private values paradigm. Restriction 4 additionally imposes that bids are symmetric given auction characteristics z. Restrictions 1-3 were imposed by HT while Restriction 4 was not. In Theorem 1 we present bounds on $F_z(\cdot)$ that refine those of HT using only Restrictions

[1] These conditions are also imposed in Appendix D of HT in which the sharpness of identified sets is discussed. As was the case in HT, with a reserve price r our analysis applies to the distribution of valuations truncated below at r.

1-3. The bounds are shown to be sharp if, additionally, Restriction 4 holds, or if the researcher only has data on bid order statistics.

The following inequalities involving the order statistics of B and V set out in Lemma 1 are a consequence of Restrictions 1 and 2. The proof of the lemma, like all other proofs, is provided in Appendix A.

Lemma 1. *Let Restrictions 1-3 hold. Then for all m and M*

$$B_{m:M} \leq V_{m:M} \tag{2.1}$$

$$B_{M:M} \geq V_{M-1:M}. \tag{2.2}$$

In similar manner to HT, our identification analysis is based on the restrictions (2.1) and (2.2) on bid and valuation order statistics.

3 Generalized Instrumental Variable models

This auction model falls in the class of Generalized Instrumental Variable (GIV) models introduced in Chesher and Rosen (2014).

We consider M-bidder auctions for some particular value of M and use the results in CR to characterize the sharp identified set of valuation distributions delivered by a joint distribution of M ordered final bids. In cases where M is not included as an element of Z, so that $F_z(\cdot)$ is restricted to be invariant with respect to M, the intersection of the sets obtained with different values of M gives the identified set of valuation distributions in situations in which there is variation in the number of bidders across auctions.

A GIV model places restrictions on a process that generates values of observed endogenous variables, Y, given exogenous variables Z and U, where Z is observed and U is unobserved. The variables (Y, Z, U) take values on $\mathcal{R}_{YZU}$ which is a subset of a suitably dimensioned Euclidean space.

GIV models place restrictions on a structural function $h : \mathcal{R}_{YZU} \to \mathbb{R}$ which defines the admissible combinations of values of Y and U that can occur at each value z of Z which has support $\mathcal{R}_Z$. Admissible combinations of values of (Y, U) at $Z = z$ are zero level sets of this function, as follows.

$$\mathcal{L}(z; h) = \{(y, u) : h(y, z, u) = 0\}$$

For each value of U and Z we can define a Y-level set

$$\mathcal{Y}(u, z; h) \equiv \{y : h(y, z, u) = 0\}$$

which will be singleton for all u and z in complete models but not otherwise.

GIV models place restrictions on such structural functions and also on a collection of conditional distributions

$$\mathcal{G}_{U|Z} \equiv \{G_{U|Z}(\cdot|z) : z \in \mathcal{R}_Z\}$$

whose elements are conditional distributions of U given $Z = z$ obtained as z varies across the support of Z. $G_{U|Z}(\mathcal{S}|z)$ denotes the probability that $U \in \mathcal{S}$ conditional on $Z = z$ under the law $G_{U|Z}$.

In the context of the auction model the observed endogenous variables, Y, are the ordered final bids, thus.[2]

$$Y_m \equiv B_{m:M}, \qquad m \in \{1, \dots, M\}$$

It is convenient to write the ordered valuations as functions of uniform order statistics. To this end let

$$\tilde{U} \equiv \left(\tilde{U}_1, ..., \tilde{U}_M\right), \qquad \tilde{U}_m \equiv F_z\left(V_m\right), m \in \{1, \dots, M\}, \qquad U \equiv \left(\tilde{U}_{1:M}, \dots, \tilde{U}_{M:M}\right).$$

The components of $\tilde{U}$ are thus identically and independently distributed uniform variates on the unit interval. The components of U order those of $\tilde{U}$ from smallest to largest, with m_{th} component U_m the m_{th} order statistic of $\tilde{U}$. Moreover, by strict monotonicity of F_z, $U_m = F_z\left(V_{m:M}\right)$, equivalently $V_{m:M} = F_z^{-1}\left(U_m\right)$. Independence of the components of $\tilde{U}$ implies that the distribution of U is uniform with constant density on its support $\mathcal{R}_U$, which is that part of the unit M-cube where $U_1 \leq U_2 \leq \cdots \leq U_M$.[3]

The restrictions (2.1) and (2.2) of Lemma 1 can be written as

$$\forall m, \quad Y_m \leq V_{m:M} = F_z^{-1}(U_m) \text{ and } Y_M \geq V_{M-1:M} = F_z^{-1}(U_{M-1})$$

and, on applying the increasing function $F_z(\cdot)$, they are as follows.

$$\forall m, \quad F_z(Y_m) \leq U_m \text{ and } F_z(Y_M) \geq U_{M-1} \tag{3.1}$$

A GIV structural function which expresses these restrictions is as follows.

$$h(Y, z, U) = \sum_{m=1}^{M} \max(\left(F_z(Y_m) - U_m\right), 0) + \max(\left(U_{M-1} - F_z(Y_M)\right), 0) \tag{3.2}$$

Thus we have cast the auction model as a GIV model in which the structural function h is a known functional of the collection of conditional valuation distributions $\{F_z\left(\cdot\right) : z \in \mathcal{Z}\}$. We use

[2] To simplify notation plain subscripts "m" rather than order statistic subscripts "$m : M$" are used for the elements of Y, and shortly, U.

[3] See Section 2.2 in David and Nagaraja (2003). The support of the order statistics U in the unit M-cube has volume $1/M!$ so the constant value of the density is $M!$.

the notation $\mathcal{F}$ to denote a collection of such conditional distribution functions. The restrictions of the auction model on the distribution of (U, Z) are: (i) U and Z are independently distributed and (ii) the distribution of U, denoted G_U, is the joint distribution of the order statistics of M independent uniform variates. This is uniform on the part of the unit M-cube in which $U_1 \leq U_2 \leq \cdots \leq U_M$.

4 Characterizing the identified set of valuation distributions

Central to the GIV analysis in CR are the U-level sets of the structural function defined as follows.

$$\mathcal{U}(y, z; h) = \{u : h(y, z, u) = 0\}$$

For a given value of z this set comprises the values of u that can give rise to a particular value y of Y.

In some econometric models this set is a singleton - the classical linear model is a leading example. There are many econometric models in which U-level sets are not singleton. Examples include models for discrete outcomes and models with more sources of heterogeneity than endogenous outcomes.

In the auction model the U-level sets are:

$$\mathcal{U}(y, z; h) = \left\{ u : \left(\bigwedge_{m=1}^{M} (u_m \geq F_z(y_m)) \right) \wedge (F_z(y_M) \geq u_{M-1}) \right\} \tag{4.1}$$

it being understood that for all m, $u_m \geq u_{m-1}$.[4] These are not singleton sets. Figure 1 illustrates for the 2 bidder case.[5] The U-level set $U((y'_1, y'_2), z; h)$ is the blue rectangle below the 45° line.

Applying Theorem 4 of CR, the identified set of valuation distribution functions $F_z(\cdot)$ comprises the set of distribution functions $F_z(\cdot)$ such that for all sets $\mathcal{S}$ in a collection of test sets $\mathsf{Q}(h, z)$ the following inequality is satisfied almost surely

$$G_U(\mathcal{S}) \geq \mathbb{P}\left[\mathcal{U}(Y, Z; h) \subseteq \mathcal{S} | Z = z\right]. \tag{4.2}$$

The collection of test sets $\mathsf{Q}(h, z)$ is defined in Theorem 3 of CR. It comprises certain unions of the members of the collection of U-level sets $\mathcal{U}(y, z; h)$ obtained as y takes values in the conditional support of Y given $Z = z$.[6] The following theorem, proven in the Appendix, provides the formal

[4] If there was a minimum bid increment Δ, then (4.1) would have $F_z(y_M + \Delta)$ in place of $F_z(y_M)$.

[5] The 2 bidder graphs shown in Figures 1 - 4 can be interpreted as projections of U-level sets for the M bidder case, regarding u_2 (u_1) as the value of the largest, U_M, (second largest, U_{M-1}) uniform order statistic.

[6] In general the collection $\mathsf{Q}(h, z)$ contains all sets that can be constructed as unions of sets, (4.1), on the support of the random set $\mathcal{U}(Y, Z; h)$. In particular models some unions can be neglected because the inequalities they deliver are satisfied if inequalities associated with other unions are satisfied. There is more detail and discussion in CR14.

result.

Theorem 1. *Let $\mathcal{F}$ be restricted to F. In the independent private values auction model in which Restrictions 1-3 hold, the set*

$$\mathcal{F}^* \equiv \{\mathcal{F} \in \mathsf{F} : \text{ for all closed } \mathcal{S} \subseteq \mathcal{R}_U,\ G_U(\mathcal{S}) \geq \mathbb{P}[\mathcal{U}(Y,Z;h) \subseteq \mathcal{S}|Z] \text{ a.s. } \mathbb{P}\}, \tag{4.3}$$

comprises bounds on the collection of conditional distributions $\{F_z(\cdot) : z \in \mathcal{Z}\}$.

If, in addition, either Restriction 4 holds or only the order statistics of the bids are observable rather than the bids themselves, then these these bounds are sharp.

In the Theorem above F is a collection of families of conditional distribution functions that embody the researcher's prior information on the distribution functions $F_z(\cdot)$. In a nonparametric auction model F would comprise the class of all strictly increasing cumulative distribution functions for each $z \in \mathcal{Z}$. If parametric restrictions are imposed then F may index the distributions $F_z(\cdot)$ with a finite dimensional parameter vector.

The Theorem states that the collection of such $\mathcal{F}$ satisfying (4.2) for all closed sets $\mathcal{S} \subseteq \mathcal{R}_U$ comprises bounds on the collection of possible valuation distributions conditional on auction characteristics z. If, additionally, bids are exchangeable, or if only the distribution of bid order statistics conditional on $Z = z$ is identified from the data, rather than the distribution of bids, then these bounds are sharp and (4.3) delivers the identified set.

In CR we characterize a sub-family of closed sets on $\mathcal{R}_U$, denoted $\mathsf{Q}(h,z)$ such that if (4.2) holds for all $\mathcal{S} \in \mathsf{Q}(h,z)$, then it must also hold for all closed $\mathcal{S} \subseteq \mathcal{R}_U$. In Section 5 below we consider the form of inequalities generated by particular sets $\mathcal{S}$, and then consider their identifying power in the examples of Sections 6 and 7.

Given a test set $\mathcal{S}$ the probability mass $G_U(\mathcal{S})$ on the left hand side of (4.2) is calculated as $M!$ times the volume of the set $\mathcal{S}$.[7]

The set $\mathcal{U}(Y,Z;h)$ in (4.2) is a random set (Molchanov (2005)) whose realizations are U-level sets as set out in (4.1). Its conditional probability distribution given $Z = z$ is determined by the probability distribution of Y given $Z = z$. In the auction setting this is the conditional distribution of ordered final bids in auctions with $Z = z$. The probability on the right hand side of (4.2) is a conditional containment functional. It is equal to the conditional probability given $Z = z$ that Y lies in the set $\mathcal{A}(\mathcal{S},z;h)$ where

$$\mathcal{A}(\mathcal{S},z;h) \equiv \{y : \mathcal{U}(y,z;h) \subseteq \mathcal{S}\}.$$

[7] This is so because the joint distribution of the uniform order statistics is uniform on the part of the unit M-cube in which $u_1 \leq u_2 \leq \cdots \leq u_M$, with density equal to $M!$.

When $\mathcal{S}$ is a U-level set, say $\mathcal{U}(y', z; h)$ with $y' \equiv (y'_1, \ldots, y'_M)$, there is

$$\mathcal{A}(\mathcal{U}(y', z; h), z; h) = \left\{ y : \left(y_M = y'_M\right) \wedge \left(\bigwedge_{m=1}^{M-1} \left(y_m \geq y'_m\right) \right) \right\} \tag{4.4}$$

it being understood that for all m, $y_m \geq y_{m-1}$.

Figures 2 and 3 show that the effect of changing the value y'_M (which produces the magenta colored rectangles in these Figures) is to produce a new level set that *is not* a subset of $\mathcal{U}(y', z; h)$, hence the equality in (4.4).

Figure 4 shows that the effect of increasing the value y'_{M-1} is to produce a new level set that *is* a subset of $\mathcal{U}(y', z; h)$, hence the weak inequalities in (4.4).

If Y is continuously distributed the equality in (4.4) causes the probability $\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', z; h)|Z = z]$ to be zero and the inequality (4.2) does not deliver an informative bound when $\mathcal{S} = \mathcal{U}(y', z; h)$. So, when bids are continuously distributed, amongst the unions of U-level sets in a collection $\mathsf{Q}(h, z)$, nontrivial bounds are only delivered by unions of a collection of level sets whose members have values of the maximum bid Y_M ranging over a set of values of nonzero measure.

We proceed to consider particular unions of this sort, as follows.

$$\mathcal{U}(y', y''_M, z; h) \equiv \bigcup_{y_M \in [y'_M, y''_M]} \mathcal{U}((y'_1, \ldots, y'_{M-1}, y'_M), z; h), \qquad y''_M \geq y'_M$$

Such unions are termed *contiguous unions* of U-level sets.[8]

The region in $\mathcal{R}_U$ occupied by such a contiguous union is

$$\mathcal{U}(y', y''_M, z; h) = \left\{ u : \left(\bigwedge_{m=1}^{M} \left(u_m \geq F_z(y'_m)\right) \right) \wedge \left(F_z(y''_M) \geq u_{M-1}\right) \right\}$$

it being understood that, for all m, $u_m \geq u_{m-1}$. Figure 5 illustrates for the 2 bidder case. The contiguous union is the region under the 45° line outlined in blue - a rectangle with its top left hand corner removed.

The probability mass placed on this region by the distribution of the uniform order statistics is:

$$G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = M! \int_{F_z(y'_M)}^{1} \int_{F_z(y'_{M-1})}^{\min(u_M, F_z(y''_M))} \int_{F_z(y'_{M-2})}^{u_{M-1}} \cdots \int_{F_z(y'_1)}^{u_2} du. \tag{4.5}$$

[8] A simple U-level set as in (4.1) is obtained on setting $y'_M = y''_M$. When Y_M is not continuously distributed this member of $\mathsf{Q}(h, z)$ may deliver nontrivial bounds.

The set of values of Y that deliver U-level sets that are subsets of $\mathcal{U}(y', y''_M, z; h)$ is

$$\mathcal{A}(\mathcal{U}(y', y''_M, z; h), z; h) = \left\{ y : \left(y'_M \leq y_M \leq y''_M\right) \wedge \left(\bigwedge_{m=1}^{M-1} \left(y_m \geq y'_m\right) \right) \right\}$$

it being understood that $y_1 \leq \cdots \leq y_M$. This region is indicated by the shaded area in Figure 5.[9] The conditional containment functional on the right hand side of (4.2) is calculated as follows.

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[\left(y'_M \leq Y_M \leq y''_M\right) \wedge \left(\bigwedge_{m=1}^{M-1} (Y_m \geq y'_m) \right) \middle| z \right] \tag{4.6}$$

This is a probability that can be estimated using data on values of ordered bids while the probability $G_U\left(\mathcal{U}(y', y''_M, z; h)\right)$ is determined entirely by the chosen values of y', y''_M, z and the distribution function of valuations, F_z, whose membership of the identified set is under consideration.

For any choice of F_z a list of values of (y', y''_M) delivers a list of inequalities on calculating (4.2) and if one or more of the inequalities is violated the candidate valuation distribution F_z is outside the identified set.

The inequalities that arise for particular choices of (y', y''_M) are now explored. The first choices to be considered deliver the inequalities in HT, then other choices are considered which deliver additional inequalities.

5 Inequalities defining the identified set

5.1 Valuations stochastically dominate bids

With $y''_M = +\infty$ and

$$y' = (-\infty, -\infty, \ldots, -\infty, \underbrace{v}_{\text{position } n}, v, \ldots, v)$$

the containment functional probability (4.6) is:

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}[Y_n \geq v|z],$$

the probability mass placed by the distribution G_U on the contiguous union, (4.5), is

$$G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = \mathbb{P}[U_n \geq F_z(v)] = \mathbb{P}[V_n \geq v|z; F_z]$$

[9] The shaded area shows values of $F_z(y_1)$ and $F_z(y_2)$ that give rise to U-level sets that are subsets of $\mathcal{U}(y', y''_M, z; h)$.

and so the condition (4.2) delivers the following inequalities.[10]

$$\forall n, \forall v: \qquad \mathbb{P}[V_n \leq v|z; F_z] \leq \mathbb{P}[Y_n \leq v|z]$$

These inequalities hold for a valuation distribution function F_z if and only if under that distribution there is the stochastic ordering of order statistics of bids and valuations required by the restriction (2.1).

The marginal distribution of the nth order statistic of M identically and independently distributed uniform variates is $Beta(n, M+1-n)$.[11] Let $Q(p; n, M)$ denote the associated quantile function. The restrictions placed on valuation distributions by the inequality (4.2) and the test sets under consideration in this Section are, written in terms of uniform order statistics:

$$\forall n, \forall v: \qquad \mathbb{P}[U_n \leq F_z(v)] \leq \mathbb{P}[Y_n \leq v|z]$$

which can be written as follows.

$$\forall v: \qquad F_z(v) \leq \min_n Q(\mathbb{P}[Y_n \leq v|z]; n, M) \tag{5.1}$$

This continuum of pointwise upper bounds must hold for all valuation distribution functions in the identified set. This is the bound given in Theorem 1 of HT.

Figures 6 and 7 show the contiguous unions of U-level sets (the regions bordered in blue) delivering these inequalities for 2 bidder auctions. The regions shaded blue indicate the values of $(F_z(y_2), F_z(y_1))$ that deliver U-level sets that are subsets of a contiguous union.

5.2 The highest bid stochastically dominates the second highest valuation

With $y''_M = v$ and $y' = (-\infty, -\infty, \ldots, -\infty)$ the containment functional probability (4.6) is:

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[Y_M \leq v|z\right],$$

the probability mass placed by the distribution G_U on the contiguous union, (4.5), is

$$G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = \mathbb{P}[U_{M-1} \leq F_z(v)] = \mathbb{P}[V_{M-1} \leq v|z; F_z]$$

so, with this choice of (y', y''_M), the condition (4.2) delivers the following inequalities.

$$\forall v: \quad \mathbb{P}[V_{M-1} \leq v|z; F_z] \geq \mathbb{P}[Y_M \leq v|z] \tag{5.2}$$

[10] The notation $\mathbb{P}[V_n \leq s|z; F_z]$ serves to remind that V_n is an order statistic of valuations which are identically and independently distributed with conditional distribution function F_z.

[11] See Section 2.3 in David and Nagaraja (2003). The density function of this Beta random variable is proportional to $u^{n-1}(1-u)^{M-n}$.

These inequalities hold for a valuation distribution function F_z if and only if under that distribution the second highest valuation is stochastically dominated by the highest bid as required by the restriction (2.2).

All valuation distribution functions in the identified set must satisfy:

$$\forall v: \qquad \mathbb{P}[U_{M-1} \leq F_z(v)] \geq \mathbb{P}\left[Y_M \leq v|z\right]$$

which is (5.2) rewritten in terms of a uniform order statistic, equivalently[12]

$$\forall v: \qquad F_z(v) \geq Q(\mathbb{P}\left[Y_M \leq v|z\right]; M-1, M). \tag{5.3}$$

This is the bound given in Theorem 2 of HT when the minimum bid increment considered there is set equal to zero.[13]

Figure 8 shows, outlined in blue, the contiguous unions of U-level sets delivering this inequality. The shaded region indicates the values of $(F_z(y_2), F_z(y_1))$ that deliver U-level sets that are subsets of this contiguous union.

5.3 Contiguous unions depending on a single value of y

In the two cases just considered contiguous unions of U-level sets are determined by (y', y''_M) in which a single value, v, of Y appears. The inequalities they deliver place a continuum of pointwise upper and lower bounds on the value of the valuation distribution function, $F_z(v)$, at a value v. When Y is continuously distributed these are the only contiguous unions determined by a single value that deliver nontrivial inequalities.

5.3.1 Bids continuously distributed

To see that this is so, first suppose that y''_M takes some finite value v as in Section 5.2. We must have $y'_M < v$ otherwise the containment functional is zero if Y is continuously distributed. The only possible value for y'_M that does not introduce a second finite value is $-\infty$ and since $y'_m \leq y'_M$ for all M we arrive at the case considered in Section 5.2.

Now suppose a single finite value v determines the vector y'. The only feasible value for y''_M is $+\infty$ because we must have $y''_M > y'_M$ to obtain a nontrivial inequality with Y continuously

[12] There is another expression:

$$\begin{aligned} G_U\left(\mathcal{U}(y', y''_M, z; h)\right) &= M! \int_0^1 \int_0^{\min(u_M, F_z(s))} \int_0^{u_{M-1}} \cdots \int_0^{u_2} du \\ &= M F_z(s)^{M-1} - (M-1) F_z(s)^M, \end{aligned}$$

and the following inequalities for all s: $MF_z(s)^{M-1} - (M-1)F_z(s)^M \geq \mathbb{P}\left[Y_M \leq s|z\right]$.

[13] With a positive minimum bid increment Δ, Y_M is replaced by $Y_M + \Delta$.

distributed. Since the elements of y' must be ordered we arrive at the case considered in Section 5.1.

5.3.2 Bids not continuously distributed

When Y is not continuously distributed the case with $y''_M = v$ and

$$y' = (-\infty, -\infty, \ldots, -\infty, \underbrace{v}_{\text{position } n}, v, \ldots, v)$$

may deliver a nontrivial inequality when $n = M$, but not when $n < M$.

With $n = M$ there is[14]

$$\begin{aligned} G_U\left(\mathcal{U}(y', y''_M, z; h)\right) &= \mathbb{P}[U_M \geq F_z(v) \geq U_{M-1}] \\ &= MF_z(v)^{M-1}(1 - F_z(v)) \end{aligned}$$

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[Y_M = v|z\right]$$

and the condition (4.2) delivers the following inequalities.

$$\forall v: \quad MF_z(v)^{M-1}(1 - F_z(v)) \geq \mathbb{P}\left[Y_M = v|z\right]$$

With $n < M$, $G_U\left(\mathcal{U}(y', y''_M, z; h)\right)$ is zero because, in the set $\mathcal{U}(y', y''_M, z; h)$ under consideration, $U_{M-1} = F_z(v)$ and U_{M-1} is the second largest uniform order statistic which is continuously distributed. When Y is not continuously distributed the probability

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[\bigwedge_{m=n}^{M} Y_M = v \,\middle|\, z\right] \tag{5.4}$$

may be positive in which case the inequality (4.2) is violated whatever candidate distribution F_z is considered.

This violation for any F_z arises because a nonzero probability (5.4) with $n < M$ cannot occur under the restrictions of the model. This is so because if for some v, $Y_{M-1} = Y_M = v$ with nonzero probability then under the restrictions of the model the second largest valuation, $V_{M-1:M}$, is equal

[14] The expression for $G_U\left(\mathcal{U}(y', y''_M, z; h)\right)$ is obtained as:

$$\begin{aligned} \mathbb{P}[U_M \geq F_z(v) \geq U_{M-1}] &= M! \int_{F_z(v)}^{1} \int_0^{F_z(v)} \int_0^{u_{M-1}} \cdots \int_0^{u_2} du \\ &= M! \int_{F_z(v)}^{1} \int_0^{F_z(v)} \frac{u_{M-1}^{M-2}}{(M-2)!} du_{M-1} du_M \end{aligned}$$

which delivers the result as stated.

to v with nonzero probability which violates the requirement that valuations are continuously distributed.

5.4 Contiguous unions depending on two values of y

The bounds (5.1) and (5.3) are the bounds developed in HT. We now show that valuation distribution functions in the identified set are subject to additional restrictions. To do this we turn to inequalities delivered by the containment functional inequality (4.2) applied to test sets $\mathcal{S}$ which are contiguous unions of U-level sets characterized by *two* values of Y.

There are just two types of contiguous union of U-level sets that are determined by two values of Y, v_1 and v_2.

5.4.1 Case 1

In this case: $y''_M = +\infty$ and, with $v_1 \geq v_2$,

$$y' = (-\infty, \ldots, -\infty, \underbrace{v_2}_{\text{position } n_2}, \ldots, v_2, \underbrace{v_1}_{\text{position } n_1}, \ldots, v_1).$$

The containment functional probability (4.6) is

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[\bigwedge_{m=n_1}^{M} Y_m \geq v_1 \wedge \bigwedge_{m=n_2}^{n_1-1} Y_m \geq v_2 \middle| z\right]$$

while (4.5) delivers

$$G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = \mathbb{P}\left[\bigwedge_{m=n_1}^{M} U_m \geq F_z(v_1) \wedge \bigwedge_{m=n_2}^{n_1-1} U_m \geq F_z(v_2)\right]$$

which, plugged into (4.2) delivers inequalities which must be satisfied by all valuation distribution functions in the identified set for all $n_2 < n_1 \leq M$ and all $v_1 \geq v_2$.

As an example, the inequalities obtained with $n_1 = M$ and $n_2 = M - 1$ for which

$$y' = (-\infty, \ldots, -\infty, v_2, v_1)$$

are as follows.

$$\forall v_1 \geq v_2: \qquad 1 - F_z(v_1)^M - MF_z(v_2)^{M-1} + MF_z(v_1)F_z(v_2)^{M-1} \geq \mathbb{P}\left[Y_M \geq v_1 \wedge Y_{M-1} \geq v_2 | z\right] \quad (5.5)$$

These inequalities must be satisfied by all valuation distribution functions F_z in the identified set.

5.4.2 Case 2

In this case: $y''_M = v_1$ and, with $v_1 > v_2$,

$$y' = (-\infty, \ldots, -\infty, \underbrace{v_2}_{\text{position } n}, \ldots, v_2).$$

The containment functional probability (4.6) is

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[(v_2 \leq Y_M \leq v_1) \wedge \left(\bigwedge_{m=n}^{M-1} (Y_m \geq v_2)\right) \middle| z\right]$$

while (4.5) delivers

$$G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = \mathbb{P}\left[\left(\bigwedge_{m=n}^{M} U_m \geq F_z(v_2)\right) \wedge (F_z(v_1) \geq U_{M-1})\right]$$

which, plugged into (4.2) delivers inequalities which must be satisfied by all valuation distribution functions in the identified set for all $n \leq M$ and all $v_1 > v_2$. As an example, the inequalities obtained with $n = M - 1$ are as follows.

$$\begin{aligned} \forall v_1 > v_2 : \qquad & F_z(v_1)^M - MF_z(v_1)F_z(v_2)^{M-1} + (M-1)F_z(v_2)^M \\ & + M(1 - F_z(v_1))(F_z(v_1)^{M-1} - F_z(v_2)^{M-1}) \geq \\ & \mathbb{P}\left[v_1 \geq Y_M \geq v_2 \wedge Y_{M-1} \geq v_2 | z\right] \quad (5.6) \end{aligned}$$

5.4.3 Discussion

The inequalities presented in this Section, of which (5.5) and (5.6) are examples, may not be satisfied by all valuation distributions which satisfy the HT bounds (5.1) and (5.3). Section 6 presents numerical calculations for two examples in which one or both of the new inequalities are binding.

The new bounds place restrictions on *pairs* of coordinates that can be connected by distribution functions in the identified set.

Contiguous unions of U level sets that are determined by n values of Y place restrictions on n-tuples of coordinates that can be connected by valuation distribution functions in the identified set, and n can be as large as $M + 1$.

There are also test sets which are unions of contiguous unions which are not themselves contiguous unions, so there are potentially restrictions on collections of many more than $M+1$ coordinates that can be connected by valuation distribution functions in the identified set.

6 Bounds in numerical examples

This Section presents graphs of the bounds for two particular joint distributions of ordered bids in 2 bidder auctions. Details of the calculations and the valuation distributions employed are given in an Annex. In both cases the bid and valuation distributions satisfy the conditions of the auction model. The Figures show *survivor* functions, $\bar{F}_z(v) = 1 - F_z(v)$, and bounds on survivor functions.

Figure 9 shows in blue the HT bounds (5.1) and (5.3) on the valuation survivor function for Example 1. The valuation survivor function employed in the example is drawn in red. It is the survivor function of a random variable which is a mixture of two lognormal distributions.

In Figure 10 two valuation values are selected, $v_1 = 14$ and $v_2 = 5$ and the upper and lower bounds on $\bar{F}_z(v)$ are marked by black circles at these two values. Figure 11 shows a unit square within which we can plot possible values of $(\bar{F}_z(14), \bar{F}_z(5))$. In this Figure the HT bounds are shown as a blue rectangle. Since $\bar{F}_z(14) \leq \bar{F}_z(5)$ the pair of ordinates must lie above the 45° line, drawn in orange. The new inequality (5.5) requires ordinate pairs $(\bar{F}_z(14), \bar{F}_z(5))$ to lie above the magenta line which first falls and then increases. This line lies below the blue rectangle and the inequality (5.5) delivers no refinement in this case. The new inequality (5.6) requires ordinate pairs to lie above the red, increasing, line which does deliver a refinement.

Figure 11 shows that valuation survivor functions in the identified set cannot take relatively low values at $v = 5$ and relatively high values at $v = 14$. Survivor functions satisfying the pointwise bounds on their ordinates (5.1) and (5.3) that are relatively flat over this range are excluded from the identified set.

In Example 2 the valuation distribution function is a mixture of normal distributions. This is drawn in red in Figure 12 with bounds drawn in blue. Figure 13 shows two selected valuation values, $v_1 = 12.5$ and $v_2 = 11.5$. The blue rectangle in Figure 14 shows the pointwise bounds on the ordinates $(\bar{F}_z(12.5), \bar{F}_z(11.5))$ which must lie above the orange 45° line since the survivor function is decreasing. The new bounds (5.5) and (5.6) deliver respectively the magenta and red lines in 14. Ordinate pairs $(\bar{F}_z(12.5), \bar{F}_z(11.5))$ of survivor functions in the identified set must lie above both lines. In this example both of the new inequalities serve to refine the pointwise bounds (5.1) and (5.3).

7 Identified sets in a parametric model

The two-bidder example considered in this Section employs a *parametric* model which restricts the distribution of valuations to be *lognormal*, $LN(\mu, \sigma^2)$ where μ and σ^2 are the mean and variance of log valuations. Identified sets for $\theta \equiv (\mu, \sigma)$ are calculated using a probability distribution of ordered bids obtained under the bidding mechanism employed in Example 1 in Section 6 as described in Annex A, but with the probability distribution of valuations simply lognormal, $LN(0, 1)$ rather than mixed lognormal as is used in the calculations reported in Section 6.

The probabilities on the right hand sides of inequalities (5.1), (5.3), (5.5) and (5.6) are calculated from 10^8 simulated two bidder auctions.

The sharp identified set for θ is defined by an uncountable infinity of inequalities. We employ a finite number chosen as follows. A sequence of values of valuations $\mathcal{V} \equiv \{v_1, \ldots, v_N\}$ is generated as a standard lognormal, $LN(0, 1)$, quantile function applied to N values $\{\frac{1}{N+1}, \frac{2}{N+1}, \ldots, \frac{N}{N+1}\}$. For a candidate value of θ the two pointwise inequalities (5.1) and (5.3) are calculated at these N values and the new two-coordinate inequalities (5.6) and (5.5) are calculated at each $v_i > v_j$ with $(v_i, v_j) \in \mathcal{V}$.

We also consider three-coordinate inequalities obtained using contiguous unions with $y''_M = v_1$ and, with $v_1 \geq v_2 \geq v_3$,

$$y' = (-\infty, \ldots, -\infty, v_3, v_2).$$

The containment functional probability (4.6) for these test sets is

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[v_1 \geq Y_M \geq v_2 \wedge Y_{M-1} \geq v_3 |\, z\right]$$

while (4.5) delivers

$$G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = \mathbb{P}\left[U_M \geq F_z(v_2) \wedge F_z(v_1) \geq U_{M-1} \geq F_z(v_3)\right]$$

which is as follows.

$$\begin{aligned} G_U\left(\mathcal{U}(y', y''_M, z; h)\right) = {} & M(1 - F_z(v_1))(F_z(v_1)^{M-1} - F_z(v_3)^{M-1}) \\ & - MF_z(v_3)^{M-1}(F_z(v_1) - F_z(v_2)) + (F_z(v_1)^M - F_z(v_2)^M) \end{aligned}$$

These three-coordinate inequalities are calculated for each triple $(v_i, v_j, v_k) \in \mathcal{V}$ with $v_i > v_j > v_k$.

We compare outer regions of the identified set for θ obtained using (i) the pointwise bounds of HT, (ii) the pointwise and two-coordinate bounds and (iii) the pointwise, two-coordinate and three-coordinate bounds. Calculations were done using a 100×100 grid of values of (μ, σ).

Figure 15 shows the sets obtained with $N = 25$. In this case there are 50 pointwise upper and lower bounds (25 of each), 600 pairwise inequalities, 300 delivered by each of (5.5) and (5.6), and

Table 1: Projections of identified sets: parameters and location measuures. The first row is obtained using only pointwise bounds. The second row additionally employs two-coordinate inequalities and the third row uses additionally three-coordinate inequalities

Inequalities	$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$	$E[V]$	$Median[V]$	$Mode[V]$
pointwise	$[-0.74, 0.53]$	$[0.79, 2.01]$	$[0.63, 1.97]$	$[0.48, 1.70]$	$[0.28, 1.28]$
& two-coordinate	$[-0.74, \mathbf{0.32}]$	$[\mathbf{0.85}, 2.01]$	$[0.63, \mathbf{1.45}]$	$[0.48, \mathbf{1.38}]$	$[0.28, \mathbf{1.24}]$
& three-coordinate	$[-0.74, 0.32]$	$[\mathbf{0.86}, \mathbf{1.99}]$	$[0.63, 1.45]$	$[0.48, 1.38]$	$[0.28, 1.24]$

2300 three-coordinate inequalities.[15]

The region colored blue is the set of values of (μ, σ) obtained using all 2950 inequalities. The union of the regions colored pink and blue is the set obtained using just the pointwise bounds. The region colored pink on the right of the picture (at higher values of μ) is the region excluded by the 600 two-coordinate inequalities. The thin sliver colored pink on the left of the picture at high values of σ is the region further excluded by the 2300 three-coordinate inequalities. On the 100×100 grid of values of (μ, σ) the pointwise HT bounds placed 1475 pairs in the identified set. The two-coordinate inequalities exclude 203 of these pairs. The three-coordinate inequalities exclude a further 7 values. None of the values excluded by the two-coordinate inequalities would be excluded by the three-coordinate inequalities if these were considered alone.

Table 1 shows the projections of the sets for individual parameters, μ and σ, and identified intervals for various measures of the location of the distribution of valuations. Values that change on considering the two- and three-coordinate inequalities in addition to the pointwise bounds are set out in bold font.

The approximate identified sets for the optimal reserve price and the maximal profit under three values of marginal cost are shown in Table 2. The bounds on the optimal reserve price are wide - this echoes the result found in HT. The new two-coordinate inequalities reduce the bounds but not to a great extent. There is some slight further reduction on additionally considering the three-coordinate inequalities. Values that change on considering the new inequalities are set out in bold font. It is possible that the shape restrictions would have greater impact in auctions with more than 2 bidders.

8 Concluding remarks

The incomplete model of English auctions studied in HT falls in the class of Generalized Instrumental Variable models introduced in CR. Results in CR characterize the identified set of structures delivered by a GIV model. Applying the results to the auction model delivers a characterization of

[15] The number of three-coordinate inequalities generated by a sequence, $\mathcal{V}$, of N distinct valuations is $\frac{1}{6}N(N-1)(N-2)$.

Table 2: Identified sets for optimal reserve price and maximised profit obtained using pointwise bounds and additionally using two-coordinate inequalities and additionally using three-coordinate inequalities

marginal cost (v_0)	inequalities	optimal reserve price	maximal profit
0	pointwise	$[1.07, 12.32]$	$[0.31, 0.87]$
	& two-coordinate	$[1.07, 12.32]$	$[0.31, \mathbf{0.77}]$
	& three-coordinate	$[1.07, \mathbf{11.54}]$	$[0.31, 0.77]$
1	pointwise	$[2.32, 16.63]$	$[0.16, 0.65]$
	& two-coordinate	$[\mathbf{2.38}, 16.63]$	$[0.16, 0.65]$
	& three-coordinate	$[2.38, \mathbf{15.75}]$	$[0.16, \mathbf{0.64}]$
2	pointwise	$[\mathbf{3.54}, 20.44]$	$[0.10, 0.61]$
	& two-coordinate	$[\mathbf{3.66}, 20.44]$	$[0.10, 0.61]$
	& three-coordinate	$[3.66, \mathbf{19.48}]$	$[0.10, \mathbf{0.60}]$

the sharp identified set for that model.

The CR development of sharp identified sets uses results from random set theory. CR shows that a structure, comprising a structural function h and distribution of unobservable random variables, G_U, is in the identified set of structures delivered by a model and distribution of observable random variables if and only if G_U is the distribution of a selection[16] of the random U-level set delivered by the structural function h and the distribution of observable values under consideration. The identified set is then characterized using necessary and sufficient conditions for this selectionability property to hold.

The characterization given in CR and applied here to the auction model delivers a complete description of the sharp identified set. This means that *all* structures admitted by the model that can deliver the distribution of ordered bids, and *only* such structures, satisfy the inequalities that comprise the characterization.

HT left the issue of the sharpness of their bounds as an open question.[17] The approach adopted in HT to determining sharpness is a constructive one, effectively searching for admissible bidding strategies which deliver the distribution of final bids used to calculate the bounds for every distribution of valuations in a proposed identified set. As noted in HT[18] this is difficult to carry through in the auction model. Constructive proofs of sharpness have the advantage that they deliver at least one of the many complete, observationally equivalent, specifications of the process under study. However they are frequently hard to obtain. The method set out in CR and applied here has the advantage that sharpness is guaranteed.

[16] A *selection* of a random set is a random variable that lies in the random set with probability one. A probability distribution is *selectionable* with respect to a random set if there exists a selection of the random set which has that probability distribution.

[17] See Section VIII of HT.

[18] See Appendix D of HT.

Gentry and Li (2014) take a constructive approach to proof of sharpness of identified sets in a model of auctions with selective entry. They produce pointwise bounds on the value of a distribution function of valuations at each value of its argument and prove pointwise sharpness. Taking the approach adopted here may lead to shape restrictions which lead to refinement of the identified set of valuation distribution functions.

The new bounds for the auction model exploit information contained in the joint distribution of ordered final bids unlike the pointwise bounds on the levels of valuation distributions which depend only on marginal distributions of ordered bids. This new information is useful because the joint distribution of ordered final bids is informative about the spacing of ordered bids which in turn is informative about the shape of the valuation distribution.

The characterization of the sharp identified set in the English auction model involves a dense system of inequalities. The inequalities restrict not only the level of the valuation distribution function at each point in its support but also the shape of the function as it passes between the pointwise bounds. In practice, with a finite amount of data and computational resource, some selection of inequalities will be required. How to make that selection is an open research question.

Placing bounds on functionals of the distribution of valuations such as optimal reserve price in a nonparametric setting can be achieved using a flexibly parametrized parametric model and the methods employed with the simple parametric lognormal model in Section 7. We think here in terms of a sieve-like approximation to the distribution of valuations with coefficients constrained by the bounds delivered by the GIV model of CR, bounds on functionals of the distribution being obtained by projection.

In research in progress we pursue the development of sharp identified sets for the distribution of valuations when the joint distribution of final bids is not exchangeable or when valuations maybe be affiliated.

References

ARADILLAS-LOPEZ, A., A. GANDHI, AND D. QUINT (2013): “Identification and Inference in Ascending Auctions with Correlated Private Values,” *Econometrica*, 81(2), 489–534.

ARMSTRONG, T. B. (2013): “Bounds in Auctions with Unobserved Heterogeneity,” *Quantitative Economics*, 4, 377–415.

ATHEY, S., AND P. HAILE (2002): “Identification of Standard Auction Models,” *Econometrica*, 70(6), 2107–2140.

CHESHER, A., AND A. ROSEN (2014): “Generalized Instrumental Variable Models,” CeMMAP working paper CWP04/14, formerly CeMMAP working paper CWP43/13.

David, H. A., and H. N. Nagaraja (2003): *Order Statistics, 3rd edition*. John Wiley, Hoboken.

Gentry, M., and T. Li (2014): "Identification in Auctions with Selective Entry," *Econometrica*, 82(1), 315–344.

Haile, P. A., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111(1), 1–51.

Lindvall, T. (1999): "On Strassen's Theorem of Stochastic Dominantion," *Electronic Communications in Probability*, 4, 51–59.

Milgrom, P., and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding.," *Econometrica*, 50(4), 1089–1122.

Molchanov, I. S. (2005): *Theory of Random Sets*. Springer Verlag, London.

Strassen, V. (1965): "The Existence of Probability Measures with Given Marginals," *Annals of Mathematical Statistics*, 46, 423–439.

Tang, X. (2011): "Bounds on Revenue Distributions in Counterfactual Auctions with Reserve Prices," *Rand Journal of Economics*, 42(1), 175–203.

A Proofs of results stated in the main text

Proof of Lemma 1. Consider a realization (b, v) of (B, V). Under Restriction 1 the number of elements of b with values greater than $v_{m:M}$ is at most $M - m$. Therefore in all realizations of (B, V), $b_{m:M} \leq v_{m:M}$ for all m and M, from which (2.1) follows immediately. The second result, (2.2), follows directly from Restriction 2. □

Proof of Theorem 1. Let Restrictions 1-3 hold and let $\mathcal{F} = \{F_z(\cdot) : z \in \mathcal{Z}\}$ be an element of the set defined in (4.3). From Theorem 4 of CR this is so if and only if for each M there exists a random vector $U^* \equiv (U_1^*, ..., U_M^*)$ and a random vector $Y^* = (Y_1^*, ..., Y_M^*)$ on the same probability space such that for almost every $z \in \mathcal{R}_Z$, we have that conditional on $Z = z$, (i) Y^* is distributed $F_{Y|Z}(\cdot|z)$, the identified conditional distribution of ordered bids in the population, (ii) the components of U^* are distributed uniformly on that part of the M dimensional unit cube with $U_1^* \leq \cdots \leq U_M^*$ (iii) for all $m = 1, ..., M$, $Y_m^* \leq F_z^{-1}(U_m^*)$ and $Y_M^* \geq F_z^{-1}\left(U_{M-1}^*\right)$ almost surely.

This establishes that given knowledge of only the distribution of bid order statistics for almost every z, $F_{Y|Z}(\cdot|z)$, but not the unordered bid distribution $F_{B|Z}(\cdot|z)$, (4.3) comprises the identified set for conditional valuation distributions $\{F_z(\cdot) : z \in \mathcal{Z}\}$.

Now suppose the conditional distributions $F_{B|Z}(\cdot|z)$ are identified, a.e. $z \in \mathcal{R}_Z$. Knowledge of $F_{B|Z}(\cdot|z)$ also reveals knowledge of $F_{Y|Z}(\cdot|z)$, so the set defined in (4.3) constitutes valid bounds on the valuation distribution. To complete the proof, it remains to show that this set is sharp if in addition Restriction 4 holds, that is that bids are exchangeable given $Z = z$.

To do this we now show that conditions (i)-(iii) above imply that, for arbitrary choice of M, there exist random M-vectors of bids B^o and valuations V^o satisfying Restrictions 1-4 such that for each $z \in \mathcal{R}_Z$, (a) $B^o \sim F_{B|Z}(\cdot|z)$, (b) $\forall m = 1, ..., M$, $B_m^o \leq V_m^o$ almost surely, and (c) $B_{M:M}^o \geq V_{M-1:M}^o$ almost surely. For the remainder of the proof we fix z at an arbitrary value in $\mathcal{R}_Z$ and proceed implicitly conditioning on $Z = z$. All distributions that follow are to be understood to be conditional on $Z = z$, with the understanding that the steps can be repeated for any choice of $z \in \mathcal{R}_Z$.

Let $\tilde{U}$ be a random M-vector with distribution equal to the distribution of M independent Uniform$(0, 1)$ variates. It follows from (i)-(iii) above that if random vector $B^* \sim F_{B|Z}(\cdot|z)$, then we have that

$$F_V^* \succsim_{sd} F_B^*, \tag{A.1}$$

where $\succsim_{sd}$ denotes first order stochastic dominance, and where F_B^* and F_V^* denote the distributions of

$$\overline{B}^* \equiv (B_{1:M}^*, ..., B_{M:M}^*, -B_{M:M}^*),$$

and

$$\overline{V}^* \equiv \left(F_z^{-1}\left(\tilde{U}_{1:M}\right), ..., F_z^{-1}\left(\tilde{U}_{M:M}\right), -F_z^{-1}\left(\tilde{U}_{M-1:M}\right)\right),$$

respectively.

Let Σ_B and Σ_U be random vectors denoting the rank order of the elements of B^* and $\tilde{U}$, respectively, from smallest to largest. Let σ denote one of the $M!$ particular orderings of $\{1, ..., M\}$. Because B^* and $\tilde{U}$ are each exchangeable, the distribution of their order statistics is invariant with respect to conditioning on any ordering of their components. Thus, the stochastic dominance condition (A.1) extends to that of the conditional distributions of $\overline{B}^*$ and $\overline{V}^*$ conditional on $\Sigma_B = \sigma$ and $\Sigma_U = \sigma$, respectively. Thus we have that

$$F^*_{V|\Sigma_U=\sigma} \succsim_{sd} F^*_{B|\Sigma_B=\sigma},$$

for all possible orderings σ. Conditional on $\Sigma_B = \sigma$, $B^*_{m:M} = B^*_{\sigma_m}$, and conditional on $\Sigma_U = \sigma$, $\tilde{U}_{m:M} = \tilde{U}_{\sigma_m}$.

It then follows from Strassen's (1965) Theorem, see also Lindvall (1999), that we can construct random $(M+1)$-vectors $\overline{B}^o$ and $\overline{V}^o$ on the same probability space as B^* such that for each ordering σ we have that conditional on $\Sigma_B = \sigma$, $\overline{B}^o \leq \overline{V}^o$ almost surely, where

$$\begin{aligned} \overline{B}^o &\sim F^*_{B|\Sigma_B=\sigma}, \quad \overline{V}^o \sim F^*_{V|\Sigma_U=\sigma}, \text{ and} \\ \forall m &= 1, ..., M\colon \quad \overline{B}^o_{m:M} = \overline{B}^o_m, \quad \overline{V}^o_{m:M} = \overline{V}^o_m. \end{aligned}$$

The distributions of $\overline{B}^o$ and $\overline{V}^o$ are such that $\overline{B}^o_{M+1} = -\overline{B}^o_M$ and $\overline{V}^o_{M+1} = -\overline{V}^o_{M-1}$.

Now define random M-vectors B^o and U^o whose components $m = 1, ..., M$ conditional on $\Sigma_B = \sigma$ are given by:

$$B^o_m \equiv \overline{B}^o_{\sigma_m}, \quad U^o_m \equiv F_z\left(\overline{V}^o_{\sigma_m}\right),$$

and define V^o as the random M-vector whose components are

$$V^o_m \equiv F_Z^{-1}\left(U^o_m\right) = \overline{V}^o_{\sigma_m}, \text{ all } m = 1, ..., M.$$

Therefore, by construction,

$$B^o \overset{d}{=} B^*|\Sigma_B = \sigma, \text{ and } U^o \overset{d}{=} \tilde{U}|\Sigma_U = \sigma.$$

The ordering of the elements of B^o and U^o are both σ implying that conditional $\Sigma_B = \sigma$,

$$B^o \leq V^o \text{ almost surely.} \tag{A.2}$$

Also, by construction,

$$\overline{B}^o_{M+1} = -B^o_{M:M} \leq \overline{V}^o_{M+1} = -F_Z^{-1}\left(U^o_{M-1:M}\right) = -V^0_{M-1:M} \text{ almost surely,}$$

equivalently

$$V^o_{M-1:M} \le B^o_{M:M} \text{ almost surely.} \tag{A.3}$$

Since the argument holds conditional on any ordering σ, (A.2) and (A.3) also hold unconditionally.

Finally, unconditional on the particular ordering σ, we have

$$B^o \stackrel{d}{=} B^*, \text{ and } U^o \stackrel{d}{=} \tilde{U},$$

since by exchangeablility of B^* and $\tilde{U}$, the probability of $\Sigma_B = \sigma$ is the same as that of $\Sigma_U = \sigma$, namely $1/M!$. □

B Calculation of bid probabilities in Section 6

The probabilities used in the two examples in Section 6 were produced by simulation using 10^7 independent draws of identically distributed independent pairs of valuations from a valuation distribution and a fully specified stochastic mechanism that delivers final bids given valuations. Distribution functions of ordered bids and the various probabilities that appear in bounds are simply calculated as proportions of simulated ordered bids that meet the required conditions. In the two examples valuations have different distributions and bids are obtained from valuations in different ways.

In Example 1 the valuation distribution is specified as a mixture of two log normal distributions, one $LN(0, 1)$ and the other $LN(2.5, 0.5^2)$ with mixture weights respectively 0.3 and 0.7.[19] In each of the 10^7 simulated auctions two independent realizations of this mixture distribution are sampled and sorted to deliver realizations of ordered valuations.

Independently distributed random variables, Θ_1 and Θ_2 with identical symmetric Beta distributions, expected value 0.5 and standard deviation 0.06, are assigned to respectively the low and high valuation bidders. A fair coin toss determines who bids first.

The first bidder bids 10% of her valuation. If this exceeds the valuation of the other bidder the auction ends and the bid of the low valuation bidder is recorded as zero. Otherwise the auction proceeds and in turn each bidder bids Θ times the bid on the table plus $(1 - \Theta)$ times her valuation with Θ equal to Θ_1, respectively Θ_2, for the low, respectively high, valuation bidder. The auction ends when the bid of the high valuation bidder exceeds the valuation of the other bidder.

In Example 2 ordered valuations are produced as in Example 1 but with the distribution from which the valuations are sampled specified as a mixture of normal distributions, one $N(10, 1)$ the other $N(12.5, 0.5^2)$ with mixture weights 0.5 attached to each distribution. The final bid of a low valuation bidder is calculated as their valuation minus an amount which is the absolute value of a independent realization of a standard normal variable. The final bid of a high bidder is simulated as

[19] $LN(\mu, \sigma^2)$ is a random variable whose logarithm is $N(\mu, \sigma^2)$.

a weighted average of the low and high valuations with the weight given by a realization of a uniform variate with support on $[0, 1]$. The intention in example 2 is just to produce final bid distributions which respect the stochastic dominance conditions of the model. Example 1 by contrast obtains bid distributions by simulating bidding behavior.

Figure 1: The U-level set $\mathcal{U}((y_1', y_2'), z; h)$ containing values of uniform order statistics, $u_2 \geq u_1$, that can give rise to order statistics of bids, (y_1', y_2'). F_z is the distribution function of valuations. As labelled this is for the 2 bidder case. In the M bidder case this shows a projection of a level set with u_2 (u_1) denoting the largest (second largest) order statistic of M i.i.d. uniform variates.

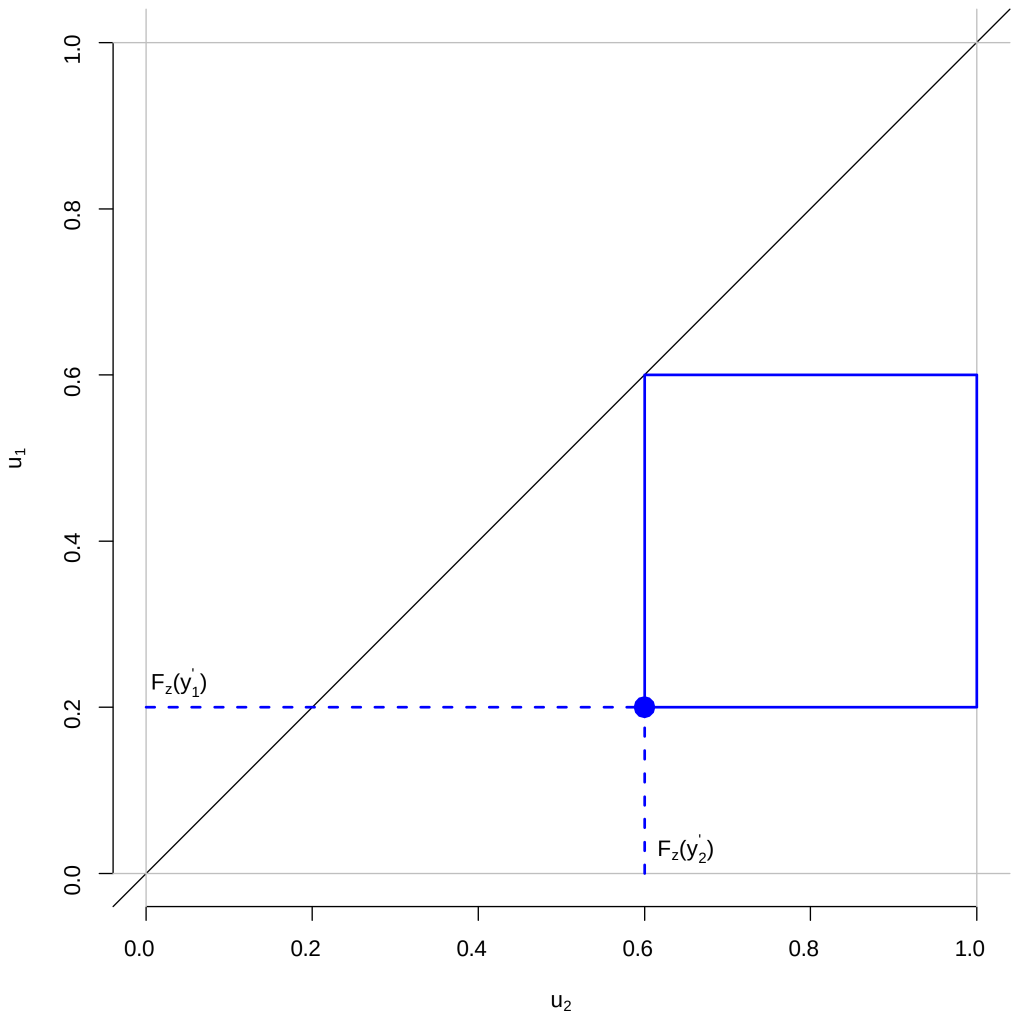

Figure 2: In blue the U-level set $\mathcal{U}((y_1', y_2'), z; h)$ containing values of uniform order statistics, $u_2 \geq u_1$, that can give rise to order statistics of bids, (y_1', y_2'). F_z is the distribution function of valuations. In magenta the U-level set obtained as y_2' is reduced as shown by the arrow. This is never a subset of the original U-level set outlined in blue. As labelled this is for the 2 bidder case.

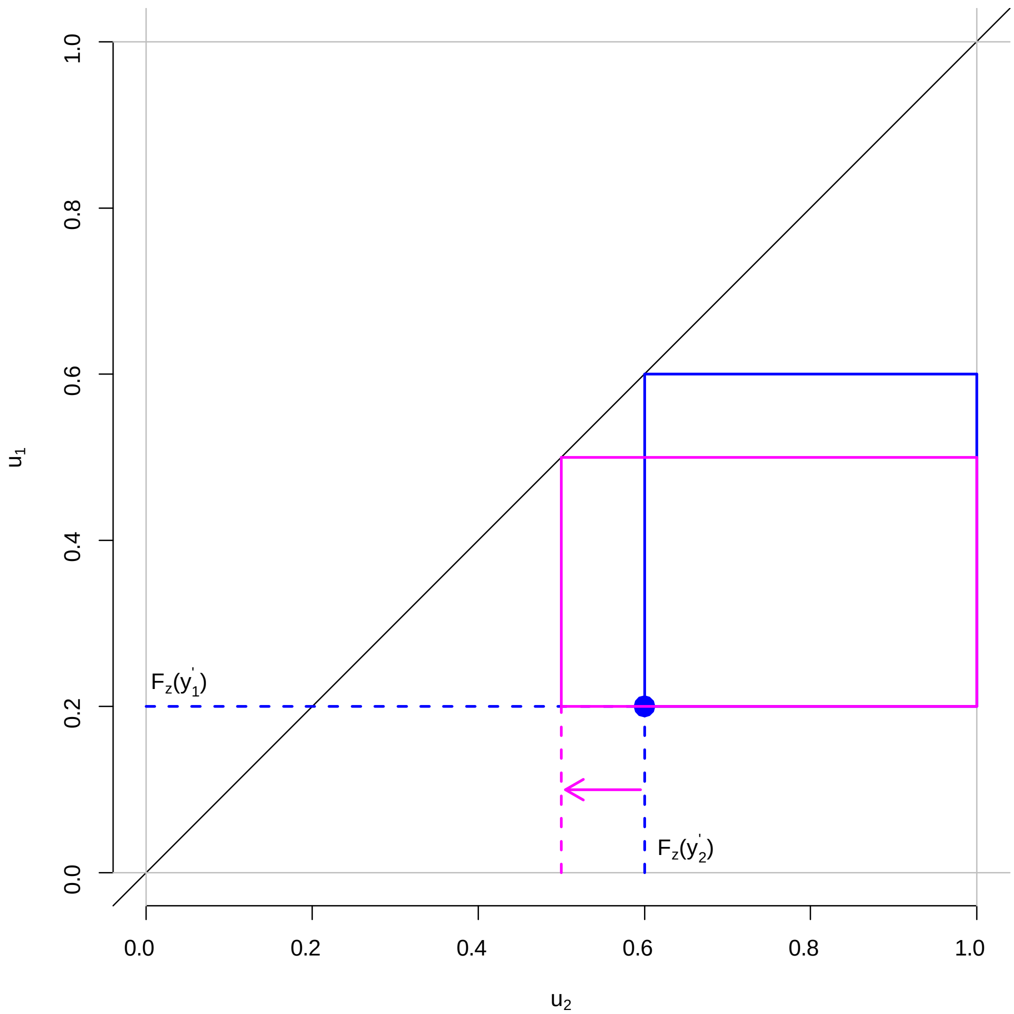

Figure 3: In blue the U-level set $\mathcal{U}((y_1', y_2'), z; h)$ containing values of uniform order statistics, $u_2 \geq u_1$, that can give rise to order statistics of bids, (y_1', y_2'). F_z is the distribution function of valuations. In magenta the U-level set obtained as y_2' is increased as shown by the arrow. This is never a subset of the original U-level set outlined in blue. As labelled this is for the 2 bidder case.

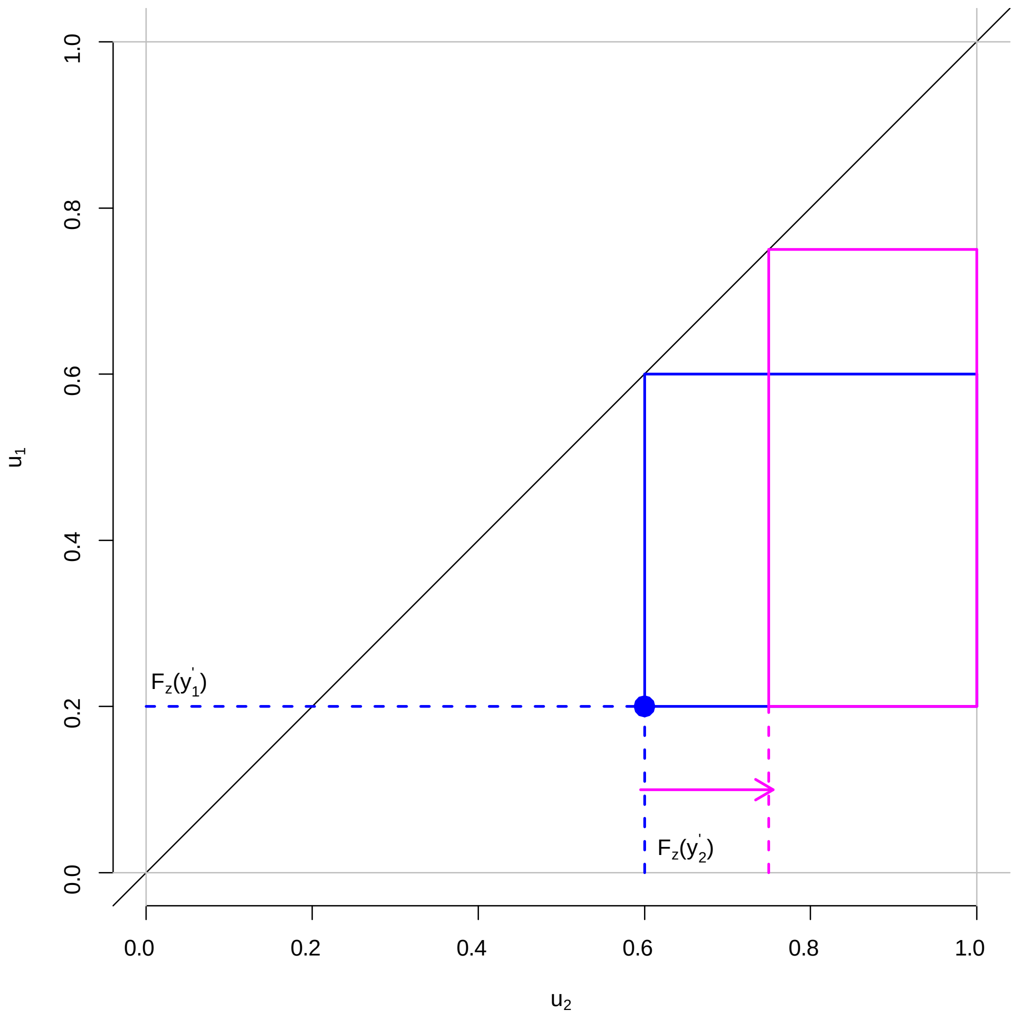

Figure 4: In blue the U-level set $\mathcal{U}((y_1', y_2'), z; h)$ containing values of uniform order statistics, $u_2 \geq u_1$, that can give rise to order statistics of bids, (y_1', y_2'). F_z is the distribution function of valuations. In magenta the U-level set obtained as y_1' is increased as shown by the arrow. This is always a subset of the original U-level set. As labelled this is for the 2 bidder case.

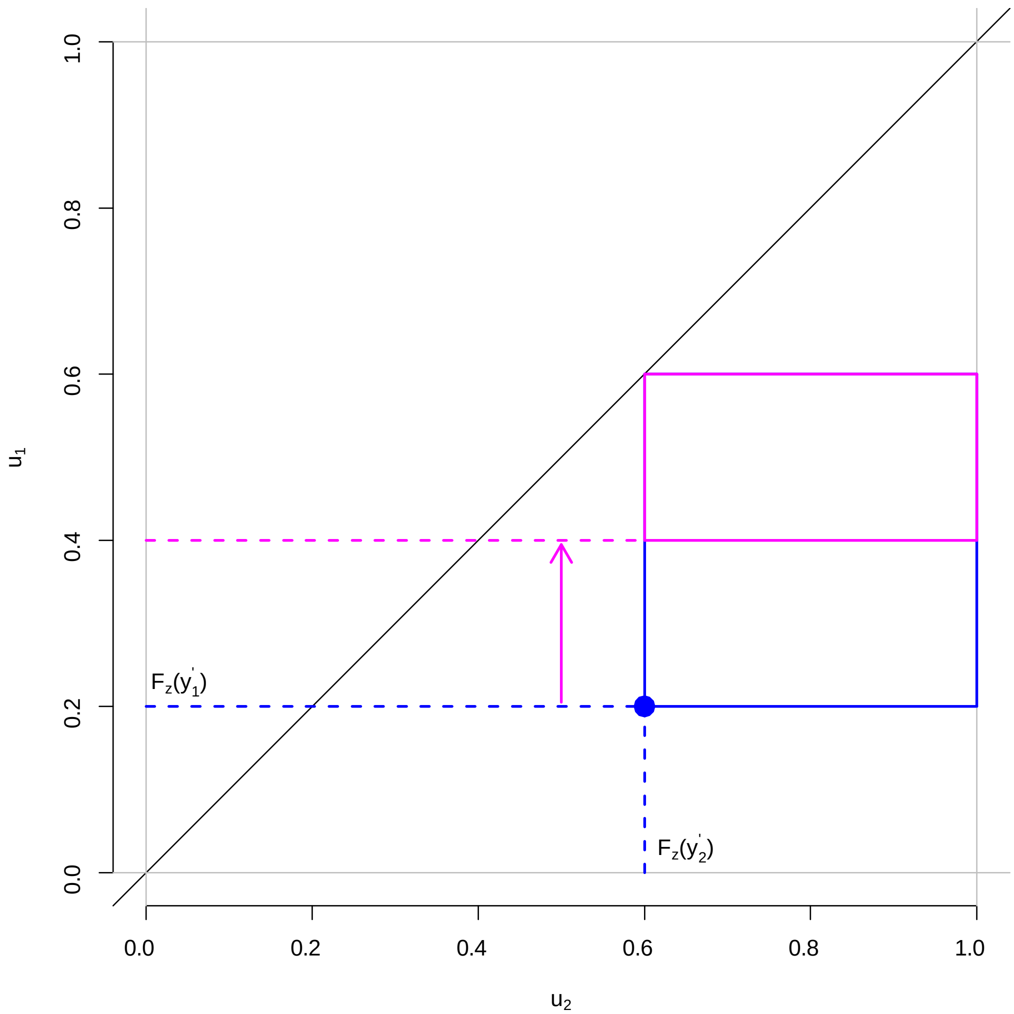

Figure 5: 2 bidder case. The contiguous union of level sets: $\mathcal{U}(y', y''_2, z; h)$ where $y' = (y'_1, y'_2)$, y'_2 and y''_2 are values taken by the maximal order statistic of bids, Y_2, and y'_1 is a value taken by the second largest order statistic of bids, Y_1. In the labels, F_z is the distribution function of valuations. The shaded area indicates the values of Y that give a U-level set which is a subset of the contiguous union.

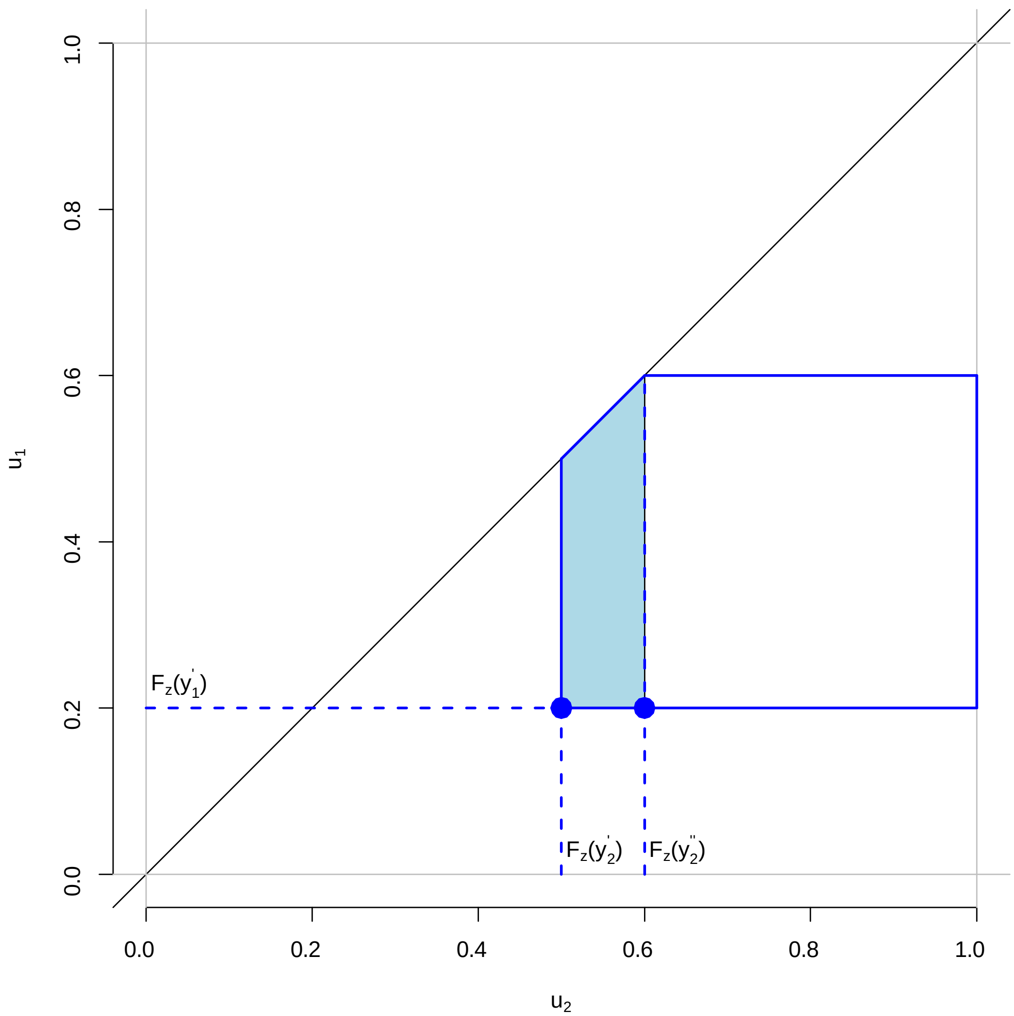

Figure 6: 2 bidder case. The triangular region outlined in blue is the contiguous union of level sets: $\mathcal{U}(y', y''_2, z; h)$ where $y'_1 = y'_2 = v$ and $y''_2 = \infty$. F_z is the distribution function of valuations. This choice of y' and y''_2 delivers the inequality requiring the second highest valuation to stochastically dominate the second highest bid. The shaded area indicates the values of Y that give a U-level set which is a subset of the contiguous union.

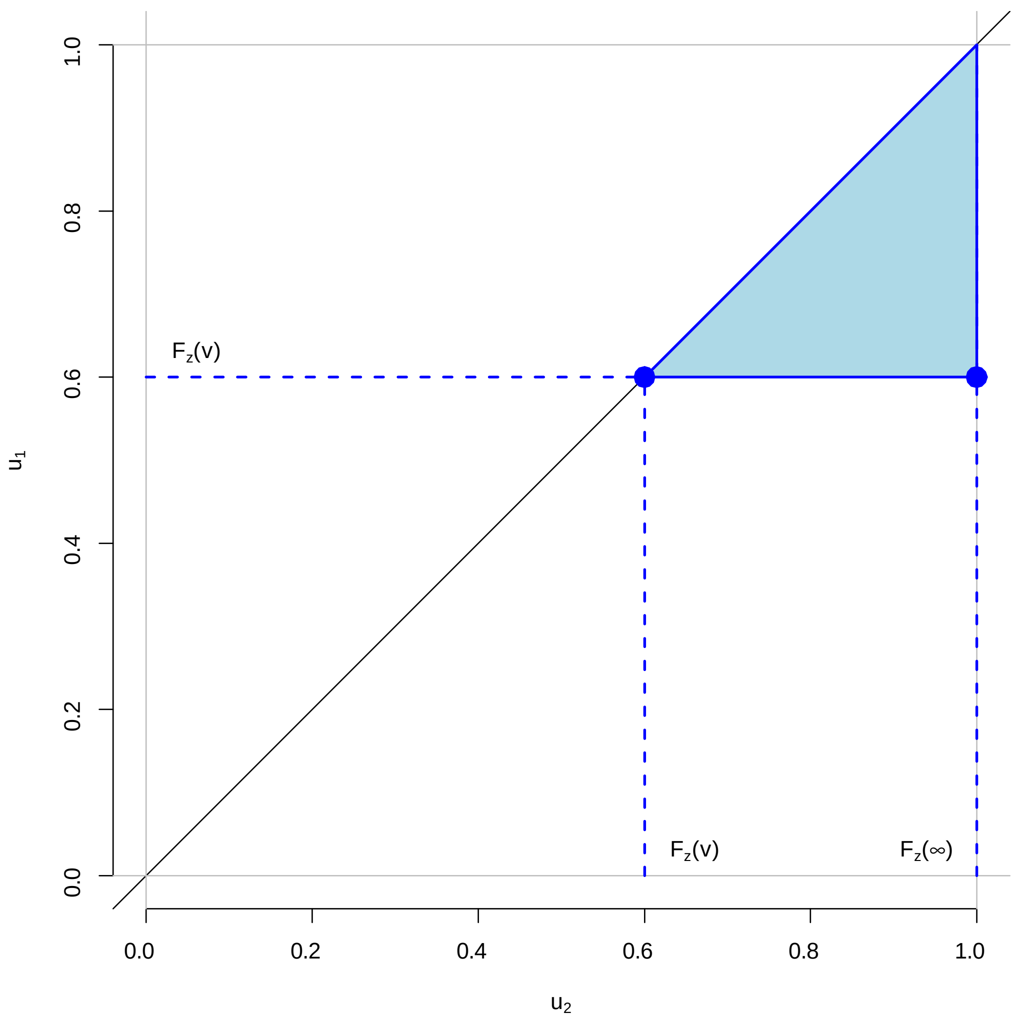

Figure 7: 2 bidder case. The trapezoidal region outlined in blue is the contiguous union of level sets: $\mathcal{U}(y', y_2'', z; h)$ where $y' = (-\infty, v)$ and $y_2'' = \infty$. F_z is the distribution function of valuations. This choice of y' and y_2'' delivers the inequality requiring the highest valuation stochastically dominates the highest bid. The shaded area indicates the values of Y that give a U-level set which is a subset of the contiguous union.

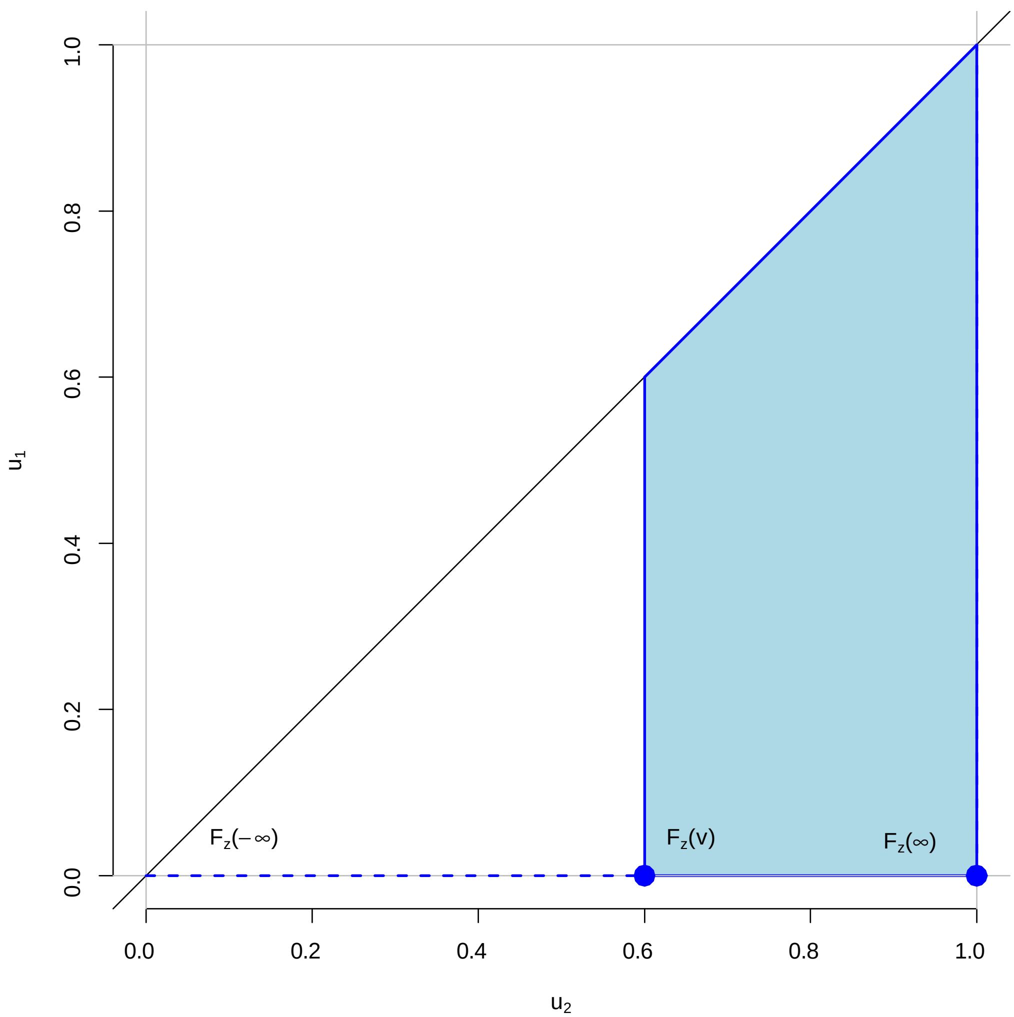

Figure 8: 2 bidder case. The trapezoidal region outlined in blue is the contiguous union of level sets: $\mathcal{U}(y', y_2'', z; h)$ where $y' = (-\infty, -\infty)$ and $y_2'' = v$. F_z is the distribution function of valuations. This choice of y' and y_2'' delivers the inequality requiring the highest bid stochastically dominates the second highest valuation. The shaded area indicates the values of Y that give a U-level set which is a subset of the contiguous union.

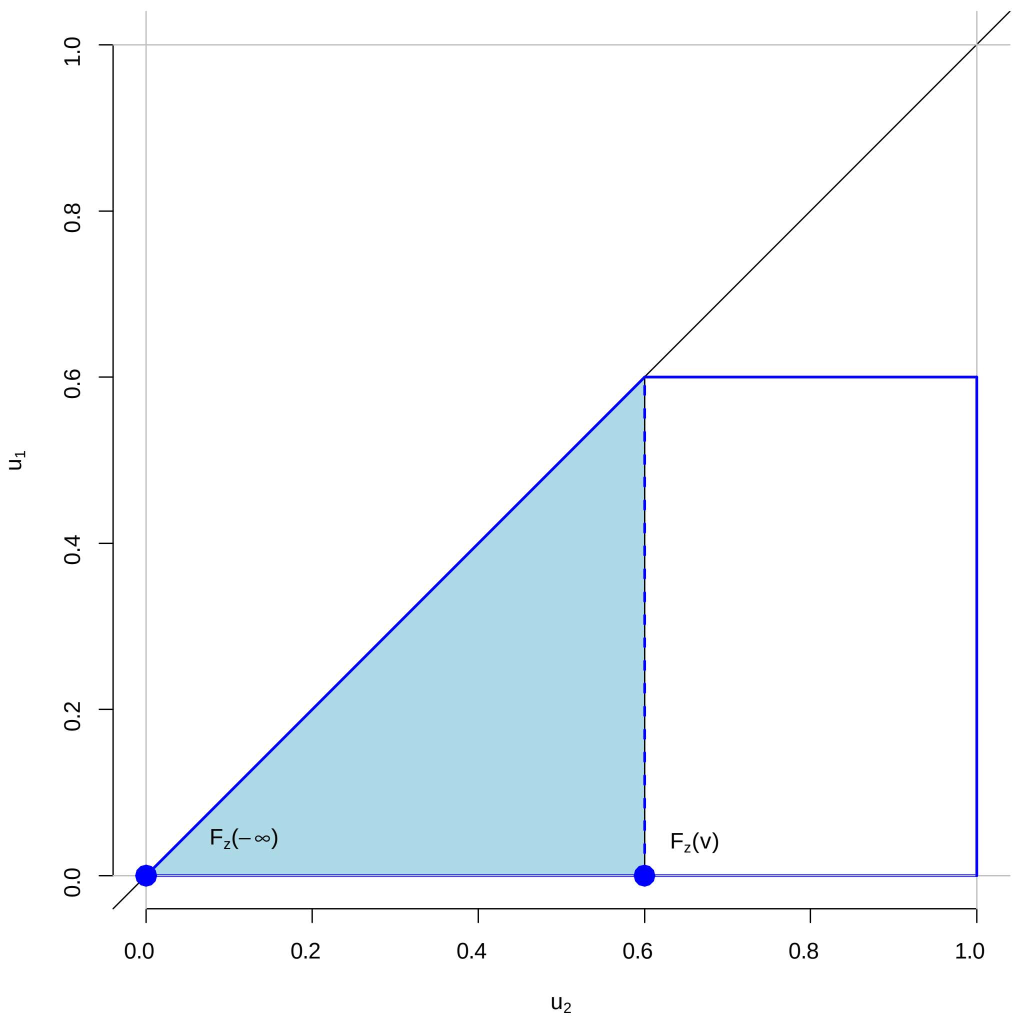

Figure 9: Example 1. Upper and lower bounds (blue) on the valuation survivor function (red).

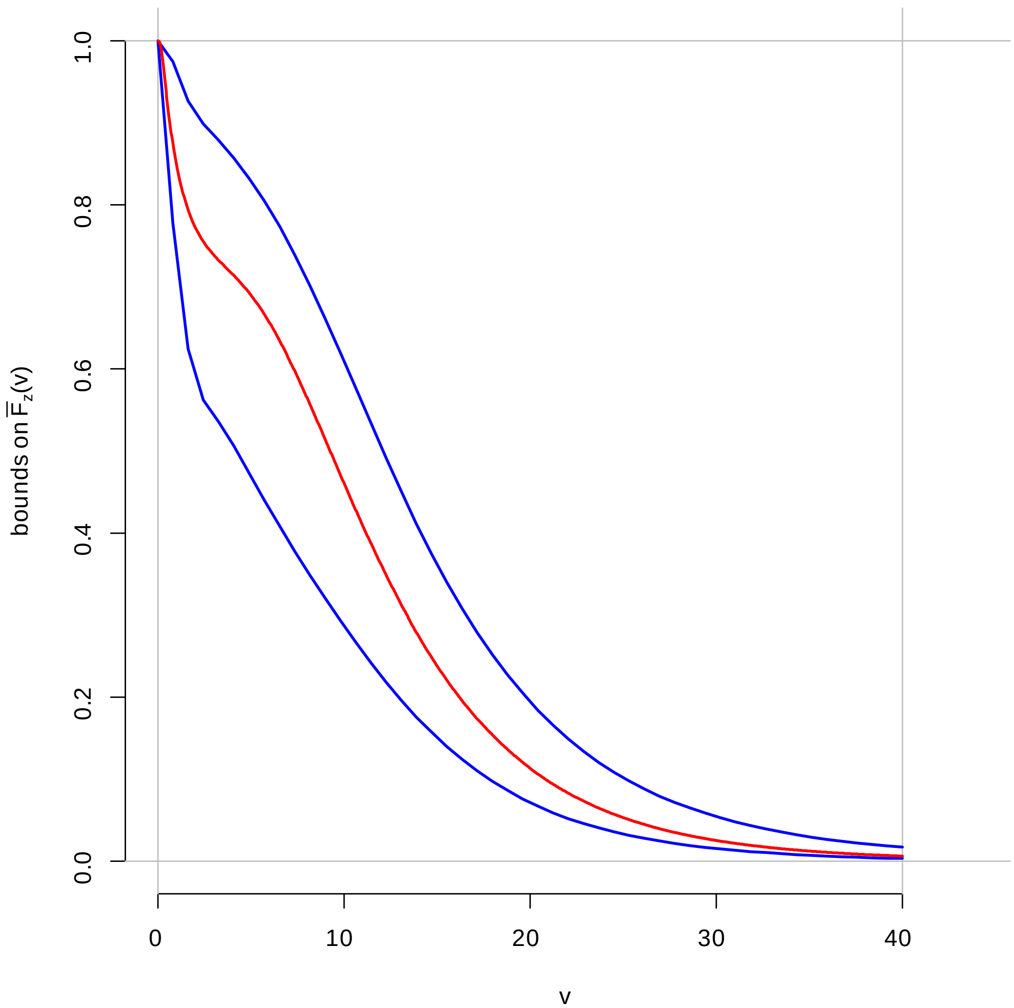

Figure 10: Example 1. Upper and lower bounds (blue) on the valuation survivor function (red). Two values of v, $v_1 = 14$ and $v_2 = 5$ are identified.

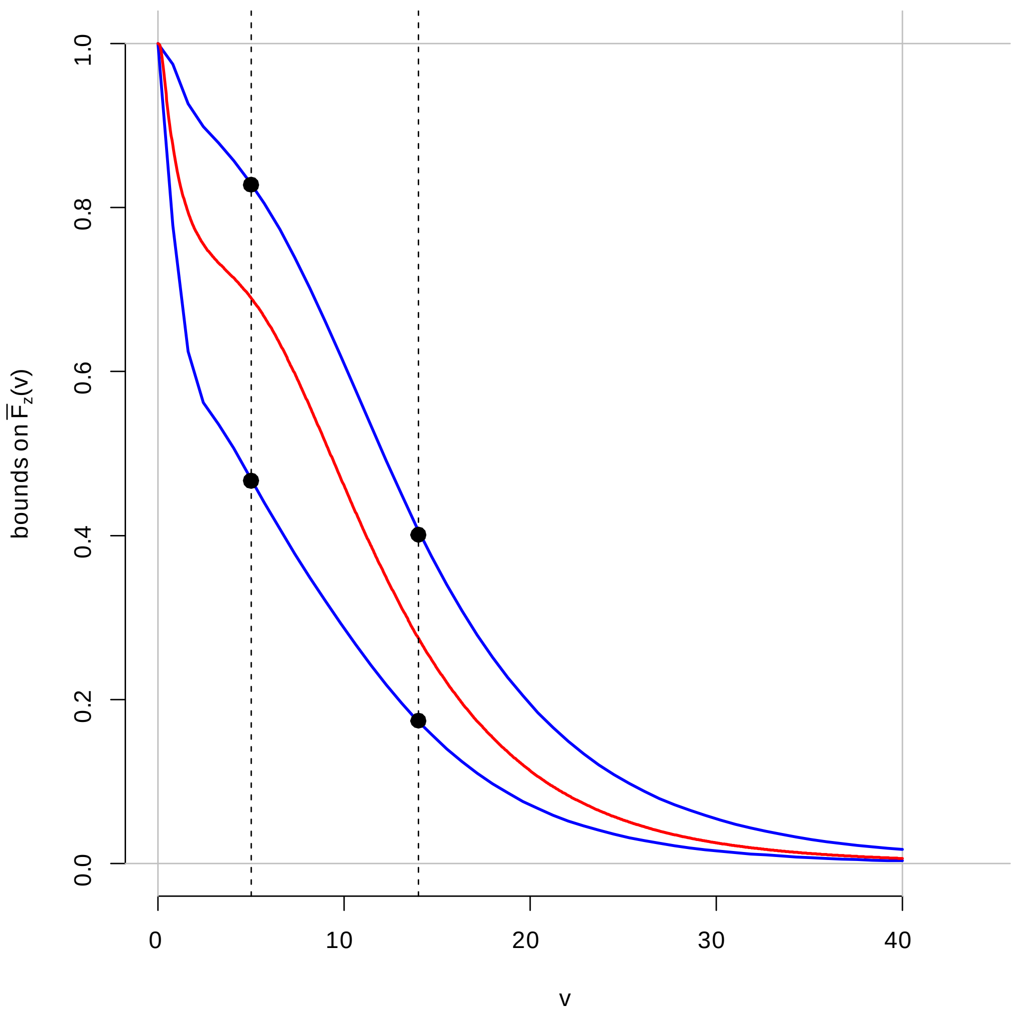

Figure 11: Example 1. The blue rectangle shows upper and lower bounds on $\bar{F}_z(v_1)$ and $\bar{F}_z(v_2)$ at $v_1 = 14$ and $v_2 = 5$. These ordinates of the valuation survivor function must lie above the 45° line (orange). The new bounds require they lie above the magenta and red lines as well. Only the red line delivered by inequality 5.6 is binding.

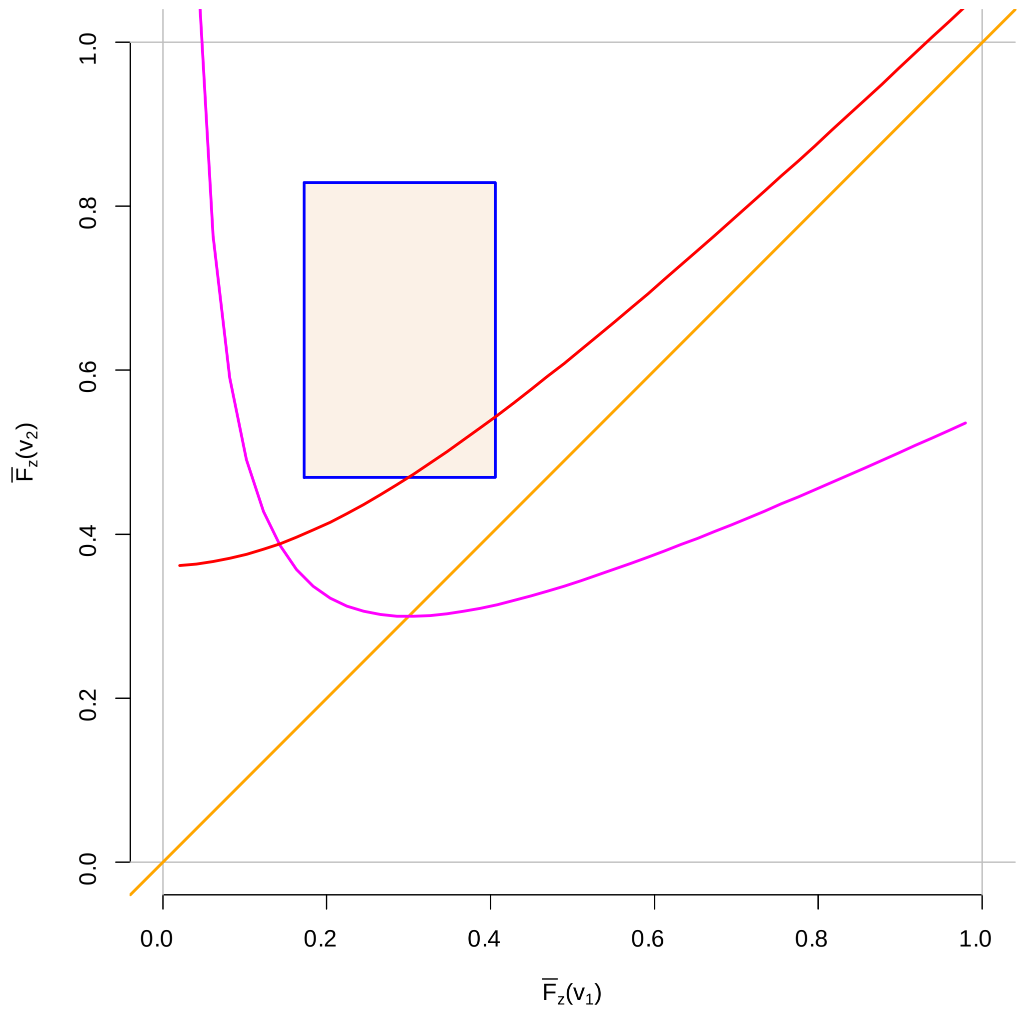

Figure 12: Example 2. Upper and lower bounds (blue) on the valuation survivor function (red).

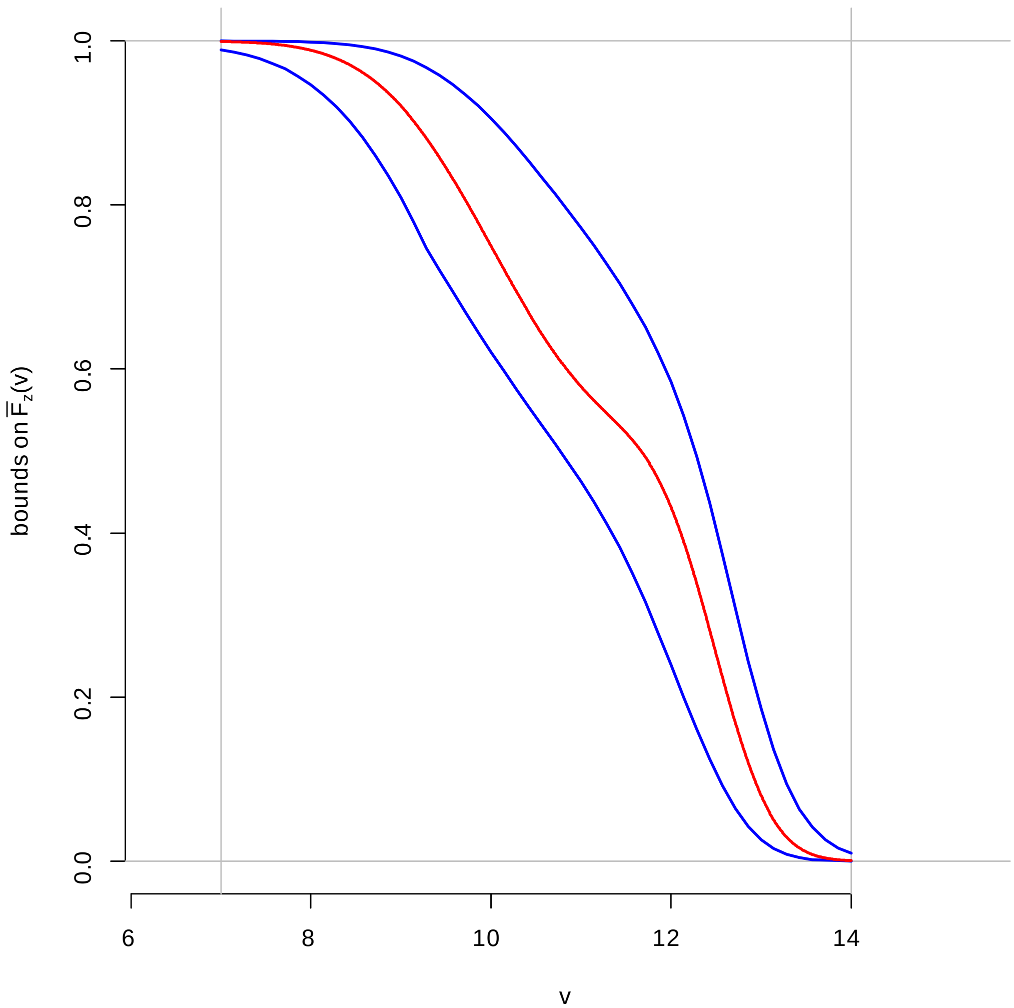

Figure 13: Example 2. Upper and lower bounds (blue) on the valuation survivor function (red). Two values of v are identified: $v_1 = 12.5$ and $v_2 = 11.5$.

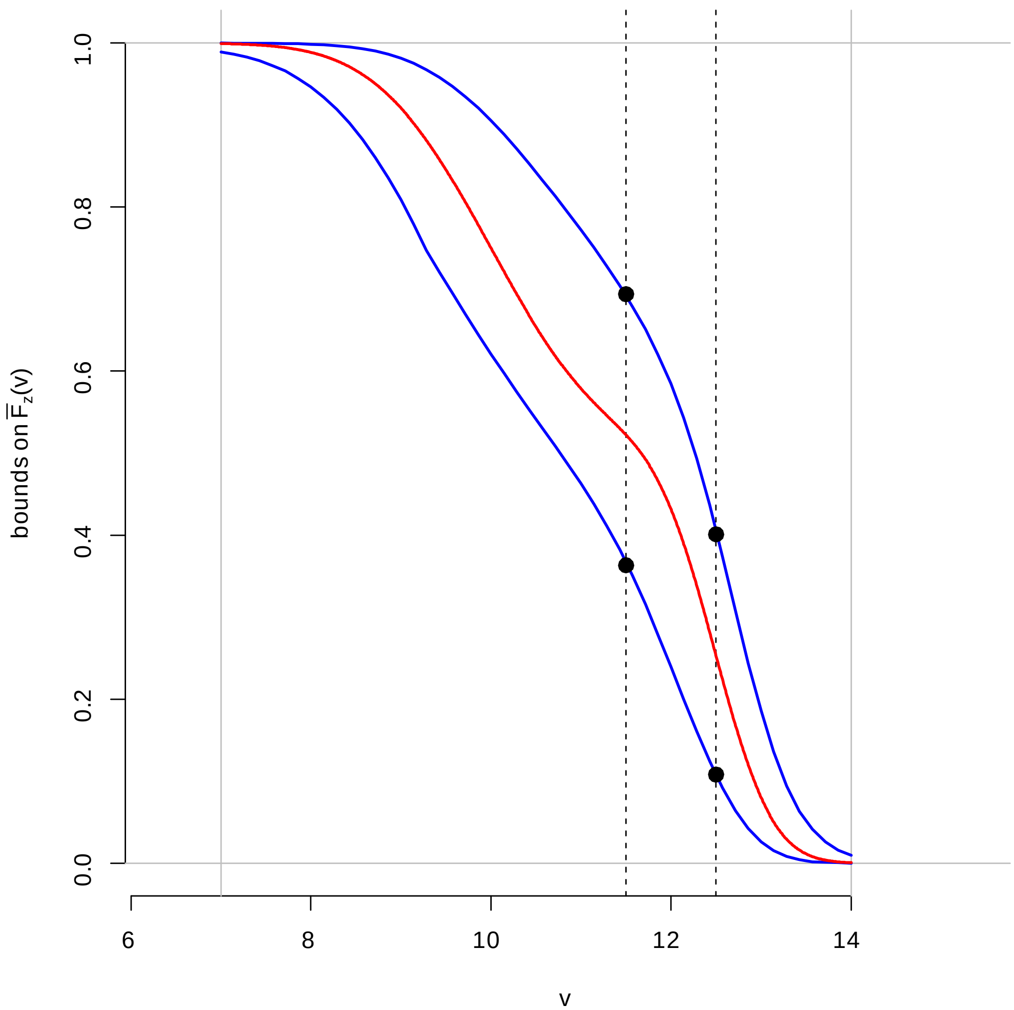

Figure 14: Example 2. The blue rectangle shows upper and lower bounds on $\bar{F}_z(v_1)$ and $\bar{F}_z(v_2)$ at $v_1 = 12.5$ and $v_2 = 11.5$. These ordinates of the valuation survivor function must lie above the 45° line (orange). The new bounds (5.5) and (5.6) require they lie above the magenta and red lines as well.

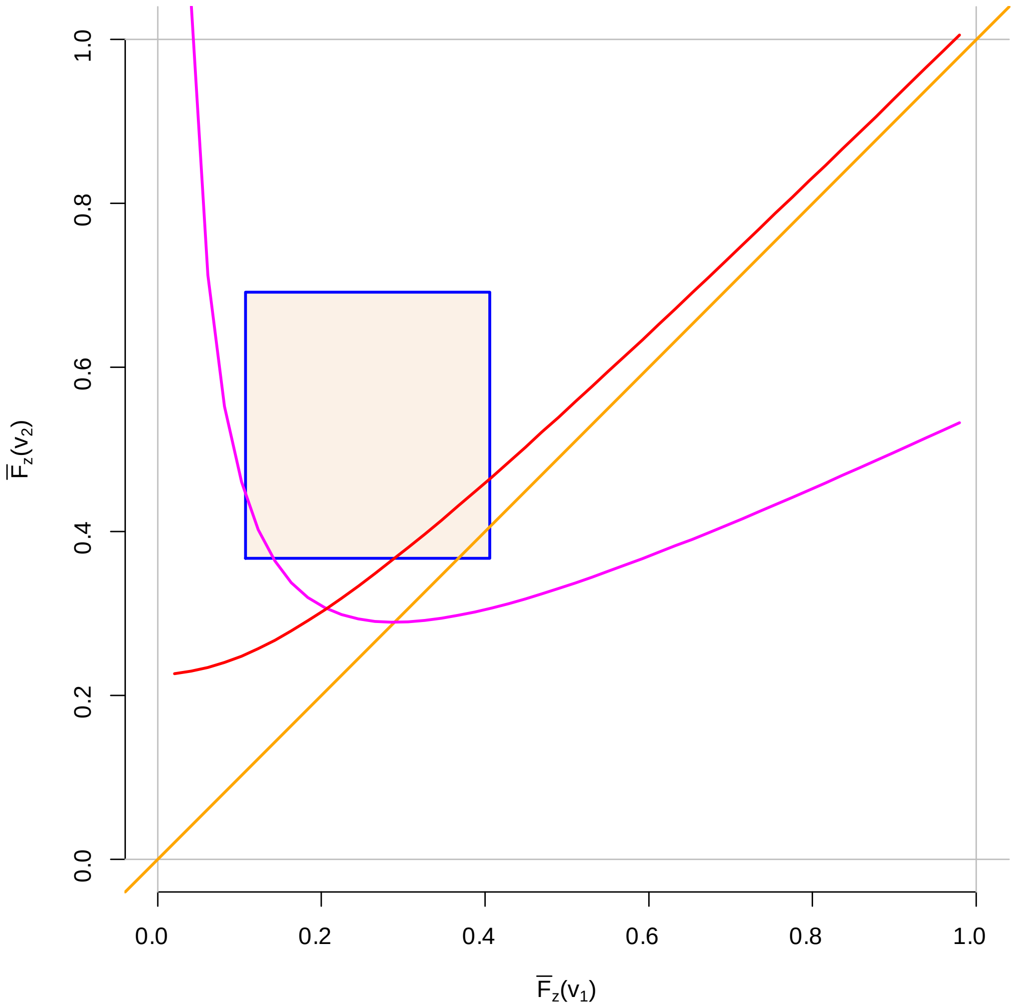

Figure 15: Outer regions for lognormal valuation distribution parameters μ and σ (mean and standard deviation of log valuations). The union of the two filled regions (blue and pink) is the identified set obtained using the HT pointwise upper and lower bounds. The **lower** pink region is excluded by the inequalities (5.5) and (5.6) and the upper pink region is excluded by new inequalities involving three values of V. This leaves just the filled blue region as the approximate identified set. The green dot marks the value of (μ, σ) used to generate the probability distribution of valuations employed in this example.