

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

Andrew Chesher and Adam Rosen

13th February 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 setting.

One innovation in the paper was the use of an incomplete model based on weak plausible restrictions on bidder behaviour, 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.

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 derives 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.

The restrictions of the model 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 Section 7 concludes.

## 2 Model

This section sets out the restrictions of the auction 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.<sup>1</sup>

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

---

<sup>1</sup>These conditions are also imposed in Appendix D of HT03 in which the sharpness of identified sets is discussed.

of bidders,  $M$ , could be an element of  $Z$ .

The  $m$ th largest value in an  $M$  element list of numbers  $x = (x_1, \dots, x_M)$  or random variables  $X = (X_1, \dots, X_M)$  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  $B = (B_1, \dots, B_M)$  and  $V = (V_1, \dots, V_M)$  such that for all  $m = 1, \dots, M$ ,  $B_m \leq V_m$  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.** In an auction for which  $Z = z$  the valuations of bidders are identically and independently continuously distributed with conditional distribution function given  $Z = z$  denoted by  $F_z(\cdot)$ .

Each auction is thus 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  $(\Omega, \mathcal{F}, \mathbb{P})$  with sigma algebra  $\mathcal{F}$  endowed with the Borel sets on  $\Omega$ . 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  $B_m \leq V_m$  in Restriction 2 and those stated in Lemma 1 below are to be understood to mean these inequalities hold  $\mathbb{P}$  almost surely.

The following inequalities involving the order statistics of  $B$  and  $V$  set out in Lemma 1 are a consequence of Restrictions 1 and 2.

**Lemma 1.** For all  $m$  and  $M$

$$B_{m:M} \leq V_{m:M} \tag{1}$$

$$B_{M:M} \geq V_{M-1:M}. \tag{2}$$

**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 or equal to  $v_{m:M}$  is at most  $m$  almost surely. Therefore, for almost every realization of  $(B, V)$ ,  $b_{m:M} \leq v_{m:M}$  for all  $m$  and  $M$ , from which (1) follows. The second result (2) restates Restriction 2.

In similar manner to HT, our identification analysis is based on the restrictions (1) and (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  $h : \mathcal{R}_{YZU} \rightarrow \mathbb{R}$  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.<sup>2</sup>

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

---

<sup>2</sup>To simplify notation plain subscripts “ $m$ ” rather than order statistic subscripts “ $m : M$ ” are used for the elements of  $Y$ , and shortly,  $U$ .

It is convenient to write the ordered valuations as functions of uniform order statistics. To this end let  $\tilde{U}$  be  $M$  identically and independently distributed uniform variates with support on the unit  $M$ -cube,  $[0, 1]^M$  and define  $U = (U_1, \dots, U_M)$  as the order statistics of  $\tilde{U}$  with  $U_M \equiv \max(\tilde{U})$  and so forth.<sup>3</sup> The joint distribution of  $U$  is uniform, with constant density on its support  $\mathcal{R}_U$  which is the part of the unit  $M$ -cube where  $U_1 \leq U_2 \leq \dots \leq U_M$ .<sup>4</sup>

The  $M$  valuations can be written as functions of  $\tilde{U}$ , thus

$$\forall m, \quad V_m = F_z^{-1}(\tilde{U}_m),$$

where  $F_z^{-1}(\cdot)$  is the inverse function of  $F_z$ . This is the quantile function of the valuation distribution. Since  $F_z^{-1}$  is strictly increasing the ordering of  $\tilde{U}$  is the same as the ordering of  $V$  so the order statistics of the  $M$  valuations are related to the uniform order statistics,  $U$ , as follows.

$$\forall m, \quad V_{m:M} = F_z^{-1}(U_m)$$

The restrictions (1) and (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.

$$\forall m, \quad F_z(Y_m) \leq U_m \text{ and } F_z(Y_M) \geq U_{M-1} \quad (3)$$

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

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

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  $U_1 \leq U_2 \leq \dots \leq U_M$ .

---

<sup>3</sup> $\tilde{U}$  is to be understood to reside on the common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  without loss of generality.

<sup>4</sup>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  $M!$ .

## 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:

$$\mathcal{U}(y, z; h) = \left\{ u : \left( \bigwedge_{m=1}^M (u_m \geq F_z(y_m)) \right) \wedge (F_z(y_M) \geq u_{M-1}) \right\} \quad (4)$$

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

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

$$G_U(\mathcal{S}) \geq \mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{S} | Z = z] \quad (5)$$

The collection of test sets  $\mathcal{Q}(h, z)$  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  $Z = z$ .<sup>6</sup>

Given a test set  $\mathcal{S}$  the probability mass  $G_U(\mathcal{S})$  on the left hand side of (5) is calculated as  $M!$  times the volume of the set  $\mathcal{S}$ .<sup>7</sup>

<sup>5</sup>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.

<sup>6</sup>In general the collection  $\mathcal{Q}(h, z)$  contains all sets that can be constructed as unions of sets, (4), 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.

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

The set  $\mathcal{U}(Y, Z; h)$  in (5) is a random set (Molchanov (2005)) whose realizations are  $U$ -level sets as set out in (4). 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 (5) 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

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

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

$$\mathcal{A}(\mathcal{U}(y', z; h), z; h) = \left\{ y : (y_M = y'_M) \wedge \left( \bigwedge_{m=1}^{M-1} (y_m \geq y'_m) \right) \right\} \quad (6)$$

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 (6).

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 (6).

If  $Y$  is continuously distributed the equality in (6) causes the probability  $\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', z; h) | Z = z]$  to be zero and the inequality (5) 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  $\mathcal{Q}(h, z)$ , 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.

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

Such unions are termed *contiguous unions* of  $U$ -level sets.<sup>8</sup>

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

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

---

<sup>8</sup>A simple  $U$ -level set as in (4) is obtained on setting  $y'_M = y''_M$ . When  $Y_M$  is not continuously distributed this member of  $\mathcal{Q}(h, z)$  may deliver nontrivial bounds.

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:

$$G_U(\mathcal{U}(y', y''_M, z; h)) = M! \int_{F_z(y'_M)}^1 \int_{F_z(y'_{M-1})}^{\min(u_M, F_z(y''_M))} \int_{F_z(y'_{M-2})}^{u_{M-1}} \cdots \int_{F_z(y'_1)}^{u_2} du. \quad (7)$$

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

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

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

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h) | z] = \mathbb{P} \left[ (y'_M \leq Y_M \leq y''_M) \wedge \left( \bigwedge_{m=1}^{M-1} (Y_m \geq y'_m) \right) \middle| z \right] \quad (8)$$

This is a probability that can be estimated using data on values of ordered bids while the probability  $G_U(\mathcal{U}(y', y''_M, z; h))$  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 (5) 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.

---

<sup>9</sup>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  $\mathcal{U}(y', y''_M, z; h)$ .

## 5 Inequalities defining the identified set

### 5.1 Valuations stochastically dominate bids

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

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

the containment functional probability (8) 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, (7), is

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

and so the condition (5) delivers the following inequalities.<sup>10</sup>

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

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 (1).

The marginal distribution of the  $n$ th order statistic of  $M$  identically and independently distributed uniform variates is  $Beta(n, M + 1 - n)$ .<sup>11</sup> Let  $Q(p; n, M)$  denote the associated quantile function. The restrictions placed on valuation distributions by the inequality (5) and the test sets under consideration in this Section are, written in terms of uniform order statistics:

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

which can be written as follows.

$$\forall v : \quad F_z(v) \leq \min_n Q(\mathbb{P}[Y_n \leq v|z]; n, M) \quad (9)$$

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.

<sup>10</sup>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  $F_z$ .

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

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  $y' = (-\infty, -\infty, \dots, -\infty)$  the containment functional probability (8) is:

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

the probability mass placed by the distribution  $G_U$  on the contiguous union, (7), is

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

so, with this choice of  $(y', y''_M)$ , the condition (5) delivers the following inequalities.

$$\forall v : \quad \mathbb{P}[V_{M-1} \leq v|z; F_z] \geq \mathbb{P}[Y_M \leq v|z] \quad (10)$$

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).

All valuation distribution functions in the identified set must satisfy:

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

which is (10) rewritten in terms of a uniform order statistic, equivalently<sup>12</sup>

$$\forall v : \quad F_z(v) \geq Q(\mathbb{P}[Y_M \leq v|z]; M-1, M). \quad (11)$$

---

<sup>12</sup>There is another expression:

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

and the following inequalities.

$$\forall s : \quad MF_z(s)^{M-1} - (M-1)F_z(s)^M \geq \mathbb{P}[Y_M \leq s|z]$$

This is the bound given in Theorem 2 of HT when the minimum bid increment considered there is set equal to zero.<sup>13</sup>

Figure 8 shows, outlined in blue, the contiguous union 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 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

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

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

---

<sup>13</sup>With a positive minimum bid increment  $\Delta$ ,  $Y_M$  is replaced by  $Y_M + \Delta$ .

With  $n = M$  there is<sup>14</sup>

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

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

and the condition (5) delivers the following inequalities.

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

With  $n < M$ ,  $G_U(\mathcal{U}(y', y''_M, z; h))$  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

$$\mathbb{P}[\mathcal{U}(Y, Z; h) \subseteq \mathcal{U}(y', y''_M, z; h)|z] = \mathbb{P}\left[\bigwedge_{m=n}^M Y_M = v \mid z\right] \quad (12)$$

may be positive in which case the inequality (5) is violated whatever candidate distribution  $F_z$  is considered.

This violation for any  $F_z$  arises because a nonzero probability (12) 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 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 (9) and (11) 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

<sup>14</sup>The expression for  $G_U(\mathcal{U}(y', y''_M, z; h))$  is obtained as:

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

which delivers the result as stated.

functional inequality (5) 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  $v_1 \geq v_2$ ,

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

The containment functional probability (8) is

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

while (7) delivers

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

which, plugged into (5) 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

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

are as follows.

$$\begin{aligned} \forall v_1 \geq v_2 : \quad & 1 - F_z(v_1)^M - M F_z(v_2)^{M-1} + M F_z(v_1) F_z(v_2)^{M-1} \\ & \geq \mathbb{P}[Y_M \geq v_1 \wedge Y_{M-1} \geq v_2 | z] \quad (13) \end{aligned}$$

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  $v_1 > v_2$ ,

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

The containment functional probability (8) 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 (7) delivers

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

which, plugged into (5) 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.

$$\begin{aligned} \forall v_1 > v_2 : \quad & F_z(v_1)^M - M F_z(v_1) F_z(v_2)^{M-1} + (M-1) F_z(v_2)^M \\ & + M(1 - F_z(v_1))(F_z(v_1)^{M-1} - F_z(v_2)^{M-1}) \geq \\ & \mathbb{P}[v_1 \geq Y_M \geq v_2 \wedge Y_{M-1} \geq v_2 | z] \quad (14) \end{aligned}$$

### 5.4.3 Discussion

The inequalities presented in this Section, of which (13) and (14) are examples, may not be satisfied by all valuation distributions which satisfy the HT bounds (9) and (11). 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 Numerical example

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 (9) and (11) 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  $45^\circ$  line, drawn in orange. The new inequality (13) 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 (13) delivers no refinement in this case. The new inequality (14) 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 (9) and (11) 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  $45^\circ$  line since the survivor function is decreasing. The new bounds (13) and (14) 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 (9) and (11).

## 7 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 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<sup>15</sup> 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 *every* structure 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.<sup>16</sup> 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<sup>17</sup> 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

---

<sup>15</sup>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.

<sup>16</sup>See Section VIII of HT03.

<sup>17</sup>See Appendix D of HT03.

are frequently hard to obtain. The method set out in CR and applied here has the advantage that sharpness is guaranteed.

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.

## ANNEX

### Calculation of bid probabilities for the numerical examples 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.<sup>18</sup> In each of the  $10^7$  simulated auctions

---

<sup>18</sup>  $LN(\mu, \sigma^2)$  is a random variable whose logarithm is  $N(\mu, \sigma^2)$ .

two independent realizations of this mixture distribution are sampled and sorted to deliver realizations of ordered valuations.

Independently distributed random variables,  $\Theta_1$  and  $\Theta_2$  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  $\Theta$  times the bid on the table plus  $(1 - \Theta)$  times her valuation with  $\Theta$  equal to  $\Theta_1$ , respectively  $\Theta_2$ , 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 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.

#### REFERENCES

- CHESHER, A.D., AND A.M. ROSEN (2014) "Generalized Instrumental Variable Models," CeMMAP Working Paper, CWP0414.
- DAVID, H.A., AND H.N. NAGARAJA (2003) *Order Statistics*, 3rd edition. Hoboken: John Wiley.
- GENTRY, M., AND TONG 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.
- MILGROM, P.R., AND R.J. WEBER (1982) "A Theory of Auctions and Competitive Bidding." *Econometrica* 50(3), 1089-1122.
- MOLCHANOV, I. (2005) *Theory of Random Sets*, London: Springer-Verlag.

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.

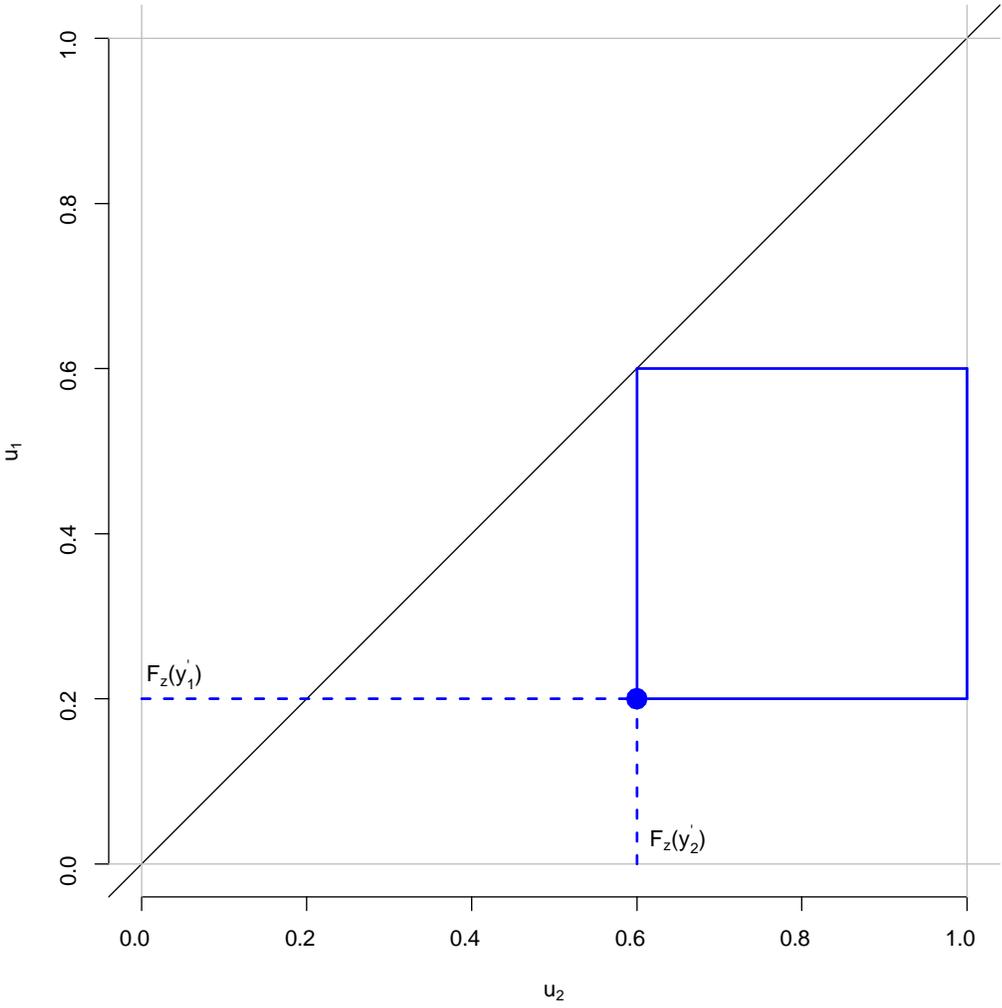


Figure 2: In blue the  $U$ -level set  $\mathcal{U}((y'_1, y'_2), z; h)$  containing values of uniform order statistics,  $u_2 \geq u_1$ , that can give rise to order statistics of bids,  $(y'_1, y'_2)$ .  $F_z$  is the distribution function of valuations. In magenta the  $U$ -level set obtained as  $y'_2$  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.

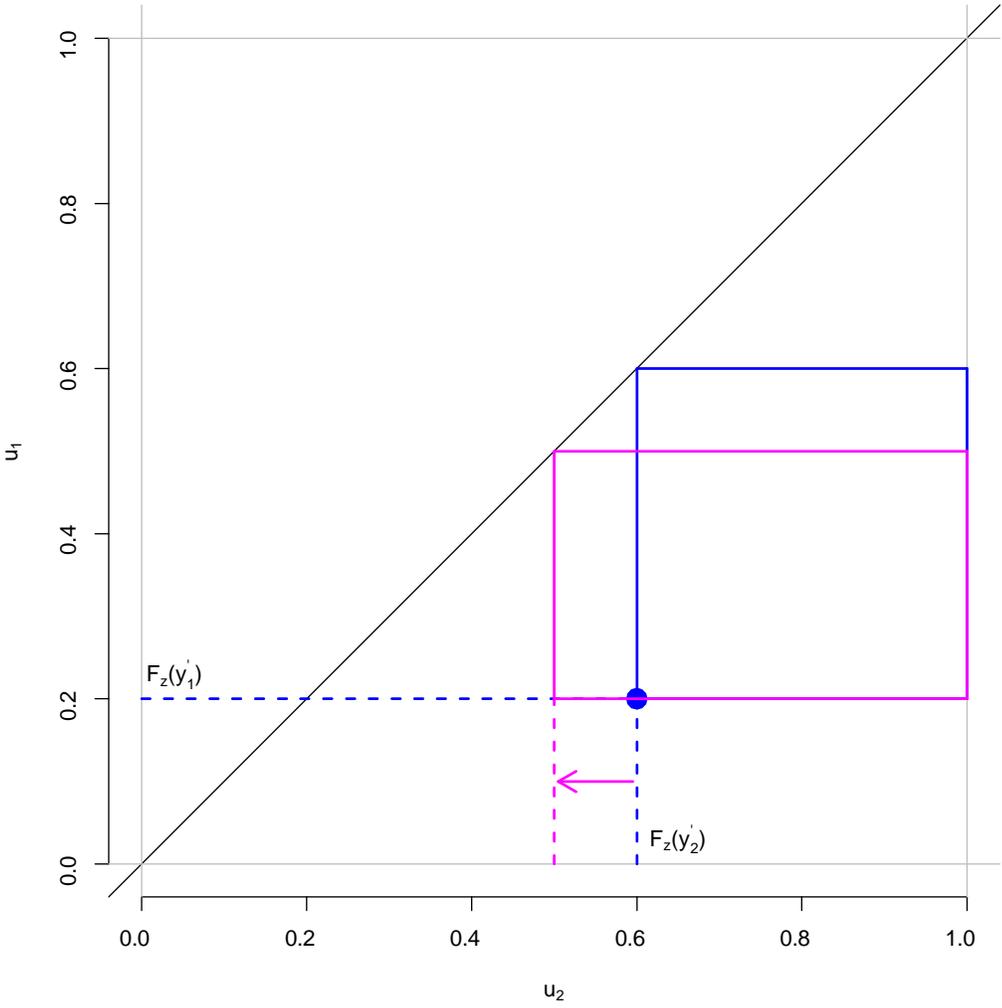


Figure 3: In blue the  $U$ -level set  $\mathcal{U}((y'_1, y'_2), z; h)$  containing values of uniform order statistics,  $u_2 \geq u_1$ , that can give rise to order statistics of bids,  $(y'_1, y'_2)$ .  $F_z$  is the distribution function of valuations. In magenta the  $U$ -level set obtained as  $y'_2$  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.

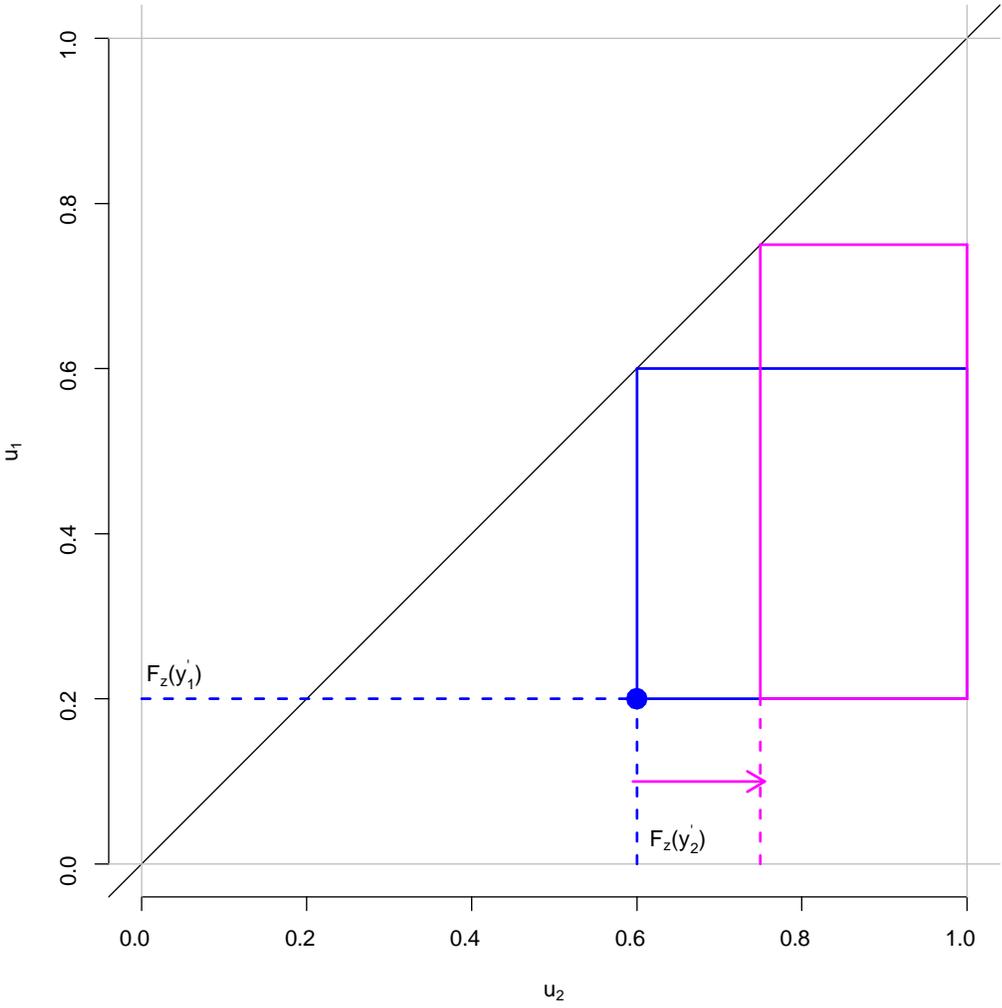


Figure 4: In blue the  $U$ -level set  $\mathcal{U}((y'_1, y'_2), z; h)$  containing values of uniform order statistics,  $u_2 \geq u_1$ , that can give rise to order statistics of bids,  $(y'_1, y'_2)$ .  $F_z$  is the distribution function of valuations. In magenta the  $U$ -level set obtained as  $y'_1$  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.

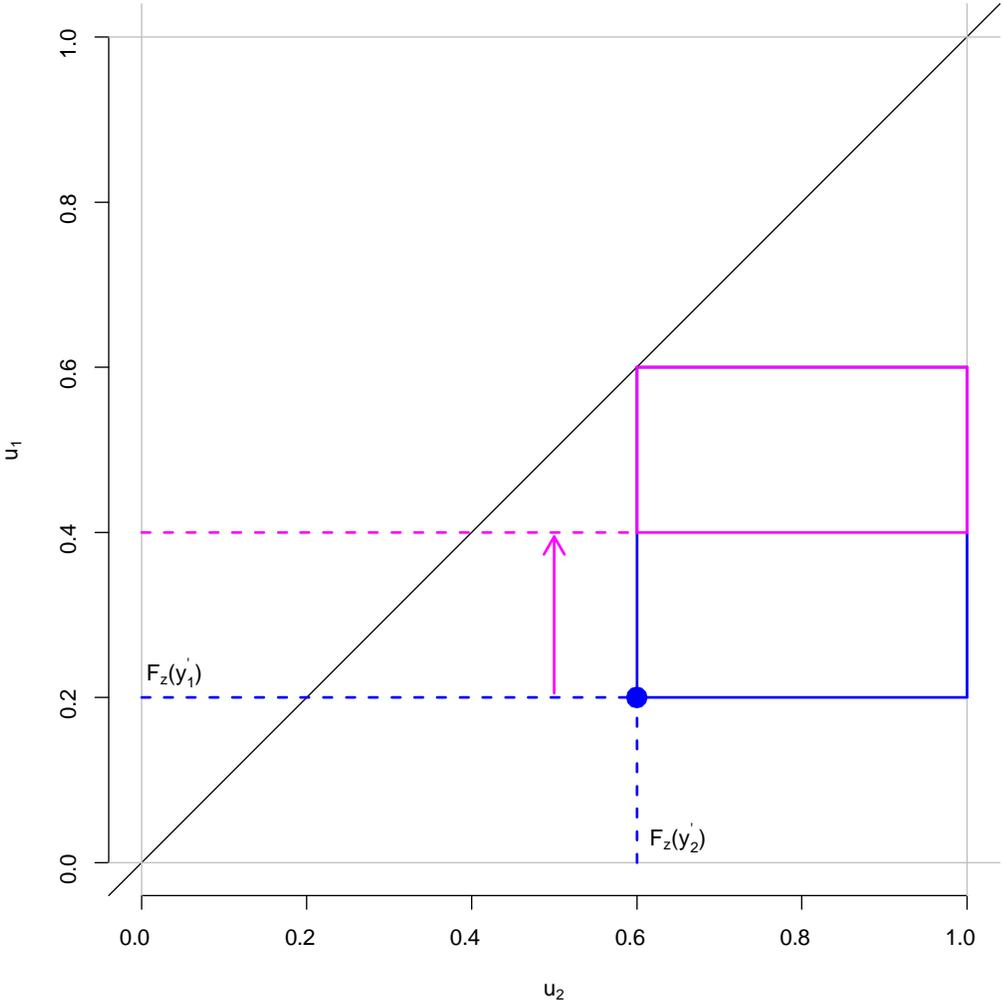


Figure 5: 2 bidder case. The contiguous union of level sets:  $\mathcal{U}(y', y_2'', z; h)$  where  $y' = (y'_1, y'_2)$ ,  $y'_2$  and  $y_2''$  are values taken by the maximal order statistic of bids,  $Y_2$ , and  $y'_1$  is a value taken by the second largest order statistic of bids,  $Y_1$ . In 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.

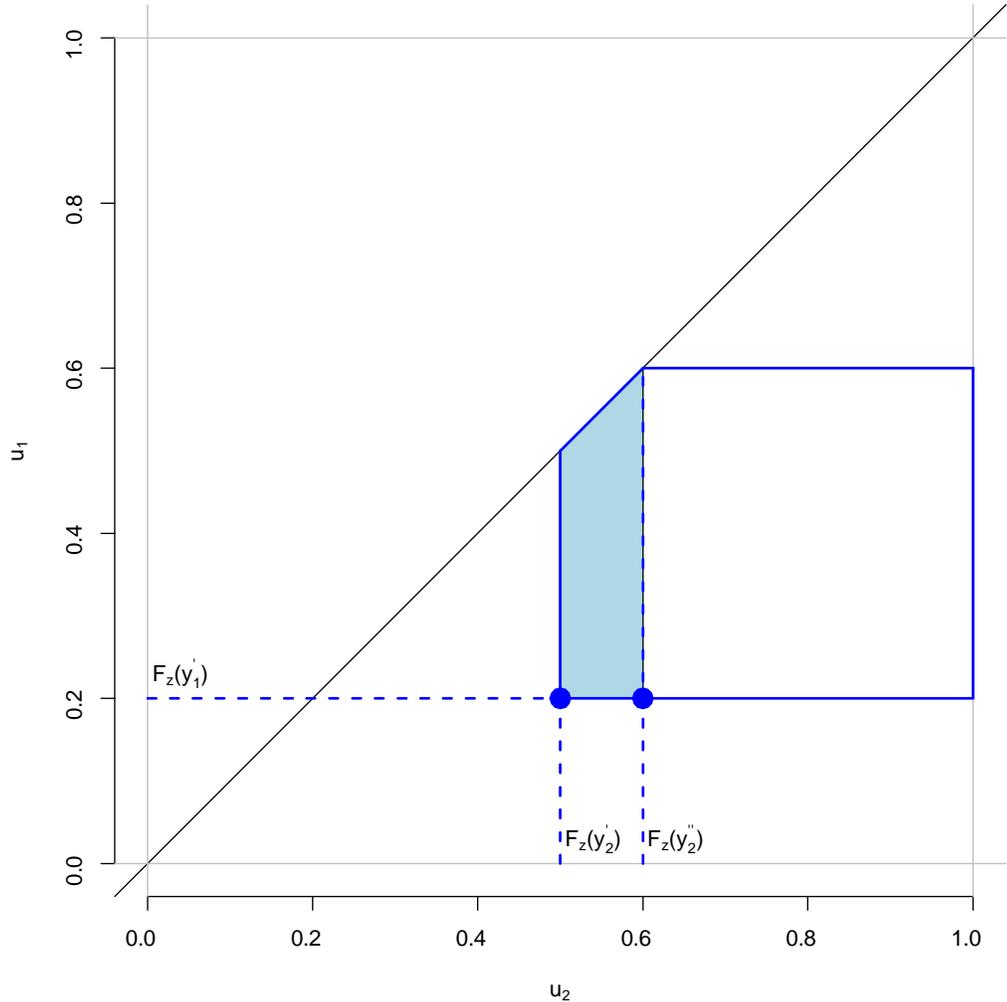


Figure 6: 2 bidder case. The triangular region outlined in blue is the contiguous union of level sets:  $\mathcal{U}(y', y_2'', z; h)$  where  $y_1' = y_2' = v$  and  $y_2'' = \infty$ .  $F_z$  is the distribution function of valuations. This choice of  $y'$  and  $y_2''$  delivers the inequality requiring 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.

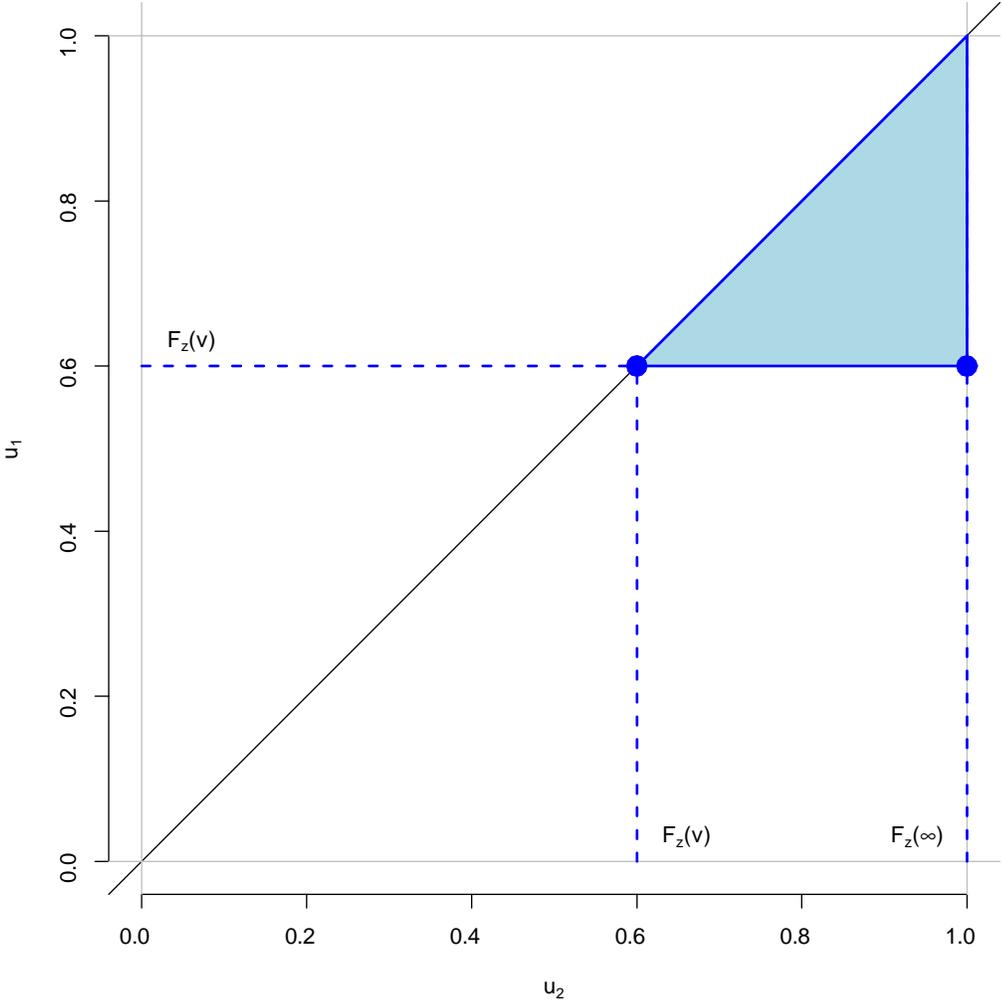


Figure 7: 2 bidder case. The trapezoidal region outlined in blue is the contiguous union of level sets:  $\mathcal{U}(y', y_2'', z; h)$  where  $y' = (-\infty, v)$  and  $y_2'' = \infty$ .  $F_z$  is the distribution function of valuations. This choice of  $y'$  and  $y_2''$  delivers the inequality 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.

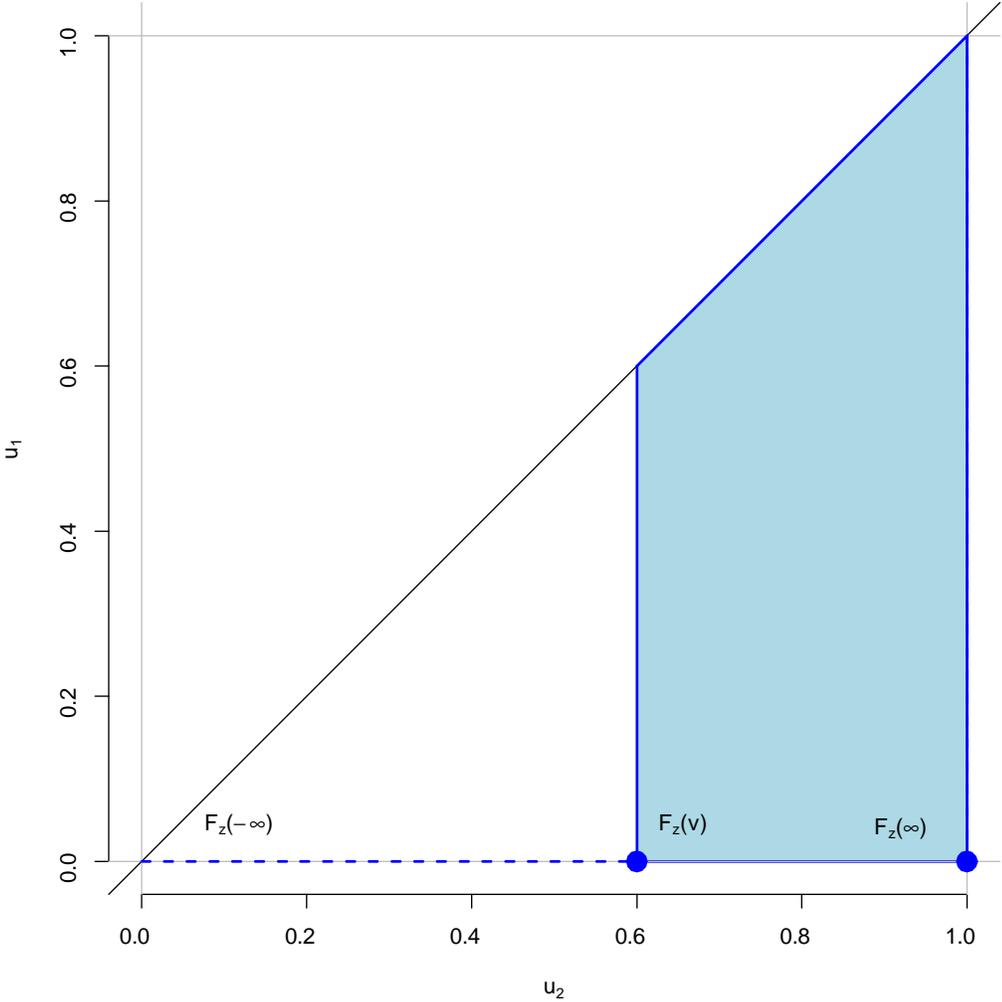


Figure 8: 2 bidder case. The trapezoidal region outlined in blue is the contiguous union of level sets:  $\mathcal{U}(y', y_2'', z; h)$  where  $y' = (-\infty, -\infty)$  and  $y_2'' = v$ .  $F_z$  is the distribution function of valuations. This choice of  $y'$  and  $y_2''$  delivers the inequality 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.

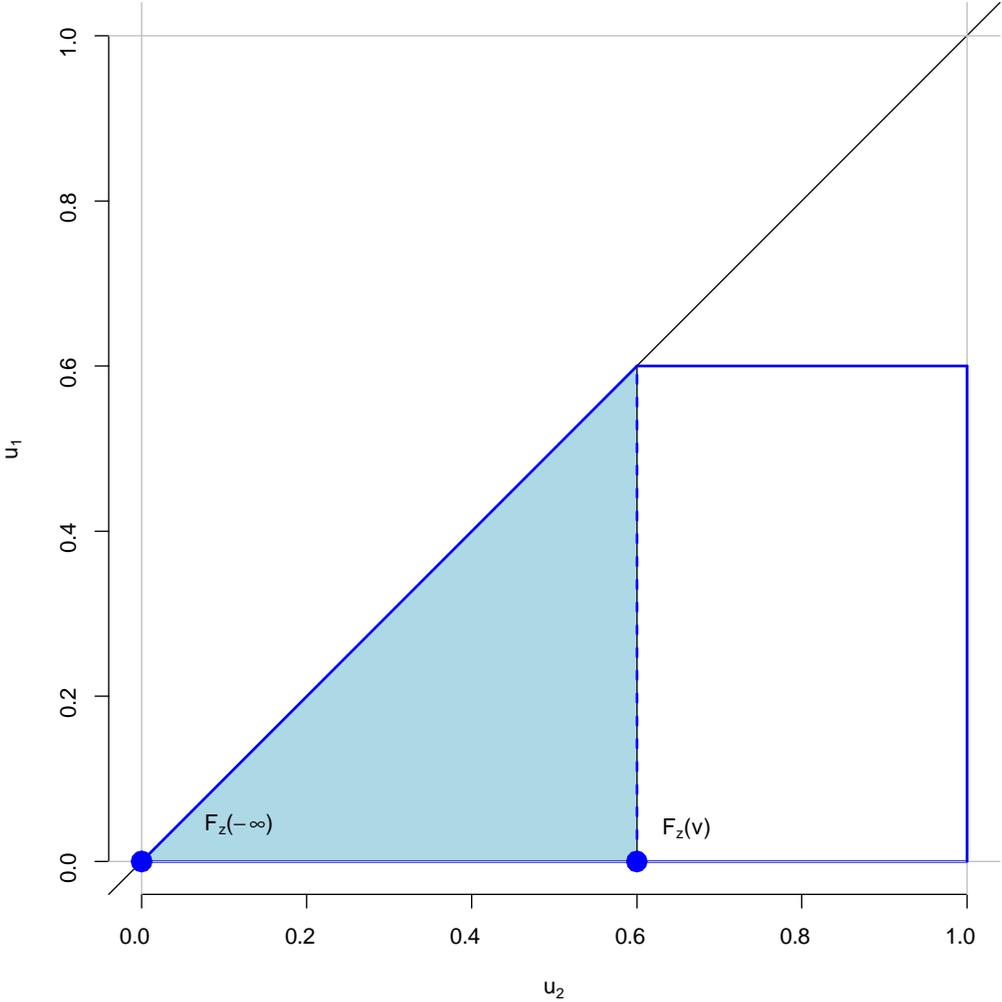


Figure 9: Example 1. Upper and lower bounds (blue) on the valuation survivor function (red).

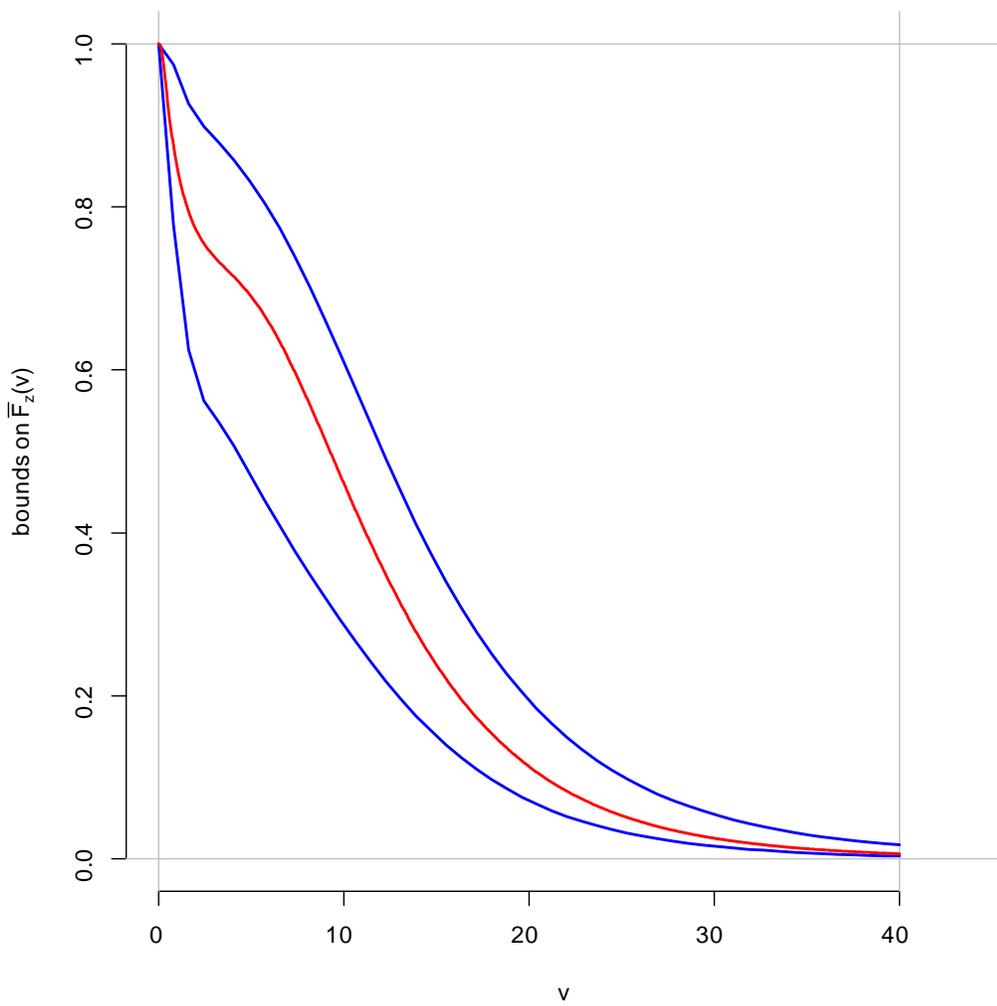


Figure 10: Example 1. Upper and lower bounds (blue) on the valuation survivor function (red). Two values of  $v$ ,  $v_1 = 14$  and  $v_2 = 5$  are identified.

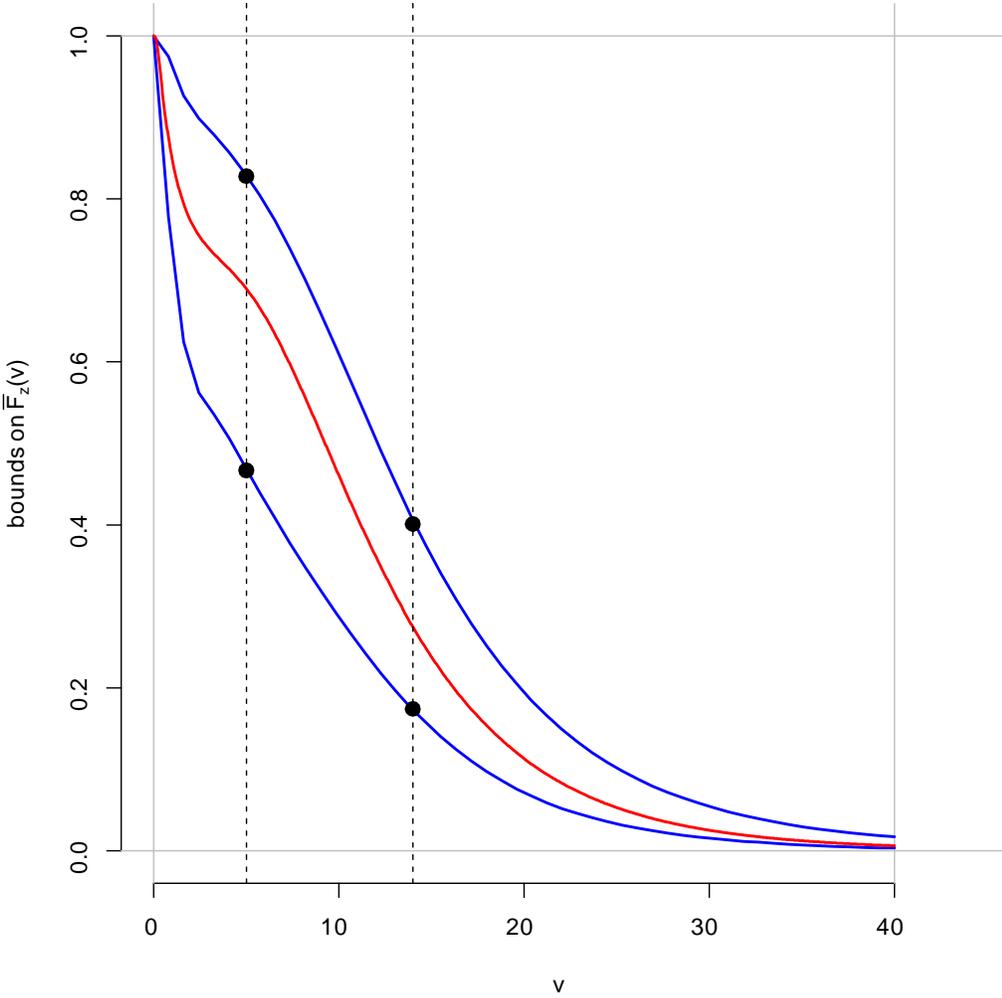


Figure 11: Example 1. The blue rectangle shows upper and lower bounds on  $\bar{F}_z(v_1)$  and  $\bar{F}_z(v_2)$  at  $v_1 = 14$  and  $v_2 = 5$ . These ordinates of the valuation survivor function must lie above the 45° line (orange). The new bounds require they lie above the magenta and red lines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

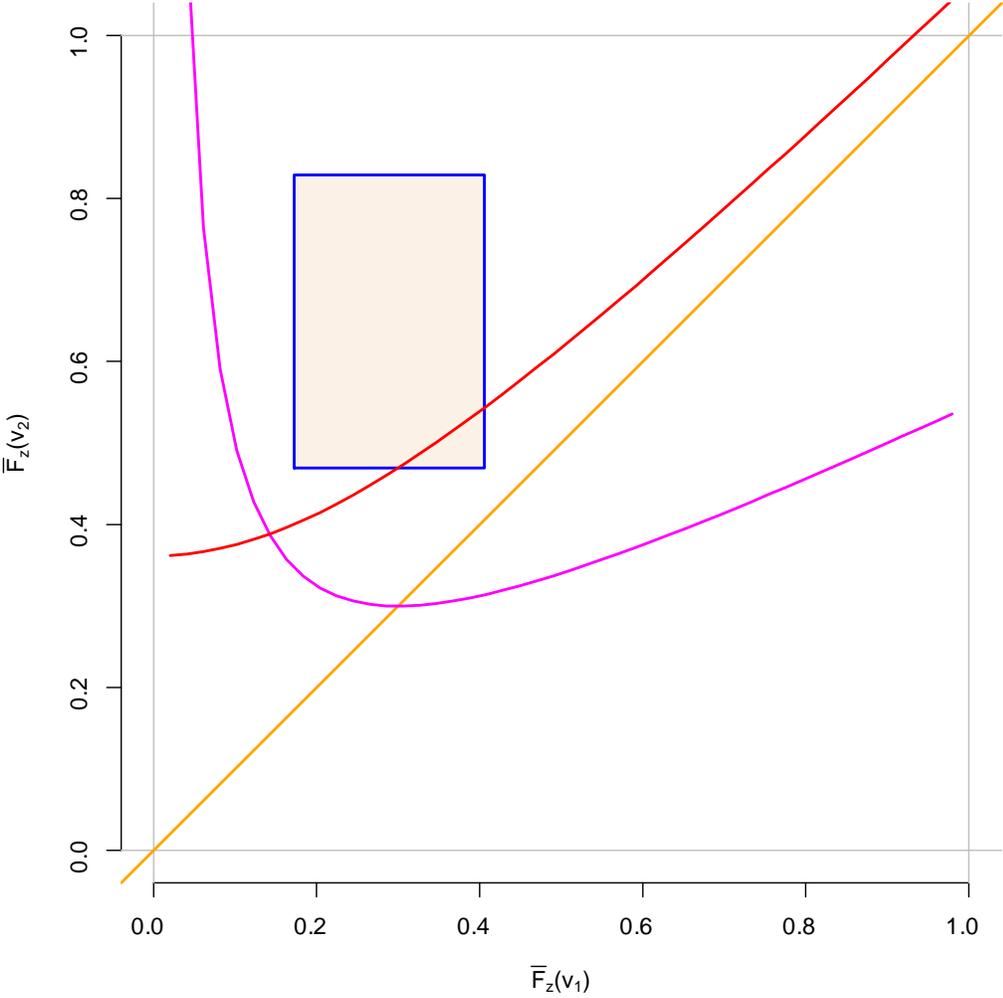


Figure 12: Example 2. Upper and lower bounds (blue) on the valuation survivor function (red).

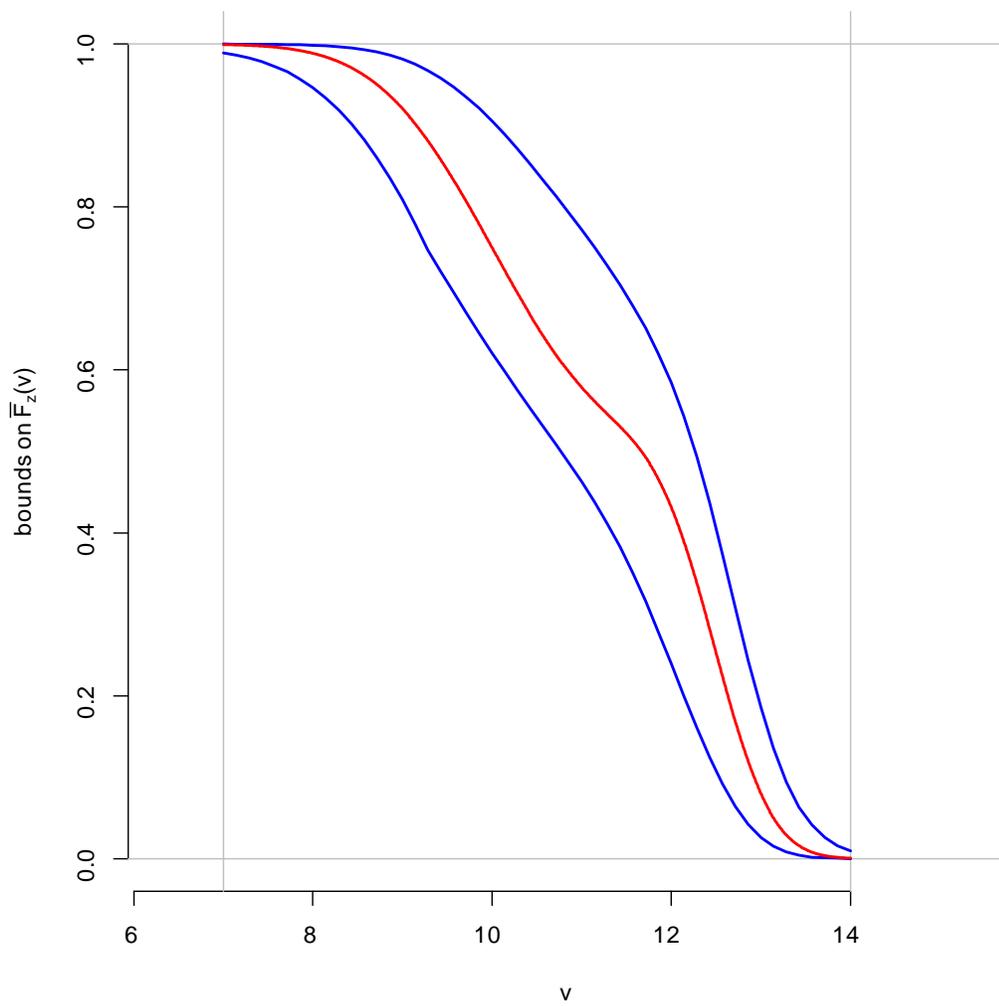


Figure 13: Example 2. Upper and lower bounds (blue) on the valuation survivor function (red). Two values of  $v$  are identified:  $v_1 = 12.5$  and  $v_2 = 11.5$ .

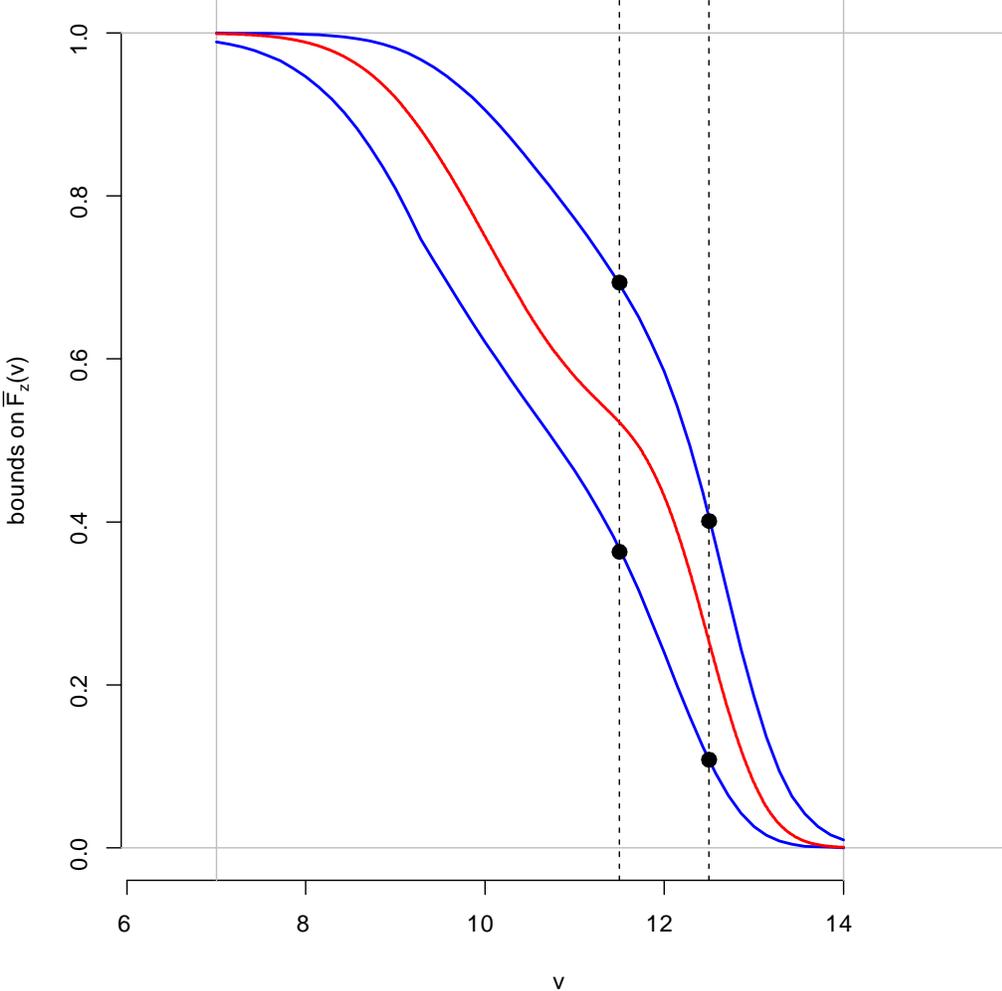


Figure 14: Example 2. The blue rectangle shows upper and lower bounds on  $\bar{F}_z(v_1)$  and  $\bar{F}_z(v_2)$  at  $v_1 = 12.5$  and  $v_2 = 11.5$ . These ordinates of the valuation survivor function must lie above the 45° line (orange). The new bounds (13) and (14) require they lie above the magenta and red lines as well.

