Nonparametric Identification of First-Price Auction with Unobserved Competition: A Density Discontinuity Framework

Emmanuel Guerre
School of Economics
University of Kent
United Kingdom

Yao Luo
Department of Economics
University of Toronto
Canada

August 2019

Acknowledgments: The authors would like to thank Lu Han, Xavier d'Haultfœuille, Quang Vuong, together with conference and seminar participants for stimulating discussions and useful suggestions. Cheok in Fok and Jiaqi Zou provided excellent research assistance. The second author thanks SSHRC for financial support. All remaining errors are our own.

Abstract

We consider nonparametric identification of independent private value first-price auction models, in which the analyst only observes winning bids. Our benchmark model assumes an exogenous number of bidders N. We show that, if the bidders observe N, the resulting discontinuities in the winning bid density can be used to identify the distribution of N. The private value distribution can be identified in a second step. A second class of models considers endogenously-determined N, due to a reserve price or an entry cost. If bidders observe N, these models are also identifiable using winning bid discontinuities. If bidders cannot observe N, however, identification is not possible unless the analyst observes an instrument which affects the reserve price or entry cost. Lastly, we derive some testable restrictions for whether bidders observe the number of competitors and whether endogenous participation is due to a reserve price or entry cost. An application to USFS timber auction data illustrates the usefulness of our theoretical results for competition analysis, showing that nearly one bid out of three can be non competitive. It also suggests that the risk aversion bias caused by a mismeasured competition can be large.

Keywords: Auction models, unobserved competition, nonparametric identification, density discontinuities, endogeneous participation, unobserved heterogeneity, discrete mixture models

JEL classification: C14, C57, D44

1 Introduction

There exists a large literature on nonparametric identification of auction models; see, e.g., Athey and Haile (2007) or Hendricks and Porter (2007) for a review. In the case of sealed-bid first price auction, a vast majority of work assumes that the analyst can observe all of the bids, or both the winning bid and the number of competitors. This may not always be observed. Lamy (2012) mentions that, in French timber auctions, only the winning bid may be available to researchers to preserve anonymity of the bidders. It is common practice in many markets that can be treated as auctions for only the winning bid (i.e. the transaction price) to be recorded. For instance, a company soliciting price quotes for a task to be completed is implicitly organizing a first-price auction. While the company may not record all quotes or the number of responses, the price charged to the winning bidder is likely to appear in accounting records. As noted in Han and Strange (2015), "bidding wars" are becoming commonplace in housing markets, where houses are sold through a competitive bidding process resembling an informal first-price auction. Governments may offer subsidies to attract firms as recently considered by Kim (2018) and Slattery (2019) using an auction framework. Observing all the subsidy offers is again unlikely. Hence, in many economic situations of interest, the records only contain the final winning bid. Therefore, the ability to identify auction primitives solely from winning bid data may enlarge the scope of auction theory applications.

A second motivation stems from misspecification considerations, because the observed number of bids can be a misleading indicator of competition. For instance, Laffont, Ossard and Vuong (1995) consider an application where bidders are agents acting for several buyers such that the number of bids underestimates competition. Alternatively, some buyers may enter an auction simply to gain information, in which case their bids will be dominated and never impact the winning bid; counting these bids would overestimate competition. When the observed number of bids as a measure of competition is at doubt, using only winning bids provides a robust approach for identifying primitives of interests. Our USFS timber application illustrates the bias in estimating risk-aversion when we use the number of bids as the measure of competition as in Lu and Perrigne (2008) or Campo, Guerre, Perrigne and Vuong (2011). Bidders are much more likely to be risk neutral when we are agnostic about the distribution of competition.

Last, the distribution of competition is a parameter of interest in itself. We view it as a latent variable and recover it from the winning bid. As noted above, the number of observed bids is often different from the number of bids actually contributing to the winning bid. In this case, the more meaningful measure is the latter, which is useful for auction design such as choosing an optimal reserve price. In our USFS timber application, it is found that nearly one bidder out of three does not bid in a competitive way. As noted by Bulow and Klemperer (1996), increasing competitive participation would give a higher seller expected revenue than choosing an optimal reserve price. Our finding therefore suggests that, among other things, the seller should consider a bidder training program, as studied in De Silva, Li and Zhao (2019), or should investigate whether the dominated bid is due to collusion, see eg Abrantes-Metz and Bajari (2009).

This paper therefore studies the identification of model primitives using only data on winning bids. Specifically, we develop a new approach for first-price auction model identification that exploits discontinuities in the density function of the winning bid. First, we identify the distribution of unknown competition. In particular, we build on an important restriction that first-price auction models impose on the data: the bid quantile function must be strictly increasing with respect to the number of bidders. Therefore, the upper boundary of the bid distribution, conditioning on the number of bidders, is strictly increasing, as well. We show that this creates discontinuities, or jumps, in the winning bid density function at these upper boundaries. A novel result of the paper is that these jumps identify the distribution of the number of bidders.

Second, we identify the value distribution function by iteratively exploiting two equilibrium mappings relating the value and bid quantile functions. Based on the location of winning bid density discontinuities, we create a sequence of expanding quantile intervals over which the private value quantile is identified. For every iteration, we start by identifying the bid quantile function in the most competitive auction, which has the largest number of bidders. This information can then be used to identify the value quantile function in the same quantile interval and further calculate the corresponding bid quantile for other competition levels.

We then extend our results to endogenous participation due to the presence of a reserve price or entry cost. When active buyers observe the number of participants, identification extends naturally. On the other hand, identification fails due to a lack of discontinuity if active buyers cannot observe the number of participants. We restore identification assuming that the analyst can instrument for the reserve price or entry cost, as in Gentry and Li (2014). In addition, it is assumed that the analyst observes auctions where the object is not sold or there is only one participant. Identification is then established exploiting the simple binomial distribution of the number of bidders.

Lastly, we derive testable restrictions of whether or not bidders observe the number of competitors, and if participation is constrained by a reserve price or entry cost. In models of endogenous participation, it is usually assumed that the buyers do not observe the number of active buyers. See, among others, Guerre, Perrigne and Vuong (2000), Marmer, Shneyerov and Xu (2013) or Gentry and Li (2014). As known since McAfee and McMillan (1987), the expected payoff of a first-price auction with reserve price or entry cost does not depend upon buyers information regarding competition. Hence, under risk neutrality of buyers and sellers, both kinds of buyer competition information are likely to hold. Whether buyers observe competition can be a known market characteristic, as in the case where active buyers sit in the same room or in housing market "bidding wars" where real estate agents often decide to reveal the number of offers. But it can also be a parameter of interest, as in the case of a company contacting contractors in a public or private way unobserved by the analyst.

Related literature

Auctions with unobserved competition. Allowing for unknown competition started early in the empirical auction literature. Laffont et al. (1995) estimate the number of buyers N as a parameter that they take to be constant across auctions. Paarsch (1997) treats unknown competition as a nuisance parameter, which is eliminated using conditional likelihood estimation. For ascending eBay auctions, Song (2004) shows that the private value distribution and a constant number of buyers are identified from winning and second highest bids, but not from winning bids alone when N is random. More pertinent to our paper is the misclassification approach of An, Hu and Shum (2010), who achieve identification from the winning bid using a proxy $N^* \leq N$ for the number of buyers and an instrument that can be a discretized second bid. Shneyerov and Wong (2011) suppose that only winning bids and the number of active bidders are observed. Recent work for ascending auctions include Quint (2015) and Freyberger and Larsen (2017).

Mixture distribution. The present paper contributes to the literature on nonparametric identification of finite mixtures; see for instance the review of Compiani and Kitamura (2016). Existing identification results require either exclusion restrictions or multiple independent measurements. A first-price auction example of the latter is d'Haultfœuille and Février (2015), who recover the distribution of an unobserved continuous auction characteristic from three bids. Hu and Sasaki (2017) obtain identification from two measurements in a model with discrete unobserved heterogeneity. In our setting, the number of buyers can be viewed as unobserved heterogeneity while the winning bid is a unique bid. Identification is however possible because the mixture components are the bid distribution given $N = n$, issued from the same private value distribution and constrained by an optimal bidding condition. When the buyers do not observe the number of competitors, this is restrictive enough to ensure identification in presence of a reserve price or entry cost instrument, without the exclusion restrictions of Compiani and Kitamura (2016).

Discontinuity design. The discontinuity design (DD) literature has expanded rapidly in recent years; interested readers are encouraged to refer to review papers by Imbens and Lemieux (2008), Kleven (2016) and Jales and Yu (2017). Recent auction applications include Coviello and Marinello (2014), and Choi, Neisheim and Razul (2016). As in the DD literature, this paper employs jump sizes for identification purposes - more specifically, to identify the probability that $N = n$. However, this paper departs from the DD literature by considering an unknown number of density discontinuities with unknown location, which identify the support of N.

The remainder of the paper is organized as follows. In Section 2, we describe the benchmark model. In Section 3, we extend our analysis to auction models with endogenous competition generated by reserve price or entry cost, for when buyers can or cannot observe competition. Section 4 reports the results of our empirical application. Section 5 concludes. We also include a proof section that contains all proofs omitted in the main text. A discontinuity detection algorithm, which also compute a discontinuous density estimator, is presented in the appendix.

2 The benchmark model

In this section, we start by describing the benchmark auction model and introduce two equilibrium mappings that are convenient for describing our discontinuity identification strategy. Next, we derive the restrictions that the model imposes on the observed winning bids, especially with respect to the formation of discontinuities. Finally, we describe our identification strategy in two steps. First, we identify the distribution of the number of buyers from the discontinuities in the winning bid density function. Second, we identify the value distribution function using the two equilibrium mappings iteratively.

2.1 The symmetric independent private values paradigm

Suppose there is a single item for sale with N active symmetric buyers bidding for the item. All buyers observe N. In contrast, the analyst does not observe N, which causes auction-specific unobserved heterogeneity. Each buyer i also observes her private valuation V_i, which is unknown to other buyers. The private values V_i are *i.i.d.* draws from a distribution $F(\cdot)$, which is known to all the buyers and is independent of N. The buyers are risk neutral and their bids B_i are formed according to a symmetric best-response strategy. In sum, the primitives are the distribution of the number of buyers N and the private value distribution.

We assume that the analyst only observes the winning bid W, i.e., the maximum bid among the N buyers in the set $\mathcal{N}$ of active buyers

$$W = \max_{i \in \mathcal{N}} B_i.$$

Hence, the analyst observes draws from the unconditional cumulative probability distribution of the winning bid $G(\cdot)$, which is a mixture of the conditional winning bid distributions given N:

$$G(b) \;=\; \sum_{n=2}^{+\infty} \mathbb{P}\left(\max_{1\leq i\leq n} B_i \leq b\right) \times \mathbb{P}\left(N=n\right) = \sum_{n=2}^{+\infty} G_n^n\left(b\right)\mathbb{P}\left(N=n\right), \tag{1}$$

where $G_n(\cdot)$ is the conditional bid distribution given $N = n$.

The two next assumptions introduce some additional conditions for the distribution of N and for the private value distribution $F(\cdot)$.

Assumption N. *The number of active buyers N is a discrete random variable with support $\{\underline{n}, \ldots, \overline{n}\}$ for some integers $2 \leq \underline{n} \leq \overline{n}$, i.e., $p_n = \mathbb{P}(N = n) > 0$ for $n = \underline{n}, \underline{n}+1, \ldots, \overline{n}$ with $\sum_{n=\underline{n}}^{\overline{n}} p_n = 1$.*

Assumption IPV. *Buyers' private values V_i are unknown to competitors and i.i.d. draws from a common knowledge distribution $F(\cdot)$. The cumulative distribution function $F(\cdot)$ has a compact support $[\underline{v}, \overline{v}]$. Its probability density function $f(\cdot)$ is continuous and strictly positive over $[\underline{v}, \overline{v}]$.*

Both theoretical and empirical literatures adopt the assumption of a private value distribution with compact support. In particular, it rules out multiple asymmetric equilibria; see Maskin and Riley (1984, Remark 2.3), who also establish that symmetric Bayesian Nash Equilibrium bids are given by a strictly increasing and continuously differentiable function of private values.

For our discontinuity approach, the compact support assumption ensures the existence of discontinuities in the density of unconditional winning bids that we exploit in this paper. In particular, the winning bid densities $g_n(\cdot)$ given $N = n$ stay bounded away from 0 at its upper boundary; see (7) below.

2.2 Bid and value quantile equilibrium mappings

In this subsection, we describe two equilibrium mappings that are repeatedly used in our identification procedure. Specifically, there is an equilibrium mapping from the value distribution and the bid distribution, and vice versa. Our discontinuity identification strategy is conveniently described using the quantile framework as in Guerre, Perrigne and Vuong (2009), Liu and Luo (2016), and Guerre and Gimenes (2019), that we recount below.

Let $V(\alpha) = F^{-1}(\alpha)$ represent the private value quantile function, where $\alpha \in [0,1]$ is the quantile level. Let $B_n(\alpha)$ denote the bid quantile function given that n buyers participate in the auction. Following Milgrom (2001)'s exposition of the identification strategy of Guerre, Perrigne and Vuong (2000), the private value quantile function $V(\cdot)$ can be viewed as the common valuation function of buyers who receive independent uniform private signals

$$A_i = F(V_i),$$

which determine their private values $V_i = V(A_i)$. By Assumption IPV, $B_i = \beta_n(A_i)$ for all i, where $\beta_n(\cdot)$ is strictly increasing and continuously differentiable. It follows that for any b in the range of $\beta_n(\cdot)$,

$$G_n(b) = \mathbb{P}(\beta_n(A_i) \leq b) = \mathbb{P}\left(A_i \leq \beta_n^{-1}(b)\right) = \beta_n^{-1}(b)$$

because A_i is uniformly distributed over $[0,1]$. Hence the best-response strategy is the bid quantile function

$$\beta_n(\alpha) = B_n(\alpha) \text{ for all } \alpha \in [0,1].$$

Now, let us relate the bid and private value quantile functions. Suppose that buyer i receives signal α but makes a suboptimal bid $B_n(a)$ for some $a \in [0,1]$. Since her opponents bid $B_n(A_j)$, the probability that her bid $B_n(a)$ wins the auction is given by $\mathbb{P}(\max_{j\neq i} A_j \leq a)$, which is equal to a^{n-1} as the signals of the $n-1$ opponents A_j, where $j \neq i$, are independent and uniform. It follows that the expected payoff of buyer i is $\left(V\left(\alpha\right) - B_n\left(a\right)\right) a^{n-1}$, which is maximized when $a = \alpha$. Since

$$\begin{aligned}\frac{\partial}{\partial a}\left[\left(V\left(\alpha\right) - B_n\left(a\right)\right) a^{n-1}\right]\Big|_{a=\alpha} &= V\left(\alpha\right)\left(n-1\right)\alpha^{n-2} - \frac{\partial\left[B_n\left(\alpha\right)\alpha^{n-1}\right]}{\partial\alpha} \\ &= \left(n-1\right)\alpha^{n-2}\left(V\left(\alpha\right) - B_n\left(\alpha\right) - \frac{\alpha B_n^{(1)}\left(\alpha\right)}{n-1}\right),\end{aligned}$$

setting this derivative to 0 gives

$$V\left(\alpha\right) = B_n\left(\alpha\right) + \frac{\alpha B_n^{(1)}\left(\alpha\right)}{n-1}. \tag{2}$$

This constitutes the equilibrium mapping from the bid quantile function to the value quantile function, which is the basis of the identification of $V(\cdot)$ with knowledge of $B_n(\cdot)$.

Now, let us consider the inverse of the mapping (2). Indeed, (2) is equivalent to $\frac{\partial\left[B_n(\alpha)\alpha^{n-1}\right]}{\partial\alpha} = V\left(\alpha\right)\left(n-1\right)\alpha^{n-2}$ and, it follows,

$$B_n\left(\alpha\right) = \frac{n-1}{\alpha^{n-1}}\int_0^{\alpha} t^{n-2}V(t)dt. \tag{3}$$

For convenient of identification that will be clarified later on, let us introduce the conditional bid upper bound

$$\bar{b}_n = B_n(1) = (n-1)\int_0^1 t^{n-2}V(t)dt,$$

which gives

$$B_n(\alpha) = \frac{n-1}{\alpha^{n-1}}\left[\frac{\bar{b}_n}{n-1} - \int_{\alpha}^1 t^{n-2}V\left(t\right)dt\right]. \tag{4}$$

This constitutes the equilibrium mapping from the value quantile to the bid quantile function. The two mappings represented in (2) and (4) are repeatedly used in our identification procedure.

2.3 Structure of the winning bid distribution

The structure of winning bid distributions compatible with a first-price auction where buyers observe N follows from the mixture expression of $G(\cdot)$ in Equation (1) and the best-response differential equation (2).

Proposition 2.1 *A c.d.f. $G(\cdot)$ is rationalized by a first-price auction model satisfying Assumptions N, IPV, and observability of N by buyers if and only if*

(i). The c.d.f. $G(\cdot)$ has a mixture structure

$$G(\cdot) = \sum_{n=\underline{n}}^{\overline{n}} p_n G_n^n(\cdot), \tag{5}$$

where the $G_n(\cdot)$ are c.d.f., $2 \leq \underline{n} \leq \overline{n}$ and the positive p_n satisfy $\sum_{n=\underline{n}}^{\overline{n}} p_n = 1$.

(ii). The quantile functions $B_n(\cdot) = G_n^{-1}(\cdot)$ are continuously differentiable over $[0,1]$ and satisfy the compatibility conditions

$$B_n(\alpha) + \frac{\alpha B_n^{(1)}(\alpha)}{n-1} = B_m(\alpha) + \frac{\alpha B_m^{(1)}(\alpha)}{m-1}$$

for all $\underline{n} \leq n, m \leq \overline{n}$ and all $\alpha \in [0,1]$. Moreover the function $V(\alpha) = B_n(\alpha) + \frac{\alpha B_n^{(1)}(\alpha)}{n-1}$ is continuously differentiable over $[0,1]$ with $V^{(1)}(\cdot) > 0$.

Proof of Proposition 2.1: It remains to be shown that (i) and (ii) are sufficient. The function $V(\cdot)$ in (ii) is a quantile function associated with a c.d.f. $F(\cdot)$ satisfying the requirements of Assumption IPV while the mixture weights p_n define a distribution for N as in Assumption N. These $\{p_n, \underline{n} \leq n \leq \overline{n}\}$ and private value quantile function $V(\cdot)$ generate a distribution for N and best response bidding strategy functions $B_n(\cdot)$ by (3), with $G(\cdot)$ as a winning bid c.d.f. □

In short, a c.d.f. $G(\cdot)$ as in Proposition 2.1 is a mixture with components constrained by compatibility conditions driven by the best response differential equation

(2). The compatibility conditions of Proposition 2.1-(ii) reflects that the mixture components $G_n(\cdot)$ are generated by the same private value distribution, an important feature for identification. In particular, our identification results rely on the constraints it imposes on the extremities of the conditional bid p.d.f. $g_n(\cdot)$, as illustrated in the next corollary. Recall $\overline{b}_n = B_n(1)$, $\underline{v} = V(0) = \underline{b}_n$ and $\overline{v} = V(1)$.

Corollary 2.1 *Suppose that the compatibility conditions of Proposition 2.1-(ii) hold. Then for all $n = \underline{n}, \ldots, \overline{n}$, $\overline{b}_n < \overline{v}$, and*

$$g_n(\underline{v}) = \frac{n}{n-1} f(\underline{v}), \quad \textit{with} \tag{6}$$

$$g_n\left(\overline{b}_n\right) = \frac{1}{(n-1)(\overline{v} - \overline{b}_n)}. \tag{7}$$

Proof of Corollary 2.1: The compatibility conditions imply that (3) holds, and integrating by parts gives

$$B_n(\alpha) = \frac{1}{\alpha^{n-1}} \int_0^\alpha V(t)\, d\left[t^{n-1}\right] = V(\alpha) - \int_0^\alpha \left(\frac{t}{\alpha}\right)^{n-1} V^{(1)}(t)\, dt.$$

Hence $\overline{b}_n = \overline{v} - \int_0^\alpha t^{n-1} V^{(1)}(t)\, dt < \overline{v}$ as $V^{(1)}(\cdot) > 0$. Note that this also gives $B_n(\alpha) < V(\alpha)$ for all $\alpha > 0$, and then $B^{(1)}(\alpha) > 0$ by (2). When α goes to 0, the following holds

$$\begin{aligned} B_n(\alpha) &= V(0) + V^{(1)}(0)\alpha + o(\alpha) - \int_0^\alpha \left(\frac{t}{\alpha}\right)^{n-1} \left(V^{(1)}(0) + o(1)\right) dt \\ &= V(0) + \frac{n-1}{n} V^{(1)}(0)\alpha + o(\alpha), \end{aligned}$$

which shows that $B_n^{(1)}(0) = \frac{n-1}{n} V^{(1)}(0)$. As $B_n^{(1)}(\cdot) > 0$, the conditional bid p.d.f. $g_n(\cdot)$ satisfies

$$g_n(b) = \frac{1}{B_n^{(1)}(G_n(b))} \text{ for all } b \in \left[\underline{v}, \overline{b}_n, \right]. \tag{8}$$

Hence $g_n(\underline{v}) = 1/B_n^{(1)}(0) = \frac{n}{n-1} 1/V^{(1)}(0) = \frac{n}{n-1} f(\underline{v})$, which is (6). For (7), (2) and (8) give

$$g_n\left(\bar{b}_n\right) = \frac{G_n\left(\bar{b}_n\right)}{(n-1)\left(V\left(G_n\left(\bar{b}_n\right)\right) - \bar{b}_n\right)} = \frac{1}{(n-1)\left(\overline{v} - \bar{b}_n\right)}$$

as $G_n\left(\bar{b}_n\right) = 1$, so that (7) holds. □

Equation (7) implies that $g_n\left(\bar{b}_n\right)$ is strictly positive. It turns out from (1) that this causes discontinuities in the winning bid p.d.f. $g(\cdot)$ at each $\bar{b}_n$, as studied in the next section. As it follows that $\bar{b}_n$ is identified, (7) shows that $g_n\left(\bar{b}_n\right)$, where $n \in \{\underline{n}, \ldots, \overline{n}\}$, are determined by the common unknown parameter $\overline{v}$. We employ this consequence of the compatibility conditions of Proposition 2.1-(ii) later on to identify the distribution of N.

2.4 Winning bid density discontinuities

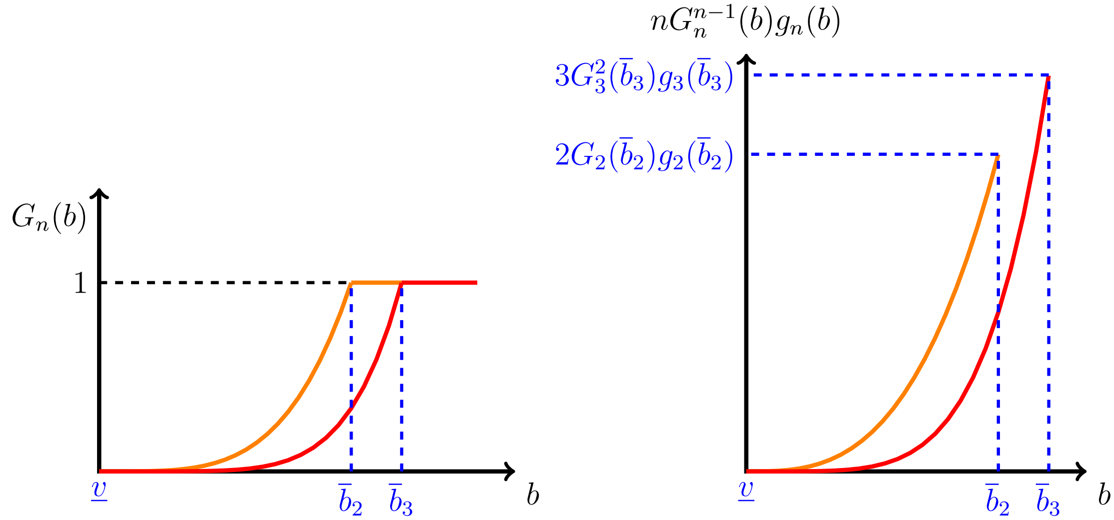
In this subsection, we introduce a numerical example to illustrate the discontinuity features of the winning bid p.d.f that follows from Corollary 2.1. This example will also be useful for introducing our identification procedure. A general lemma completes the example.

2.4.1 Numerical example

Consider the private value c.d.f. $F(v) = v^2$ for all v in $[0,1]$ and a number of buyers $N = \{2,3\}$ with equal probability. As $V(\alpha) = \alpha^{1/2}$, it follows that

$$B_n(\alpha) = \frac{n-1}{\alpha^{n-1}} \int_0^{\alpha} t^{n-2+\frac{1}{2}} dt = \frac{n-1}{n-\frac{1}{2}} \alpha^{1/2}.$$

Hence $\bar{b}_n = \frac{n-1}{n-\frac{1}{2}}$ and (8) yields the conditional bid p.d.f $g_n(b)$, given $N = n$, is equal to $2b/\bar{b}_n^2$ on $[0, \bar{b}_n]$ and vanishes outside this interval. Figure 1 displays the conditional c.d.f. and p.d.f. of the winning bid when $N = \{2,3\}$. Note that the support of



(a) c.d.f. $G_n^n(\cdot)$

(b) p.d.f. $nG_n^{n-1}(\cdot)g_n(\cdot)$

Figure 1: Conditional winning bid distribution, where $N = \{2, 3\}$ and $V(\alpha) = \sqrt{\alpha}$

the conditional density function increases with the number of buyers. Both densities jump to zero at their upper boundaries as expected from (7).

Let us now turn to the winning bid, the observation of the analyst. As expected from Figure 1b, the unconditional p.d.f. $g(b) = \frac{1}{2} \cdot 2G_2(b)g_2(b) + \frac{1}{2} \cdot 3G_3^2(b)g_3(b)$ displayed in Figure 2b is discontinuous at $\bar{b}_2$ and $\bar{b}_3$, with jump sizes Δ_2 and Δ_3 respectively. The resulting winning bid c.d.f. exhibits kinks at these values, as illustrated by Figure 2a. In this example, Figure 2b exhibits two discontinuities (and Figure 2a exhibits two kinks) because N takes two potential values here.

2.4.2 The general case

The increasing support property observed in Figure 1 and the winning bid p.d.f discontinuities in Figure 2b are generic, as shown in the upcoming lemma. Lemma 2.1-(i) recalls more generally that bids increase with competition — a key feature of

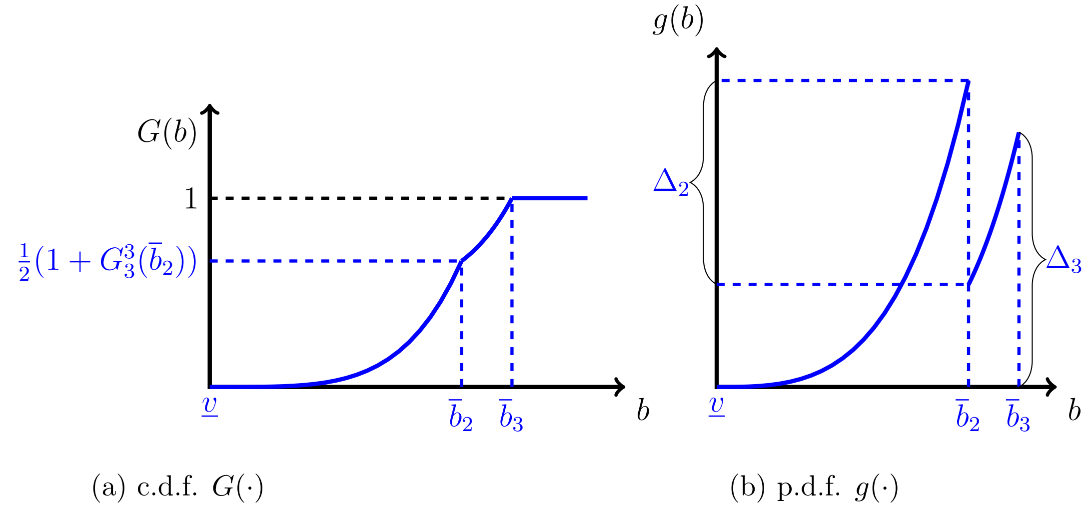


Figure 2: Winning bid distribution ($V(\alpha) = \sqrt{\alpha}$ and $\mathbb{P}(N = 2) = \mathbb{P}(N = 3) = 1/2$)

first-price auctions that does not hold in ascending or second-price ones.[1] Lemma 2.1-(ii) focuses on the winning bid p.d.f. discontinuities and its jumps.

Lemma 2.1 *Suppose Assumptions N and IPV hold. Then, all of the following hold.*

(i). For all α in $(0,1]$, $B_{\underline{n}}(\alpha) < \cdots < B_{\overline{n}}(\alpha) < V(\alpha)$ with $B_n(0) = V(0)$ for all n. In particular, for $\overline{b}_n = B_n(1)$, $\overline{b}_{\underline{n}} < \cdots < \overline{b}_{\overline{n}} < \overline{v}$.

(ii). The c.d.f. $G(\cdot)$ has a p.d.f. $g(\cdot)$ with $g(\underline{v}) = 0$, which is continuous over the straight line with the exception of the $\overline{n} - \underline{n} + 1$ discontinuity points $\overline{b}_{\underline{n}} < \cdots < \overline{b}_{\overline{n}}$, the interval $[\underline{v}, \overline{b}_{\overline{n}}]$ being the support of $G(\cdot)$. For $\underline{n} \leq n \leq \overline{n}$, the jumps

[1] This feature also does not hold when buyers do not not observe N, in which case the model primitives are not identified. For the sake of brevity, we do not establish here that observation of competition by buyers is essential for purposes of identification. This can be done as in Proposition 3.2-(ii), which considers a reserve price model wherein active buyers cannot observe competition. Haile, Hong and Shum (2003) have also used support variation to test a common value setting against a private value one.

$\Delta_n = \lim_{t\downarrow 0}\left(g(\bar{b}_n - t) - g(\bar{b}_n + t)\right)$ *satisfy*

$$\Delta_n = \frac{np_n}{(n-1)(\bar{v} - \bar{b}_n)}. \tag{9}$$

(iii). It holds that $\underline{n} = \lim_{t\downarrow 0} \frac{\log G(\underline{v}+t)}{\log t}$.

Proof of Lemma 2.1: As

$$B_n(\alpha) = \frac{1}{\alpha^{n-1}} \int_0^{\alpha} V(t)\, d\left[t^{n-1}\right] = V(\alpha) - \int_0^{\alpha} \left(\frac{t}{\alpha}\right)^{n-1} V^{(1)}(t)\, dt$$

differentiating with respect to n gives

$$\frac{\partial B_n(\alpha)}{\partial n} = -\int_0^{\alpha} \left(\frac{t}{\alpha}\right)^{n-1} \log\left(\frac{t}{\alpha}\right) V^{(1)}(t)\, dt \geq 0,$$

the inequality is strict except when $\alpha = 0$, in which case $B_n(0) = \underline{v}$ for all n. It follows that the bid c.d.f. $G_n(\cdot)$ given that $N = n$ has a support $[\underline{v}, \bar{b}_n]$, with an upper bound $\bar{b}_n = B_n(1)$ which is strictly increasing with respect to n and strictly smaller than $\bar{v} = \lim_{n\uparrow\infty} \bar{b}_n$. Hence, this proves Part (i). For part (ii) The expression for jumps (9) follows from (5), which shows that the winning bid p.d.f. is

$$g(b) = \sum_{k=\underline{n}}^{\bar{n}} p_k k G_k^{k-1}(b)\, g_k(b), \tag{10}$$

with $g_k(b) = 0$ for $b > \bar{b}_n$ when $k \leq n$ by Lemma 2.1-(i). This gives

$$\begin{aligned}
&g(\bar{b}_n - t) - g(\bar{b}_n + t) \\
&\quad = \sum_{k=n}^{\bar{n}} p_k k G_k^{k-1}\left(\bar{b}_n - t\right) g_k\left(\bar{b}_n - t\right) - \sum_{k=n+1}^{\bar{n}} p_k k G_k^{k-1}\left(\bar{b}_n + t\right) g_k\left(\bar{b}_n + t\right) \\
&\quad \to np_n g_n\left(\bar{b}_n\right) = \Delta_n
\end{aligned}$$

when t goes to 0. The equality (7) for $g_n\left(\bar{b}_n\right)$ then gives (9). For part (iii), continuity of $B_n^{(1)}(\cdot)$, which is bounded away from 0 and infinity, and (8) shows that $g_n(\cdot)$ is continuous with $g_n(\underline{v}) > 0$ by (6). This gives when t goes to 0

$$G(\underline{v}+t) = \sum_{n=\underline{n}}^{\overline{n}} p_n \left(\int_{\underline{v}}^{\underline{v}+t} g_n(u)du \right)^n = \sum_{n=\underline{n}}^{\overline{n}} p_n g_n^n(\underline{v}) t^n (1+o(1))$$

$$= p_{\underline{n}} g_{\underline{n}}^{\underline{n}}(\underline{v}) t^{\underline{n}} (1+o(1))$$

as $p_{\underline{n}} g_{\underline{n}}^{\underline{n}}(\underline{v}) > 0$, which implies $\underline{n} = \lim_{t\downarrow 0} \frac{\log G(\underline{v}+t)}{\log t}$. □

Lemma 2.1 is an important building block for identifying the competition distribution. Part (iii) is a tail identification result for $\underline{n}$ as in Hill and Shneyerov (2013). Lemma 2.1-(ii) shows that the jumps in the winning bid p.d.f. identify $\mathbb{P}(N = n)$ up to the unknown $\overline{v}$.

2.5 Identification

Here, we first illustrate our identification procedure using the numerical example of Section 2.4.1. We then turn to the identification of competition and the private value distribution.

2.5.1 Numerical example (cont'd)

By Lemma 2.1-(iii), $\underline{n} = 2$ is identified. The winning bid c.d.f. of Figure 2a has two kinks, which identify the upper bounds $\overline{b}_2 < \overline{b}_3$, and then $\overline{n} = 3$. Consider now the jumps Δ_2 and Δ_3 of the winning bid p.d.f. in Figure 2b. Expression (9) for the jumps allows for the identification of $p_1 = p_2 = \frac{1}{2}$; see Lemma 2.2.

Let us now turn to the identification of the private value distribution, which is based on the winning bid c.d.f.

$$G(b) = \frac{1}{2} \left(G_2^2(b) + G_3^3(b) \right)$$

displayed in Figure 2a. Since $G_2^2(b) = 1$ on $[\overline{b}_2, \overline{b}_3]$,

$$G_3(b) = (2G(b) - 1)^{\frac{1}{3}}, \quad b \in [\overline{b}_2, \overline{b}_3].$$

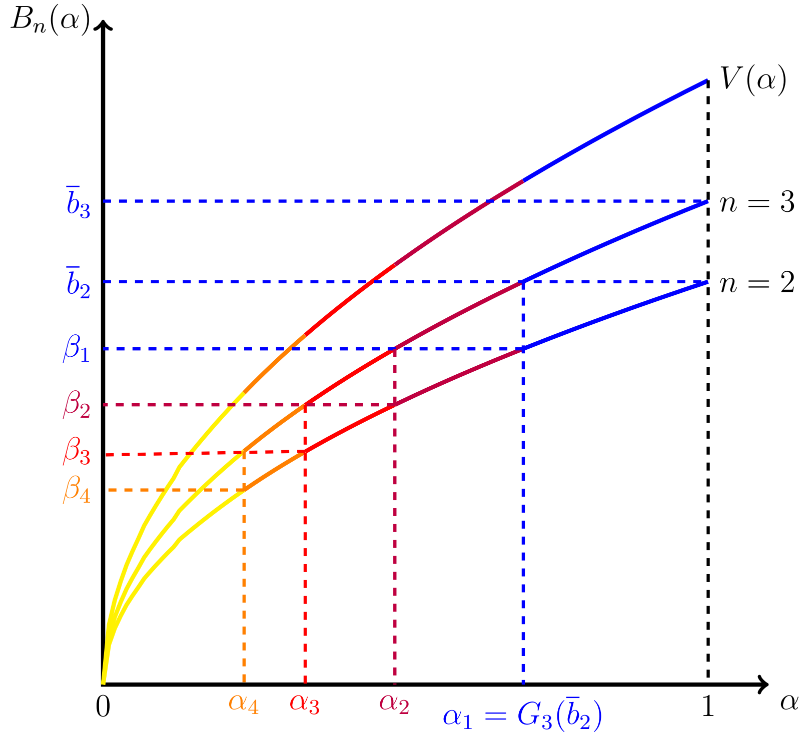


Figure 3: Iterative identification of $B_n(\alpha)$ and $V(\alpha)$ from $G(\cdot)$, as in Figure 2.

It follows that $B_3(\cdot)$ is identified on $[\alpha_1, 1]$, where $\alpha_1 = G_3(\bar{b}_2)$, using the top portion of the winning bid distribution; see Figure 2a when $G(b) \in [\frac{1}{2}(1 + G_3^3(\bar{b}_2)), 1]$. Using the mapping (2) from the bid quantile function to the private value one gives

$$V(\alpha) = B_3(\alpha) + \alpha B_3^{(1)}(\alpha),$$

and $V(\cdot)$ is identified on $[\alpha_1, 1]$. Additionally, using the mapping (4) from the private value quantile function gives

$$B_2(\alpha) = \frac{1}{\alpha}\left[\bar{b}_2 - \int_\alpha^1 V(t)dt\right]$$

so that $B_2(\cdot)$ is also identified on $[\alpha_1, 1]$. The identified $B_2(\alpha)$, $B_3(\alpha)$, and $V(\alpha)$, where $\alpha \in [\alpha_1, 1]$, are displayed in blue in Figure 3.

Next, we enlarge the interval $[\alpha_1, 1]$ over which $V(\cdot)$ is identified. For this purpose, let $\beta_1 = B_2(\alpha_1) < \bar{b}_2$ and observe that $G_2(b)$ is identified for $b \geq \beta_1$. Since

$$G_3(b) = \left(G_2^2(b) - 2G(b)\right)^{\frac{1}{3}},$$

$G_3(b)$ is identified for $b \geq \beta_1$, as $B_3(\alpha)$ is identified for $\alpha \geq \alpha_2 = G_3(\beta_1)$. Figure 3 shows that $\alpha_2 < \alpha_1$ and arguing as above gives us identification of $V(\cdot)$ and $B_2(\cdot)$ on $[\alpha_2, 1]$. Three portions of $V(\cdot)$, $B_3(\cdot)$, and $B_2(\cdot)$ are identified through three iterations and plotted in purple, red and orange, respectively, in Figure 3. Furthermore, Figure 3 suggests that additional iterations of this identification procedure should allow us to recover any $V(\alpha)$.

2.5.2 Identification of the distribution of N

In this subsection, we describe the identification of the support of N and its distribution using the discontinuity points and jump sizes. To identify the support, we exploit two implications of Lemma 2.1: (1) the minimum number of buyers $\underline{n}$ is identified from the winning bid distribution tail near the lower boundary; (2) each number of buyers n generates a discontinuity in the winning bid distribution, which identifies the difference $\overline{n} - \underline{n}$. More specifically, Lemma 2.1-(ii) identifies $\underline{n}$ and $\overline{n}$ through $\underline{n} = \lim_{t \downarrow 0} \frac{\log G(\underline{v}+t)}{\log t}$ and

$$\overline{n} = \underline{n} + \mathsf{Card}\left\{b; g\left(\cdot\right) \text{ is discontinuous at } b\right\} - 1.$$

This also identifies the support of the distribution of N as $\mathbb{P}(N = n) > 0$ for all n with $\underline{n} \leq n \leq \overline{n}$ by Assumption N.

Next, we exploit the jumps in the p.d.f. to identify $p_n = \mathbb{P}(N = n)$. Recall that Equation (9) identifies p_n up to the private value upper bound $\overline{v}$,

$$p_n = \frac{n-1}{n} \Delta_n \left(\overline{v} - \bar{b}_n\right).$$

But $\sum_{\underline{n}}^{\overline{n}} p_n = 1$ implies

$$\overline{v} = \frac{1 + \sum_{n=\underline{n}}^{\overline{n}} \frac{n-1}{n} \Delta_n \overline{b}_n}{\sum_{\underline{n}}^{n=\overline{n}} \frac{n-1}{n} \Delta_n}. \tag{11}$$

Hence p_n satisfies

$$p_n = \frac{\frac{n-1}{n}\Delta_n}{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k} + \frac{n-1}{n}\Delta_n \left(\frac{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k \overline{b}_k}{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k} - \overline{b}_n \right), \quad n = \underline{n}, \dots, \overline{n}, \tag{12}$$

and is identified because the discontinuity points $\overline{b}_k$ and jump sizes Δ_k are. We summarize these identification results in the next lemma.

Lemma 2.2 *Suppose Assumptions N and IPV hold. Then $\underline{v}$, $\overline{n}$, $\underline{n}$, $\overline{b}_{\underline{n}} < \cdots < \overline{b}_{\overline{n}}$, $\overline{v}$ and the probabilities p_n, $n = \underline{n}, \dots, \overline{n}$, are identified.*

The identifying equations (11) and (12) can also be used to derive inequality constraints satisfied by the jumps sizes Δ_n, discontinuity locations $\overline{b}_n$, the lowest and largest numbers of bidders $\underline{n}$ and $\overline{n}$. Indeed $\overline{v} > \overline{b}_{\overline{n}}$ and $0 \leq p_n \leq 1$ are equivalent to the equations[2]

$$\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k \left(\overline{b}_{\overline{n}} - \overline{b}_k\right) \leq 1,$$

$$1 + \sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k \left(\overline{b}_k - \overline{b}_n\right) \leq \frac{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k}{\frac{n-1}{n}\Delta_n}, \quad n = \underline{n}, \dots, \overline{n}$$

given that it must also holds that $\Delta_n > 0$ by Lemma 2.1-(ii). A violation of any of these inequalities indicate that the model is not correct.

2.5.3 Identification of the private value quantile function

We now obtain identification of the value quantile function by iteratively exploiting the two equilibrium mappings in (2) and (4). We proceed in three steps:

[2] It is easily seen that $\overline{b}_{\overline{n}} < \overline{v}$, which is equivalent to the first inequality, implies $0 \leq p_n$, $n = \underline{n}, \dots, \overline{n}$. The second inequalities are equivalent to $p_n \leq 1$, $n = \underline{n}, \dots, \overline{n}$.

Step 1. Note that the winning bid distribution satisfies

$$G(b) = 1 - p_{\overline{n}} + p_{\overline{n}} G_{\overline{n}}^{\overline{n}}(b) \text{ for all } b \text{ in } \left[\overline{b}_{\overline{n}-1}, \overline{b}_{\overline{n}}\right]$$

so that $G_{\overline{n}}(\cdot)$ is identified over $\left[\overline{b}_{\overline{n}-1}, \overline{b}_{\overline{n}}\right]$ as follows

$$G_{\overline{n}}(b) = \left(\frac{G(b) - (1 - p_{\overline{n}})}{p_{\overline{n}}}\right)^{1/\overline{n}} \text{ for } b \text{ in } \left[\overline{b}_{\overline{n}-1}, \overline{b}_{\overline{n}}\right].$$

Set

$$\alpha_1 = G_{\overline{n}}\left(\overline{b}_{\overline{n}-1}\right).$$

It follows that $B_{\overline{n}}(\cdot)$ is identified on $[\alpha_1, 1]$, i.e.,

$$B_{\overline{n}}(\alpha) = W\left[(1 - p_{\overline{n}}) + p_{\overline{n}}\alpha^{\overline{n}}\right],$$

where $W(\cdot) = G^{-1}(\cdot)$ is the winning bid quantile function.

Using the mapping from the bid quantile function to the value quantile function (2) shows that the private value quantile function satisfies, for all $\alpha \in [\alpha_1, 1]$,

$$\begin{aligned} V(\alpha) &= B_{\overline{n}}(\alpha) + \frac{1}{\overline{n}-1}\alpha B_{\overline{n}}^{(1)}(\alpha) \\ &= W\left[(1 - p_{\overline{n}}) + p_{\overline{n}}\alpha^{\overline{n}}\right] + \frac{\overline{n}p_{\overline{n}}}{\overline{n}-1}\alpha^{\overline{n}} W^{(1)}\left[(1 - p_{\overline{n}}) + p_{\overline{n}}\alpha^{\overline{n}}\right], \end{aligned}$$

and $V(\cdot)$ is identified over $[\alpha_1, 1]$.

Using the mapping from the value quantile function to the bid quantile function (4) shows that the bid quantile functions $B_n(\cdot)$, $n = \underline{n}, \ldots, \overline{n} - 1$ are also identified over $[\alpha_1, 1]$. Hence $\{B_n(\alpha), \alpha \in [\alpha_1, 1]\}$ and $\left\{G_n(b), b \in [B_n(\alpha_1), \overline{b}_n]\right\}$ are identified, for all $n = \underline{n}, \ldots, \overline{n}$.

Step 2. We now expand the interval over which $G_{\overline{n}}(\cdot)$ is identified using an iterative argument. Define

$$\beta_1 = B_{\overline{n}-1}(\alpha_1),$$

which is identified from the last step. Note that $\beta_1 < \overline{b}_{\overline{n}-1}$ whenever $\alpha_1 > 0$ because, by Lemma 2.1-(i),

$$\beta_1 = B_{\overline{n}-1}\left[G_{\overline{n}}\left(\overline{b}_{\overline{n}-1}\right)\right] < B_{\overline{n}}\left[G_{\overline{n}}\left(\overline{b}_{\overline{n}-1}\right)\right] = \overline{b}_{\overline{n}-1}.$$

The definition of $G(\cdot)$ implies that

$$G_{\overline{n}}(b) = \left(\frac{G(b) - \sum_{n=\underline{n}}^{\overline{n}-1} p_n G_n^n(b)}{p_{\overline{n}}}\right)^{1/\overline{n}}, \tag{13}$$

where $G(\cdot)$ and p_n are identified, and $G_n(\cdot)$ are identified on $[B_n(\alpha_1), \overline{b}_{\overline{n}}]$ for all $n = \underline{n}, \ldots, \overline{n}-1$. Since $B_{\underline{n}}(\alpha_1) < \ldots < B_{\overline{n}-1}(\alpha_1) = \beta_1$, $[\beta_1, \overline{b}_{\overline{n}}] \subseteq [B_n(\alpha_1), \overline{b}_{\overline{n}}]$ for all n. Therefore, the conditional bid distribution $G_{\overline{n}}(b)$ is identified on $[\beta_1, \overline{b}_{\overline{n}}]$.

Step 3. We now identify $V(\cdot)$ on a growing interval $[\alpha_k, 1]$ using an induction argument and the identified $V(\cdot)$ on $[\alpha_1, 1]$. For an integer $k \geq 2$, define

$$\alpha_k = G_{\overline{n}}(\beta_{k-1}) = G_{\overline{n}}\left[B_{\overline{n}-1}(\alpha_{k-1})\right], \quad \beta_k = B_{\overline{n}-1}(\alpha_k).$$

Identification of $V(\cdot)$ on the growing interval $[\alpha_k, 1]$ is established in Lemma 2.3 below.

Lemma 2.3 *Suppose Assumptions N and IPV hold. Then*

(i). the sequences $\{\alpha_k, k \geq 1\}$ and $\{\beta_k, k \geq 1\}$ are decreasing sequences with

$$\lim_{k\to\infty} \alpha_k = 0.$$

(ii). $\{\alpha_k, k \geq 1\}$ is identified. For any integer number $k \geq 2$, $\{V(\alpha), \alpha \in [\alpha_k, 1]\}$ is identified if $\{V(\alpha), \alpha \in [\alpha_{k-1}, 1]\}$ is identified.

Proof of Lemma 2.3: See Section 6.1 in the proof section.

Combining Lemmas 2.2 and 2.3 shows that the joint distribution of private values and the number of active buyers is identified.

Theorem 2.1 *Suppose Assumptions N and IPV hold and that the buyers observe the number of active buyers N. Then $F(\cdot)$ and the distribution of N are identified.*

Proof of Theorem 2.1: Lemma 2.2 shows that the distribution of N is identified. Since N is independent of private values, it remains to be shown that the private value quantile function $V(\cdot)$ is identified. Take $\alpha > 0$. By Lemma 2.3-(i), there exists k such that $\alpha > \alpha_k$ and Lemma 2.3-(ii) yields identification of $V(\alpha)$. Since $V(0) = \underline{v}$ is identified by Lemma 2.2, the Theorem is proven. $\square$

3 Endogenous competition

So far, we have assumed exogenous participation, i.e., the number of buyers being independent of private values. In many cases of interest, participation arises endogenously, such as when the seller imposes a reserve price below which bids will not be considered, or when buyers face entry costs. This section extends our results to endogenous competition due to a reserve price or entry cost.

In these cases, the breadth of possible outcomes is richer than in our benchmark model. In particular, the auctioned object may not be sold when buyers decide not to enter the auction, as accounted in the next Definition of the *outcome distribution*.

Definition 3.1 *A distribution $\mathbb{G}(\cdot)$ is an* ***outcome distribution*** *if and only if it assigns probabilities solely to the events in the σ-field generated by*

$$\{\textit{the object is not sold}\}, \quad \{\textit{the object is sold at a price less than } b\}, \quad b \in \mathbb{R}.$$

In the sequel, these two events will be abbreviated as $\{\text{Not Sold}\} = \{N = 0\}$, also refereed a *failed auction* later on, and $\{W \leq b\}$, respectively, where W stands for the winning bid or, more generally, the transaction price. The winning bid distribution of the benchmark model, which attributes zero probability to {Not Sold} is an example of a continuous auction outcome distribution.

Additionally, Definition 3.1 allows for the following two kinds of discrete components. First, the object may remain unsold with positive probability. Second, $\mathbb{G}(\cdot)$ can have discrete components even when the object is sold. For instance, if a unique buyer enters an auction with a reserve price R and observes an absence of competitors, she can win the auction by bidding R, such that $\mathbb{P}(W = R) > 0$ if $\mathbb{P}(N = 1) > 0$. In this reserve price example, as in many others, the (renormalized) continuous component of $\mathbb{G}(\cdot)$, denoted as $G(\cdot)$, is the distribution of the winning bid given that there is at least two participants.

That the model allows for these additional outcomes raises the issue of econometric selection. Consider, for instance, winning bids data derived from firms' accounting books. A failed auction with an unsold object will be absent from the records as no transaction took place. Furthermore, a seller may be tempted to cancel auctions that are attended by a sole buyer who wins simply by bidding the reserve price. In this case, identification of the model should be based not on $\mathbb{G}$ but on a conditional outcome distribution given that the events above do not happen. In the reserve price and entry cost examples considered below, the latter conditional outcome distribution coincides with the continuous component $G(\cdot)$ of $\mathbb{G}(\cdot)$.

3.1 First-price auction with reserve price

In this subsection, we consider endogenous competition due to the presence of a reserve price. Assume that the seller will not sell the item if the maximum bid is lower than a reserve price R, which is known to the buyers. Only buyers with $V_i \geq R$, or equivalently $A_i \geq F(R)$ can win, and are called *participants* or active buyers. We assume that the analyst only observes the winning bid $\max_{i \in \mathcal{N}} B_i$ when the auctioned good is sold, does not know the reserve price R, the number $\overline{n} \geq 2$ of potential buyers[3] or the number N_R of participants (except when failed auction with $N_R = 0$ are observed).

The number of participants N_R has a binomial distribution with parameter $(\overline{n}, 1 - F(R))$; that is

$$\mathbb{P}(N_R = n) = \frac{\overline{n}!}{n!(\overline{n} - n)!} [1 - F(R)]^n F^{\overline{n}-n}(R).$$

Because the buyers participating in the auction receive a signal $A_i \geq F(R)$, it is convenient to renormalize the signal as

$$A_{i,R} = \frac{A_i - F(R)}{1 - F(R)} \tag{14}$$

assuming $\underline{v} \leq R < \overline{v}$ from now on. Given $N_R = n \geq 1$, the auction participants' signals $A_{i,R}$ are *i.i.d.* draws from the $[0, 1]$ uniform distribution. Let $F_R(\cdot)$ and $V_R(\cdot)$ be the conditional c.d.f. and quantile function of the private values V_i given $V_i \geq R$, such that

$$\begin{aligned} F_R(v) &= \begin{cases} 0 & \text{if } v < R \\ \frac{F(v)-F(R)}{1-F(R)} & \text{if } R \leq v \leq \overline{v} \\ 1 & \text{if } v > \overline{v} \end{cases}, \\ V_R(\alpha) &= V\left[F(R) + (1 - F(R))\,\alpha\right]. \end{aligned}$$

As discussed for the benchmark model, the winning bid p.d.f is discontinuous when the participants observe N_R before submitting bids, a property that plays an important role for identification. We therefore consider separately the cases where participation is known or unknown to active buyers.

[3] There is no need to identify the minimal number of participants $\underline{n}$ among the parameters as the lower bound of N_R is known to be 0.

3.1.1 Participation known to active buyers

First, let us recall the expression of the best-response bidding strategy when the participants observe $N_R = n$. If $n = 0$, the auctioned object is not sold. If $n = 1$, it is optimal for the unique participant to bid the reserve price R. If $n \geq 2$, the optimal bidding strategy is a strictly increasing function of $A_{i,R}$ under Assumption IPV, which is equal to the bid quantile equation $B_{n,R}(\cdot) = B(\cdot|N = n, R)$, given $N_R = n \geq 2$. Arguing as in (2), except using $A_{i,R}$ and $V_R(\cdot)$ in place of A_i and $V(\cdot)$, yields the best response differential equation

$$B_{n,R}(\alpha) + \frac{\alpha}{n-1} B_{n,R}^{(1)}(\alpha) = V_R(\alpha). \tag{15}$$

It follows that

$$B_{n,R}(\alpha) = \frac{n-1}{\alpha^{n-1}} \int_0^\alpha t^{n-2} V_R(t) dt = V_R(\alpha) - \int_0^\alpha \left(\frac{t}{\alpha}\right)^{n-1} V_R^{(1)}(t) dt. \tag{16}$$

Structure of the outcome distribution. We describe the constraints on the distribution of observables imposed by the reserve price. The continuous component $G(\cdot)$ of the outcome distribution is the conditional winning bid distribution given $W > R$,[4] such that

$$G(b) = \mathbb{G}\left(W \leq b \,|W > R\right). \tag{17}$$

Proposition 3.1 *An outcome distribution $\mathbb{G}(\cdot)$ is rationalized by a first-price auction model satisfying Assumptions IPV where buyers observe the number of participants and with a reserve price R in $(\underline{v}, \overline{v})$ if and only if*

(i). There is a probability $q \in (0,1)$, an integer number $\overline{n} \geq 2$ with $\mathbb{G}(\text{Not sold}) = q^{\overline{n}}$, and a unique $R > 0$ such that $\mathbb{G}(W = R) > 0$ and is equal to $\overline{n} q^{\overline{n}-1}(1-q)$.

[4] Equivalently, $N_R \geq 2$. Using the event $\{W \geq R\}$ instead is more convenient for identification purposes, as W is observed by the analyst and R is identified as the lower bound of the support of $\mathbb{G}(\cdot)$.

(ii). Let $G(\cdot)$ be as in (17). There exists some c.d.f. $G_n(\cdot)$ such that

$$G(\cdot) = \sum_{n=2}^{\overline{n}} p_n G_n^n(\cdot), \quad p_n = \frac{\overline{n}!}{n!\,(\overline{n}-n)!} \frac{q^{\overline{n}-n}(1-q)^n}{1-q^{\overline{n}}-\overline{n}q^{\overline{n}-1}(1-q)}.$$

(iii). The quantile functions $B_n(\cdot) = G_n^{-1}(\cdot)$ satisfy the compatibility conditions of Proposition 2.1-(ii) with $B_n(0) = R$.

Proof of Proposition 3.1: See Section 6.2 in the proof section.

Proposition 3.1 is very similar to Proposition 2.1 of the benchmark model, up to the mixture weights p_n which are now given by a binomial distribution.

Identification results. The compatibility condition of Proposition 3.1-(iii) induces discontinuities in the p.d.f. of $G(\cdot)$. This gives the next identification corollary, which mostly follows from Theorem 2.1. Corollary 3.1-(ii) shows that the model primitives are identified from well functioning auctions with more than two participants provided the number of potential buyers is greater than or equal to three.

Corollary 3.1 *Consider a first-price auction satisfying Assumption IPV with a reserve price R in $(\underline{v}, \overline{v})$ and N_R participants, where N_R is known to the buyers but not to the analyst. Then*

(i). the reserve price R, number of potential buyers $\overline{n}$ and conditional private value distribution $F_R(\cdot)$ are identified from the c.d.f. $G(\cdot)$ defined in (17).

(ii). if $\overline{n} \geq 3$, the private value c.d.f. $\{F(v), v \geq R\}$ is identified from the c.d.f. $G(\cdot)$.

(iii). if $\overline{n} = 2$, the c.d.f $\{F(v), v \geq R\}$ is identified from the outcome distribution $\mathbb{G}(\cdot)$ but not from $G(\cdot)$.

Proof of Corollary 3.1: See Section 6.3 in the proof section.

Corollary 3.1-(i) follows from Theorem 2.1 as $F_R(\cdot)$ is the updated distribution of the private value given participation. Identification of $F(R)$ from $G(\cdot)$ requires at least three potential buyers. This discrepancy between $\overline{n} = 2$ and $\overline{n} \geq 3$ can be understood using Lemma 2.2, which shows here that, for $\overline{n} \geq 3$, the probabilities

$$p_n(R) = \mathbb{P}\left(N_R = n \,|N_R \geq 2\right) = \frac{\frac{\overline{n}!}{n!(\overline{n}-n)!} F^{\overline{n}-n}(R)\left[1 - F(R)\right]^n}{1 - F^{\overline{n}}(R) - \overline{n}F^{\overline{n}-1}(R)\left[1 - F(R)\right]}, \quad n = 2, \ldots, \overline{n}$$

are identified through the discontinuities of the p.d.f. $g(\cdot)$ of $G(\cdot)$. This is sufficient to identify $F(R)$ and then to identify $F(v)$ for $v \geq R$.

This contrasts with $\overline{n} = 2$, in which case $g(\cdot)$ is continuous over its inner support. The c.d.f. $G(\cdot)$ is helpful for recovering $F_R(\cdot)$ but not informative enough regarding the screening level $F(R)$ due to the lack of density discontinuities. As Corollary 3.1-(i) shows that $\overline{n}$ is identified, $F(R)$ can be recovered from $\mathbb{G}\left(\text{Not sold}\right)$, and observing failed auctions is necessary to identify $\{F(v), v \geq R\}$.[5]

More generally, the proof of Corollary 3.1-(ii,iii) shows that $G(\cdot)$ overidentifies $F(R)$ when $\overline{n} \geq 3$, while $\mathbb{G}(\cdot)$ overidentifies $F(R)$ for all $\overline{n} \geq 2$. This can be used to test the null hypothesis that a reserve price restricts participation.

3.1.2 Participation unknown to active buyers

Winning bid density discontinuities arise because buyers observe the number of participants N_R before submitting their bids. If they do not observe N_R, the winning bid density becomes continuous on its inner support. However, it is shown that identification still holds if the analyst observes whether the object is left unsold, whether $N_R = 1$ or if an instrument for the reserve price is available.

[5]Since $\mathbb{G}\left(W = R\right) = 2F\left(R\right)\left(1 - F\left(R\right)\right)$ only identifies the set $\left\{F\left(R\right), 1 - F\left(R\right)\right\}$ when $\overline{n} = 2$, , auctions with a good sold at the reserve price to a unique bidder cannot be used to recover $F(R)$.

We first recall the quantile expression of the optimal bidding strategy. Because buyers participate in the auction if and only if $A_i \geq R$, the optimal bidding strategy $B_R(\cdot)$ is a strictly increasing function of the normalized uniform signals $A_{i,R}$ in (14), which is also the bid quantile function. Suppose now that Buyer 1 participates and decides to place a bid $B_R(a)$ given her signal $A_{1,R} = \alpha$. Given Buyer 1's participation, the number of participants in $\mathcal{N} \setminus \{1\}$, $N_R - 1$, has a Binomial distribution with parameter $(\overline{n} - 1, 1 - F(R))$ and is independent of participant signals $A_{i,R}$. It follows that the probability that Buyer 1 wins is

$$\begin{aligned} &\mathbb{P}\left(\max_{i \in \mathcal{N} \setminus \{1\}} A_{i,R} \leq a \,|1 \text{ participates} \right) \\ &= \sum_{n=0}^{\overline{n}-1} a^n \frac{(\overline{n}-1)!}{n!(\overline{n}-1-n)!} \left[1 - F(R)\right]^n F^{\overline{n}-1-n}(R) \end{aligned}$$

which is equal to $\left[F(R) + (1 - F(R))\, a\right]^{\overline{n}-1}$. Hence the expected payoff of Buyer 1 is

$$(V_R(\alpha) - B_R(a)) \left[F(R) + (1 - F(R))\, a\right]^{\overline{n}-1}.$$

Arguing as for the best response differential equation (2) shows that

$$B_R(\alpha) + \left(\alpha + \frac{F(R)}{1 - F(R)} \right) \frac{B_R^{(1)}(\alpha)}{\overline{n} - 1} = V_R(\alpha), \quad B_R(0) = R, \tag{18}$$

of which the unique solution is

$$B_R(\alpha) = V_R(\alpha) - \int_0^{\alpha} \left(\frac{F(R) + [1 - F(R)]t}{F(R) + [1 - F(R)]\alpha} \right)^{\overline{n}-1} V_R^{(1)}(t) dt. \tag{19}$$

Structure of the outcome distribution. Proposition 3.2 and Lemma 3.1 below summarize the model implications for the auction outcome distribution. In these results, $G_R(\cdot)$ stands for $B_R^{-1}(\cdot)$ and $G(\cdot)$ for the continuous component $\mathbb{G}(\cdot|\text{Sold})$ of the outcome distribution. Lemma 3.1 extends a result of Guerre et al. (2000) on the divergence of the bid p.d.f. at the vicinity of the reserve price.

Proposition 3.2 *An outcome distribution* $\mathbb{G}(\cdot)$ *is rationalized by a first-price auction model satisfying Assumptions IPV, with a reserve price* R *in* $(\underline{v}, \overline{v})$*, and buyers who observe the number of participants if and only if*

(i). There is a probability $q \in (0,1)$*, and an integer number* $\overline{n} \geq 2$ *with* $\mathbb{G}(Not\ sold) = q^{\overline{n}}$*. The c.d.f.*

$$G_R(b) = \frac{\left[(1-q^{\overline{n}})\,\mathbb{G}\left(W \leq b\,|Sold\right) + q^{\overline{n}}\right]^{1/\overline{n}} - q}{1-q}$$

has a support $[R, \overline{b}_R]$ *for some* $R < \overline{b}_R < \infty$*.*

(ii). $B_R(\cdot) = G_R^{-1}(\cdot)$ *is continuously differentiable with* $B_R^{(1)}(0) = 0$*. The function*

$$B_R(\alpha) + \left(\alpha + \frac{q}{1-q}\right)\frac{B_R^{(1)}(\alpha)}{\overline{n}-1}$$

has a continuous derivative which is strictly positive over $[0,1]$*.*

Proof of Proposition 3.2: See Section 6.4 in the proof section.

Lemma 3.1 *The c.d.f.* $G_R(b)$ *and* $G(b)$ *have p.d.f.* $g_R(b)$ *and* $g(b)$*, respectively, which are continuous over* $(R, \overline{b}_R]$ *and diverge when* b *goes to* R *with*

$$0 < \lim_{b \downarrow R}\left\{(b-R)^{1/2} g_R(b)\right\} < \infty, \quad 0 < \lim_{b \downarrow R}\left\{(b-R)^{1/2} g(b)\right\} < \infty.$$

Proof of Lemma 3.1: See Section 6.5 in the proof section.

Compared to the case where the participants observe N_R, the outcome distribution $\mathbb{G}(\cdot)$ puts no mass at the reserve price, $\mathbb{G}(W = R) = 0$. Proposition 3.2-(ii) implies that the winning bid p.d.f. $g(\cdot)$ and $g_R(\cdot)$ are continuous on their inner support. Lemma 3.1 shows, however, that these winning bid p.d.f. are infinite at the reserve price. These features can be used to test whether or not the buyers observe the level of competition.

Identification results. Proposition 3.2-(i) shows that the winning bid c.d.f. $G(\cdot)$ has a mixture structure; that is, since $q = F(R)$,

$$\begin{aligned} G(b) &= \frac{\left[F(R) + (1 - F(R))\, G_R(b)\right]^{\overline{n}} - F^{\overline{n}}(R)}{1 - F^{\overline{n}}(R)} \\ &= \sum_{n=1}^{\overline{n}} \frac{\overline{n}!}{n!(\overline{n} - n)!} \frac{\left[1 - F(R)\right]^{n} F^{\overline{n}-n}(R)}{1 - F^{\overline{n}}(R)} G_R^n(b). \end{aligned}$$

The c.d.f. mixture components are powers of the same c.d.f. $G_R(\cdot)$, so that this winning bid distribution looks simpler than it counterpart (17) of the case where buyers observe participation. However, due to the absence of support variation across mixture components, it lacks the p.d.f discontinuities that are crucial to identifying the distribution of N_R when buyers observe participation. This explains the non identification result of Proposition 3.3-(ii).

Proposition 3.3 *Consider a first-price auction satisfying Assumption IPV with a reserve price R in $(\underline{v}, \overline{v})$ and a number N_R of participants unknown to buyers and the analyst. Then*

(i). the reserve price R is identified from the winning bid c.d.f. $G(\cdot)$.

(ii). the private value c.d.f. $\{F(v), v \geq R\}$ and the maximal number $\overline{n}$ of participants are not identified from the outcome distribution $\mathbb{G}(\cdot)$.

(iii). $\{F(v), v \geq R\}$ and $\overline{n}$ are identified from $\mathbb{G}(\cdot)$ and $\mathbb{P}(N_R \geq 2)$ if the analyst observes the auction outcomes and whether there were at least two participants.

Proof of Proposition 3.3: See Section 6.6 in the proof section.

As seen from the simple mixture structure of the distribution $G(\cdot)$, $G_R(\cdot)$ can be recovered as soon as $F(R)$ and $\overline{n}$ are identified, which ensure identification of the quantile function $V_R(\cdot)$ by (18). This is used in Proposition 3.3-(iii) to restore identification as $F(R)$ and $\overline{n}$ can be identified from

$$\begin{aligned} F^{\overline{n}}(R) &= \mathbb{P}(N_R = 0) = \mathbb{G}(\text{Not sold}), \\ \overline{n}F^{\overline{n}-1}(R)\,[1 - F(R)] &= \mathbb{P}(N_R = 1) = 1 - \mathbb{P}(N_R \geq 2) - \mathbb{G}(\text{Not sold}). \end{aligned}$$

Alternatively, identification also holds under an exclusion restriction, which considers an instrument z that affects the reserve price but not the other primitives of the model, as assumed later on for the cost of an entry model.

Assumption R. *The reserve price is a non-constant function $R(z)$ of the instrument $z \in \mathcal{Z}$. The private value c.d.f. $F(\cdot)$ and $\overline{n}$ do not depend upon z.*

Proposition 3.4 shows that, thanks to variation in the instrument, model primitives can be identified from the conditional outcome distribution $\{\mathbb{G}(\cdot|z), z \in \mathcal{Z}\}$, without additional observation on N_R as in Proposition 3.3-(iii) but with the assumption that failed auctions where the object remains unsold are observable to the analyst.

Proposition 3.4 *Consider a first-price auction satisfying Assumptions IPV and R with a reserve price $R(z) \in (\underline{v}, \overline{v})$ for all $z \in \mathcal{Z}$ and N_R participants, where N_R is unknown to the buyers and the analyst. Then*

(i). $R(\cdot)$ is identified from the conditional winning bid distribution $G(\cdot|\cdot)$.

(ii). the private value c.d.f $\{F(v), v \geq \inf_{z\in\mathcal{Z}} R(z)\}$ and the maximal number $\overline{n}$ of participants are identified from the conditional outcome distribution $\mathbb{G}(\cdot|\cdot)$.

Proof of Proposition 3.4: See Section 6.7 in the proof section.

As Theorem 2.1 shows that the updated private value c.d.f. $\{F_{R(z)}(\cdot), z \in \mathcal{Z}\}$ is identified, it also follows that $F(\cdot)$ is identified from the collection $\{G(\cdot|z), z \in \mathcal{Z}\}$ of conditional winning bid c.d.f. in auctions with at least two bidders if $\inf_{z\in\mathcal{Z}} R(z) = \underline{v}$.

3.2 First-price auction with entry costs

Endogenous variation of the number of buyers can arise from entry costs. This section considers the affiliated-signal entry model of Ye (2007) considered in Marmer, Shneyerov and Xu (2013) and Gentry and Li (2014). Each auction has two stages. In the first stage, buyer i observes a private signal S_i for her unknown private value, $i = 1, \ldots, \overline{n}$. Each buyer decides whether to enter the auction. Entry involves payment of an entry cost c. In the second stage, the N_c buyers who decide to enter observe their private values V_i, which are independent draws from the conditional c.d.f. $F(\cdot|S_i)$. The active buyers then submit bids B_i in a first-price auction.

Gentry and Li (2014) assume for identification purposes that the entry cost depends upon an observed auction-specific continuous variable, $c = c(Z)$, and that Z affects neither the signals nor the private values. The signals S_i are assumed to be *i.i.d.*, from a standard uniform distribution without loss of generality. Given S_i, the private value V_i is independent of N_c, Z, and the other private values and signals. Following Gentry and Li (2014), we will use the following assumption:

Assumption E. *The number of potential entrants $\overline{n}$ is greater than or equal to 2 and does not depend upon z. The support of $F(\cdot|s,z) = F(\cdot|s)$ is $[\underline{v}, \overline{v}]$ for all $s \in [0,1]$. $F(\cdot|\cdot)$ is continuous over $[\underline{v}, \overline{v}] \times [0,1]$ with a continuous p.d.f. $f(\cdot|\cdot)$ bounded away from 0 over this set. For any $v \in [\underline{v}, \overline{v}]$, $F(v|s)$ decreases with $s \in [0,1]$. The cost function $c(\cdot)$ is non-constant and continuous over $\mathcal{Z}$, which is a closed connected set with non-empty interior.*

Gentry and Li (2014, Proposition 1) show that any symmetric Nash equilibrium of the two-stage auction game has a payoff equivalent equilibrium in which Stage 1 entry decisions involve a signal threshold $s(z) = \sigma_{\overline{n}}(c(z))$ that we detail now. Let

$F_c(v|s(z))$ be the updated private value distribution given that the buyer participates,

$$F_c(v|s) = \mathbb{P}(V_i \leq v | S_i \geq s) = \frac{1}{1-s}\int_s^1 F(v|t)\,dt.$$

The signal threshold is characterized through backward induction. For the moment, suppose that buyers do not observe the number of entrants N_c. Since the symmetric second-stage bidding strategy is increasing under Assumption IPV, the probability that a buyer with private value v wins the auction is

$$(s + (1-s)F_c(v|s))^{\overline{n}-1},$$

where $s = s(z)$ is a threshold below which buyers do not enter the auction. A buyer's post-entry expected profit is

$$\int_0^v (s + (1-s)F_c(t|s))^{\overline{n}-1}\,dt;$$

see Riley and Samuelson (1981, Eq. 8). Thus, the first-stage expected profit of a buyer with a signal S is

$$\Pi_{\overline{n}}(S;s) = \int_0^{\overline{v}} [1 - F(v|S)]\,(s + (1-s)F_c(v|s))^{\overline{n}-1}\,dv. \tag{20}$$

This expression is also valid when buyers observe the number of entrants.

By Gentry and Li (2014)'s Proposition 1, all buyers enter if $c(z) \leq \Pi_{\overline{n}}(0;0)$ and none enter if $c(z) \geq \Pi_{\overline{n}}(1;1)$. If $\Pi_{\overline{n}}(0;0) \leq c(z) \leq \Pi_{\overline{n}}(1;1)$, the threshold $s(z)$ solves the break-even condition $\Pi_{\overline{n}}(s;s) = c(z)$; that is

$$c(z) = \int_0^{\overline{v}} [1 - F(v|s(z))] \times [s(z) + (1-s(z))F_c(v|s(z))]^{\overline{n}-1}\,dv, \tag{21}$$

a condition which is useful for identifying the cost. Under the conditions imposed on $\mathcal{Z}$ in Assumption E, it follows from Gentry and Li (2014) that $s(\mathcal{Z})$ is a closed interval of $[0,1]$, which is not a singleton if $0 < s(z) < 1$, as will be assumed for the remainder of this paper.

Let us now study the identification of model primitives for each of when buyers can and cannot observe the number of entrants N_c. The analyst observes the winning bid but does not observe the number of entrants N_c. The distribution of N_c given $Z = z$ is binomial $(\overline{n}, 1 - s(z))$ such that

$$\mathbb{P}(N_c = n | Z = z) = \frac{\overline{n}!}{n!(\overline{n} - n)!}[1 - s(z)]^n s^{\overline{n}-n}(z).$$

Buyer signals are

$$A_{i,c} = F_c(V_i|s(Z)), i = 1, \ldots, N_c \tag{22}$$

which are $i.i.d.$ from the $[0, 1]$ uniform distribution and independent of N_c and Z. The updated second stage private value quantile function is $V_c(\cdot|s) = F_c^{-1}(\cdot|s)$.

3.2.1 Entry known to active buyers

Finding the best response bidding strategy when buyers observe $N_c = n$ follows the same steps as those of the reserve price case with observed participation. When $n = 0$ the object is not sold and when $n = 1$, the unique entrant can win the object with a bid of 0. When $n \geq 2$, the optimal bidding strategy is a function of n, Z, and $A_{i,c}$ from (22), which is strictly increasing with respect to the latter. The optimal bidding strategy is equal to the conditional bid quantile function $B_{n,c}(\cdot|z) = B(\cdot|N_c = n, Z = z)$. Arguing as for (2) yields

$$B_{n,c}(\alpha|z) + \frac{\alpha B_{n,c}^{(1)}(\alpha|z)}{n-1} = V_c(\alpha|s(z)), \quad B_{n,c}(0|z) = \underline{v}, \tag{23}$$

$$B_{n,c}(\alpha|z) = \frac{n-1}{\alpha^{n-1}} \int_0^\alpha t^{n-2} V_c(t|s(z))\, dt = V_c(\alpha|s(z)) - \int_0^\alpha \left(\frac{t}{\alpha}\right)^{n-1} V_c^{(1)}(t|s(z))\, dt. \tag{24}$$

Structure of the outcome distribution. The next proposition summarizes the implications of this entry model for the outcome distribution. For the remainder of this section, $G(\cdot)$ will be the continuous component of the outcome distribution, such that

$$G(b|z) = \mathbb{G}\left(W \leq b \,|W > 0, Z = z\right). \tag{25}$$

Proposition 3.5 *A conditional outcome distribution $\{\mathbb{G}(\cdot|z), z \in \mathcal{Z}\}$ is rationalized by a first-price auction model with entry cost satisfying Assumption E, where buyers observe N_c, with a continuous threshold $s(\cdot)$ such that $s(\mathcal{Z}) \subset (0,1)$, if and only if*

(i). there is a continuous $q(\cdot)$, where $q(\mathcal{Z}) \subset (0,1)$, and an integer number $\overline{n} \geq 2$ such that $\mathbb{G}(\text{Not sold}|z) = q(z)^{\overline{n}}$ and $\mathbb{G}(W = 0|z) = \overline{n} q(z)^{\overline{n}-1}(1 - q(z))$.

(ii). Let $G(\cdot|\cdot)$ be as in (25). There exists some conditional c.d.f. $G_n(\cdot|q(\cdot))$ such that

$$G(\cdot|z) = \sum_{n=2}^{\overline{n}} p_n(z)\, G_n^n(\cdot|q(z)), \quad p_n(z) = \frac{\frac{\overline{n}!}{n!(\overline{n}-n)!} q(z)^{\overline{n}-n}(1-q(z))^n}{1 - q(z)^{\overline{n}} - \overline{n} q(z)^{\overline{n}-1}(1-q(z))}.$$

(iii). The quantile functions $B_n(\cdot|q(z)) = G_n^{-1}(\cdot|q(z))$ are continuously differentiable with

$$B_n(\alpha|q(z)) + \frac{\alpha B_n^{(1)}(\alpha|q(z))}{n-1} = B_m(\alpha|q(z)) + \frac{\alpha B_m^{(1)}(\alpha|q(z))}{m-1} = V_c(\alpha|q(z))$$

for all $2 \leq n, m \leq \overline{n}$ and all $\alpha \in [0,1]$, where $V_c(\cdot|\cdot)$ belongs to the class $\mathcal{V}_{q,\underline{v}}$ of $V(\cdot|\cdot)$ satisfying: (1) $\inf_{(\alpha,z)\in[0,1]\times\mathcal{Z}} \frac{\partial V(\alpha|q(z))}{\partial \alpha} > 0$, $V(0|q(z)) = \underline{v}$ over $\mathcal{Z}$; (2) The function $-\frac{\partial}{\partial q}\left[(1-q)\, V_c^{-1}(v|q)\right]$ decreases with q over the closed interval $q(\mathcal{Z}) \subset (0,1)$ and is a conditional c.d.f. with positive density on its support closure.

Proof of Proposition 3.5: See Section 6.8 in the proof section.

The probability $q(z)$ of Proposition 3.5 is an entry probability, or equivalently, a threshold signal. Proposition 3.5-(ii) reveals that entry costs constrain the dependence between auction outcomes and entry probability: $\mathbb{G}(\cdot|z)$ is equal to the conditional outcome distribution given $q(z)$.[6] A similar sufficiency property recently appears in Liu, Vuong and Xu (2017), who characterize Bayesian Nash equilibrium in entry games.

Identification results. The identification results in the next corollary follows from the density discontinuities generated by the compatibility condition of Proposition 3.5-(iii). Corollary 3.2 parallels Corollary 3.1, which holds for reserve price first-price auction where buyers observe participation. As in the latter result, Corollary 3.2 bases identification on the winning bid c.d.f. $G(\cdot|z)$ of first-price auctions that attract at least two buyers.

Corollary 3.2 *Consider a first-price auction model with entry cost satisfying Assumption E with $s(\mathcal{Z}) \subset (0,1)$ and N_c number of entrants, where N_c is observable to buyers but not the analyst. Then*

(i). the number of potential buyers $\overline{n}$ and $\{F_c(\cdot|s(z)), z \in \mathcal{Z}\}$ are identified from the winning bid c.d.f. $G(\cdot|z)$ in (25).

(ii). if $\overline{n} \geq 3$, $\{s(z), z \in \mathcal{Z}\}$, $\{c(z), z \in \mathcal{Z}\}$ and $\{F(\cdot|s), s \in s(\mathcal{Z})\}$ are identified from the winning bid c.d.f. $G(\cdot|z)$

(iii). if $\overline{n} = 2$, $c(z)$ and $s(z)$ are identified for all $z \in \mathcal{Z}$ and $F(\cdot|s)$ is identified for all $s \in s(\mathcal{Z})$ from the outcome distribution $\mathbb{G}(\cdot|z)$.

Proof of Corollary 3.2: See Section 6.9 in the proof section.

Corollary 3.2-(i) follows from Theorem 2.1 and the fact that the second stage of this auction is a first-price auction with private values drawn in $F_c(\cdot|s(z))$. Part (ii,iii),

[6] A similar result holds in presence of a reserve price; see Proposition 3.4 and the discussion at the end of Section 3.3

which establishes identification of the primitives $F(\cdot|\cdot)$ and $c(\cdot)$ when $s(\mathcal{Z}) = [0,1]$, follows from (i) and Gentry and Li (2014, Section 3.2), who have noted that the cost function can be identified from $F_c(\cdot|s(z))$ and (21).

3.2.2 Entry unknown to active buyers

If the buyers submit their bids without observing N_c, arguing as for reserve price with unknown participation shows that the bid quantile $B_c(\cdot|z)$ satisfies

$$B_c(\alpha|z) + \left(\alpha + \frac{s(z)}{1-s(z)}\right)\frac{B_c^{(1)}(\alpha|z)}{\overline{n}-1} = V_c(\alpha|s(Z)), \quad B_R(0) = V_c(0|s(z)) = \underline{v}, \tag{26}$$

for which the unique solution is

$$B_c(\alpha|z) = V_c(\alpha|s(z)) - \int_0^\alpha \left(\frac{s(z)+[1-s(z)]t}{s(z)+[1-s(z)]\alpha}\right)^{\overline{n}-1} V_c^{(1)}(t|s(z))dt. \tag{27}$$

Structure of the outcome distribution. The next proposition and lemma summarize the implications of this entry model for the outcome distribution.

Proposition 3.6 *$\{\mathbb{G}(\cdot|z), z \in \mathcal{Z}\}$ is rationalized by a first-price auction model with entry cost satisfying Assumption E, buyers do not observe N_c, and there is a continuous threshold $s(\cdot)$ such that $s(\mathcal{Z}) \subset (0,1)$, if and only if*

(i). There is a continuous $q(\cdot)$, where $q(\mathcal{Z}) \subset (0,1)$, and an integer number $\overline{n} \geq 2$ such that $\mathbb{G}(Not\ sold|z) = q(z)^{\overline{n}}$ and $\mathbb{G}(\cdot|z) = \mathbb{G}(\cdot|q(z))$. The c.d.f.

$$G_c(b|q(z)) = \frac{\left[(1-q^{\overline{n}}(z))\,\mathbb{G}(W \leq b|Sold, z) + q^{\overline{n}}(z)\right]^{1/\overline{n}} - q(z)}{1-q(z)}$$

has a support $[\underline{v}, \overline{b}_c(z)]$ for some $\underline{v} < \overline{b}_c(\cdot) < \infty$.

(ii). $B_c(\cdot|z) = G_c^{-1}(\cdot|z)$ is continuously differentiable with $B_c^{(1)}(0|z) = 0$. The function

$$V_c(\alpha|q(z)) = B_c(\alpha|z) + \left(\alpha + \frac{q(z)}{1-q(z)}\right)\frac{B_c^{(1)}(\alpha|z)}{\overline{n}-1}$$

belongs to $\mathcal{V}_{q,\underline{v}}$.

Lemma 3.2 *Let the c.d.f.* $G_c(\cdot|z)$ *and* $G(b|z) = \mathbb{G}(W \leq b\,|Sold, Z = z)$ *be as in Proposition 3.6. Then* $G_R(\cdot|z)$ *and* $G(\cdot|b)$ *have p.d.f.* $g_R(\cdot|z)$ *and* $g(\cdot|z)$*, respectively, which are continuous over* $(\underline{v}, \overline{b}_R(\cdot)]$ *and diverge when* b *goes to* $\underline{v}$ *with*

$$0 < \lim_{b \downarrow \underline{v}} \left\{ (b - \underline{v})^{1/2} g_R(b|z) \right\} < \infty, \quad 0 < \lim_{b \downarrow \underline{v}} \left\{ (b - \underline{v})^{1/2} g(b|z) \right\} < \infty.$$

The proofs of these results are omitted as they involve arguments similar to those of Propositions 3.5, 3.2 and Lemma 3.1.

Identification results. As for reserve price, variation of the instrument is sufficient for identification of the entry model when entrants do not observe competition.

Proposition 3.7 *Consider a first-price auction model with entry cost satisfying Assumption E, where the buyers and analyst do not observe the number of entrants* N_c*. Suppose in addition that the maximal bid* $B_c(1|z)$ *varies with* $z \in \mathcal{Z}$*.*

Then the maximal number of entrants $\overline{n}$*, the threshold signal* $\{s(z), z \in \mathcal{Z}\}$*, the c.d.f* $\{F(\cdot|s), s \in s(\mathcal{Z})\}$*, the signal and cost function* $\{c(z), z \in \mathcal{Z}\}$ *are identified from the outcome distribution* $\{\mathbb{G}(\cdot|z), z \in \mathcal{Z}\}$*.*

Proof of Proposition 3.7: See Section 6.10 in the proof section.

The variation condition on $B_c(1|z)$ is testable. While the proof of Proposition 3.4 establishes that the maximal bid $B_R(1|z)$ under reserve price $R(z)$ is not constant, showing that $B_c(1|z)$ satisfies such a requirement seems out of reach in full generality.[7] Interestingly, the important role played by the upper bound of the bid distribution in Propositions 3.4 and 3.7 parallels the role of the conditional bid distribution's upper bounds for achieving identification when the buyers observe competition.

[7] This amounts to studying the variations of $B_c(1|z)$ with respect to $s(z)$, as in Li and Zheng (2009,2012). But these authors consider different entry models, where, in particular, at least two bidders enter.

3.3 Information and entry models identification

The preceding sections considered four auction models whether competition is restricted by a reserve price or an entry cost and whether or not active buyers observe the number of competitors. We now show that the winning bid distribution characterizations obtained above can be used to decide which of these models hold. To decide whether entry is restricted by a reserve price or an entry cost, it is convenient to rely on an instrument satisfying the following assumption.

Assumption ER. *Buyers are restricted either by a reserve price $R(\cdot)$ as in Assumptions IPV and R, or an entry cost $c(\cdot)$, as in Assumption E. The instrument $z \in \mathcal{Z}$ satisfies Assumption R with $R(\mathcal{Z}) \subset (\underline{v}, \overline{v})$ or Assumption E with $c(\mathcal{Z}) \subset (\Pi_{\overline{n}}(0,0), \Pi_{\overline{n}}(1,1))$ and $\underline{v} > 0$, respectively.*

The instrument z can stack instruments specific to reserve price and entry cost, i.e. $z = (z_R, z_c)$ with $R(z) = R(z_R)$ and $c(z) = c(z_c)$.[8]

Our first result deals with buyer information. Corollary 3.3-(i,ii) allows to identify whether buyers observe or not competition from the continuous component of the conditional outcome distribution $\mathbb{G}(\cdot|z)$, that is using auction data with at least two active buyers. Corollary 3.3 also holds in the absence of instrument.

Corollary 3.3 *Under Assumption ER, active buyers observe the number of participants if any of the three following conditions hold:*

(i). If $\overline{n} > 3$ and for each $z \in \mathcal{Z}$, the continuous component of $\mathbb{G}(\cdot|z)$ has a p.d.f. $g(\cdot|z)$ which has at least one discontinuity on its inner support.

[8] Assumption ER restricts to cases where reserve price or entry cost is binding, so that the two participation models have an empty intersection. Although it is feasible to assume a nonbinding reserve price or entry cost, we have not done so here for the sake of brevity.

(ii). For each $z \in \mathcal{Z}$, the p.d.f. $g(\cdot|z)$ of the winning bid distribution is bounded at its support lower bound $\underline{w}(z)$.

(iii). There is a $\varrho(z) \leq \underline{w}(z)$ such that $\mathbb{G}(W = \varrho(z)|z) > 0$ for each $z \in \mathcal{Z}$. If so, entry is restricted by a reserve price $\varrho(z)$.

Proof of Corollary 3.3: (i,ii) follows from Lemma 2.1-(ii) and the compatibility conditions in Propositions 3.1-(iii) and 3.6-(ii), which give that $g(\cdot|z)$ is continuous on its support, except possibly, at a finite number of discontinuity points when $\underline{n} > 3$, and when buyers observe competition, while Lemmas 3.1 and 3.2 give that $g(\cdot|z)$ diverges at its lower boundary. (iii) is a consequence of Propositions 3.1-(i), 3.2-(i), 3.5-(i) and 3.6-(i). □

The next corollary investigates whether entry is restricted by a reserve price or entry cost using instruments variations. In this result, $\underline{w}(z)$ stands for the lower support bound of the continuous component $G(\cdot|z)$ of the outcome distribution $\mathbb{G}(\cdot|z)$.

Corollary 3.4 *Suppose Assumption ER holds. Then $\{\mathbb{G}(\cdot|z), z \in \mathcal{Z}\}$ is rationalized by an entry cost if any of the following conditions hold:*

(i). $\underline{w}(\cdot)$ does not depend upon on the instrument, in which case $\underline{w}(\cdot) = \underline{v}$.

(ii). Either $\mathbb{G}(W = 0|z) > 0$ for all $z \in \mathcal{Z}$, or for $\underline{v} > 0$ independent of z as in (i), $g(\cdot|z)$ diverges at $\underline{v}$ with $0 < \lim_{b \downarrow \underline{v}} \left\{(b - \underline{v})^{1/2} g(b|z)\right\} < \infty$ for all $z \in \mathcal{Z}$.

Proof of Corollary 3.4: This follows from Propositions 3.1-(ii,iii), 3.1-(i), 3.5-(ii,iii), 3.5-(i), and Corollary 3.2, observing that, for a reserve price, $\underline{w}(z) = R(z) > 0$ must depend upon z by Assumption R, that $\mathbb{G}(W = 0|z) = 0$, and either $\mathbb{G}(W = R(z)|z) > 0$ or $g(\cdot|z)$ diverges at $R(z)$ by Corollary 3.1. □

Corollary 3.4 uses support variation to characterize entry cost models. An alternative characterization of reserve price and entry cost models is as follows. Liu et al. (2017) note that conditional outcome distribution $\mathbb{G}(\cdot|z)$ in entry games only depends upon the entry probability $\pi(z)$, i.e. $\mathbb{G}(\cdot|z) = \mathbb{G}(\cdot|\pi(z))$. As noted after Proposition 3.5, a similar sufficiency property holds under entry cost or reserve price. Using the identified entry updated private value c.d.f. and the identified probability of entry give functional forms that can be tested. Under a reserve price model, the updated private value c.d.f. is, for $v \geq R(z)$ and $\pi(z) = F(R(z))$,

$$F_R(v|z) = \frac{F(v) - \pi(z)}{1 - \pi(z)},$$

and $(1-\pi(z))F_R(v|z)+\pi(z)$ does not depend upon z. This contrasts with the updated private value distribution under an entry model, for which $\pi(z) = s(z)$ as

$$F_c(v|\pi(z)) = \frac{1}{1 - \pi(z)} \int_{\pi(z)}^{1} F(v|s)ds.$$

Hence observing $(1 - \pi(z)F_c(v|\pi(z)) + \pi(z)$ depending upon z is evidence against a reserve price model.

4 Applications to USFS timber auctions

This section illustrates applications of our identification results to USFS timber auctions. These auctions have been studied extensively in the literature. However, no existing work allows the actual competition differ from the observed number of bids. We study the first-price auction data used in Lu and Perrigne (2008). The data contain 107 two-bid auctions and 108 three-bid auctions, and report the appraisal value and timber volume of each auctioned lot. See Lu and Perrigne (2008) for further information on this dataset. Since the literature argues that the USFS reserve prices are too low, we also assume that they are nonbinding.

Our first exercise illustrates detection of pdf discontinuities and estimation of the distribution of competition N from winning bids. Hence N should be interpreted as the number of bidders contributing to the winning bid, or the number of competitive, or equivalently equilibrium bids, which may differ from the distribution of the number of bids if, for instance, some bids are dominated. A robustness analysis using only three-bid auctions confirms this finding. Our second exercise studies bidders' risk attitude as in Lu and Perrigne (2008). Treating competition as a latent variable, we provide a robust approach for bounding the CRRA parameter.

4.1 Competition analysis

Let W_ℓ be the winning bid in auction ℓ and $X_\ell = (X_{1\ell}, X_{2\ell})$ be the associated auctioned lot covariate, which includes the appraisal value and volume. Being agnostic about the competition, we consider only a winning bid sample computed from the two samples with two and three bids. The number of bidders is considered as unobserved heterogeneity and the sample size is $L = 215$. We first estimate the lowest number of bidders $\underline{n}$ and then focus on discontinuities in the conditional winning bid pdf and on the conditional distribution of N.

4.1.1 The lowest number of bidders

The lowest number of bidders $\underline{n}$ is assumed to be independent of the covariate X, but the presence of X complicates its identification and estimation. Without auction-specific covariates X, $G(b) = g_{\underline{n}}^{\underline{n}}(0)(b - \underline{v})^{\underline{n}}(1 + o(1))$ when b decreases to $\underline{v}$. We can estimate $\underline{n}$ as a lower tail index of the cdf $G(\cdot)$, see Lemma 2.1-(iii). In this case, the Hill estimator as in Hill and Shneyerov (2013) can be applied. With auction-specific covariates X, $\underline{n}$ can be identified as a conditional lower tail index when b decreases to the common conditional lower bound $\underline{v}(X)$

$$G(b|X) = g_{\underline{n}}^{\frac{n}{\underline{n}}}(0|X)\,(b - \underline{v}(X))^{\underline{n}}\,(1 + o(1)).$$

However, the standard Hill estimator may not be consistent except when $\underline{v}(\cdot)$ is

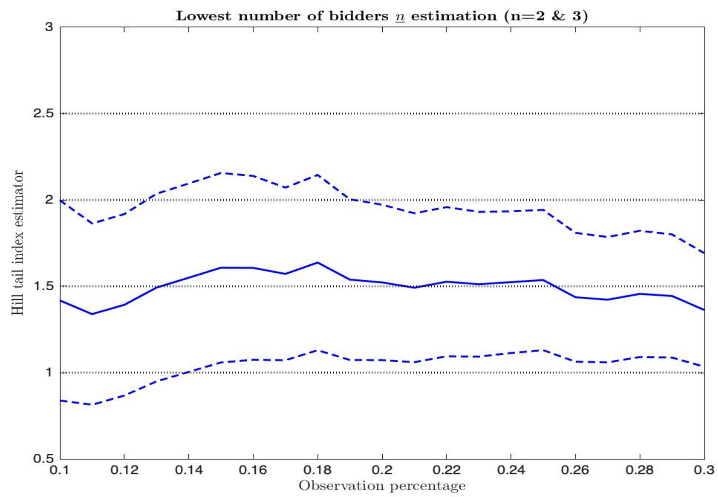


Figure 4: Hill estimators $\widetilde{\underline{n}}$: two-bid and three-bid auctions

constant. To address this issue, we first estimate the lower bound $\underline{v}(X)$ and use it to normalize the winning bids. Applying to the normalized sample of winning bid, the Hill estimator is consistent.

Note that auction theory predicts that the lower bound of value $\underline{v}(X)$ is the lower bound of the conditional bid distribution. Recall that auction-specific covariates consists in the appraisal value and volume, say $X_\ell = (X_{1\ell}, X_{2\ell})$. For a growing integer number $K = K_L$ and $k = 0, \ldots, K$, let $\hat{X}_1(k/K)$ $\hat{X}_{2\ell}(k/K)$ be the sample quantiles of order k/K of $X_{1\ell}$ and $X_{2\ell}$. Let $\mathcal{X}_{\ell K}$ be the set $(\hat{X}_1((k-1/K), \hat{X}_1(k/K)] \times (\hat{X}_2((k-1/K), \hat{X}_2(k/K)]$ to which X_ℓ belongs and set[9]

$$\widehat{\underline{v}}_\ell = \min\left(\min\left\{B_{ij}, i = 1, 2 \text{ and } X_j \text{ in } \mathcal{X}_{\ell K}\right\}, \min\left\{B_{ij}, i = 1, 2, 3 \text{ and } X_j \text{ in } \mathcal{X}_{\ell K}\right\}\right)$$

[9]Note that the lower boundary estimator $\widehat{\underline{v}}_\ell$ depends upon the bids and not the winning bids, which is permitted because the aim here is to analyze competition and not to recover primitives

We use these lower bound estimates to normalize the winning bids as

$$W_\ell^\dagger = \frac{W_\ell/\widehat{\underline{v}}_\ell}{\min_{1\leq\ell\leq L}\left(W_\ell/\widehat{\underline{v}}_\ell\right)} - 1.$$

The proposed estimator of the lowest number of bidders $\widehat{\underline{n}}$ is a rounding of the tail index Hill $\widetilde{\underline{n}}$ estimator of Hill and Shneyerov (2013) using $W_\ell^\dagger$

$$\frac{1}{\widetilde{\underline{n}}} = \ln W_{(M)}^\dagger - \frac{1}{M-1}\sum_{m=2}^{M} W_{(m)}^\dagger, \quad M = M_L \geq 2 \text{ with } M_L = o(L). \tag{28}$$

The lowest number of bidders estimator $\widehat{\underline{n}}$ is equal to the integer number k whenever $k - .5 < \widetilde{\underline{n}} \leq k + .5$. Figure 4 reports Hill estimation values $\widetilde{\underline{n}}$ when M ranges from 21 to 65 and $K = 2$ in solid line, with 95% confidence interval bounds in dashed line. The most common estimate of the lowest number of bidders $\widehat{\underline{n}}$ is $\widehat{\underline{n}} = 2$, with $\widehat{\underline{n}} = 1$ for only small M yielding noisy $\widehat{\underline{n}}$, or large M which gives biased estimates. Hence it seems very likely that $\underline{n} = 2$ for this sample.

4.1.2 Discontinuities and competition distribution

Conditioning variables. The presence of conditioning variables also affects detection discontinuities. In principle, the conditional winning pdf $g(\cdot|X)$ can be estimated using observations with covariate X_ℓ close to X. In view of our small sample size, we adopt a data-driven approach and consider three subsamples labelled "Low", "Medium" and "High" defined as follows:

- Low: auctions with appraisal value and volume both smaller than their median values (45 auctions, among which 23 have two bidders);
- Medium: auctions with appraisal value and volume both between their 25% and 75% quantiles (53 auctions, among which 28 have two bidders);

---

from winning bids. Using the winning bids to estimate $\widehat{\underline{v}}_\ell$ gives normalized winning bids which are too concentrated. Alternative Hill estimation procedures for conditional tail index, which does not use such normalization, can be found in Gardes and Stuffler (2014) and the references therein.

- High: auctions with covariates above their median values (44 auctions, among which 20 have two bidders).

Although these subsamples are small, detecting discontinuities and their locations can be done with super efficient rates, see Oudshoorn (1998) and Gayraud (2002).

Detection of discontinuities. We propose in the appendix a multiple testing approach, inspired by Chu and Cheng (1996) and Oudshoorn (1998), for detecting discontinuities, locate them and compute a corresponding discontinuous conditional pdf estimator. To cope with an irregular distribution of the winning bids across the straight line, we adopt a k-nearest neighbor (k-NN) approach in place of Chu and Cheng (1996) kernel approach. The idea is to conclude a point of discontinuity when the difference between the densities on the two sides of this point is larger than a threshold. Naturally, this threshold depends on the magnitude of the density function in the neighborhood.

Implementing the discontinuity detection algorithm gives a unique discontinuity inside the subsample support for each subsample, so that $\widehat{\overline{n}} = 3$ in all cases. Figure 5 reports the discontinuous conditional density estimators obtained for each subsample. In each case, the estimated discontinuity $\widehat{\Delta}_2$, in the middle of the subsamples, is quite large compared to the end discontinuity $\widehat{\Delta}_3$. Table 1 gives, for each subsample, the estimation of the probability that $N = 2$ derived from the discontinuities in the winning bid pdf using the conditional version of (12) (line "From W") and using the number of bids observed in the sample (line "Observed"). The value inferred from the estimated winning bid distribution is much higher than the one obtained from the bid distribution, consistently across subsample. This suggests that, for most auctions with three reported bids, only two bidders contribute to the winning bid while the remaining one makes a dominated bid.

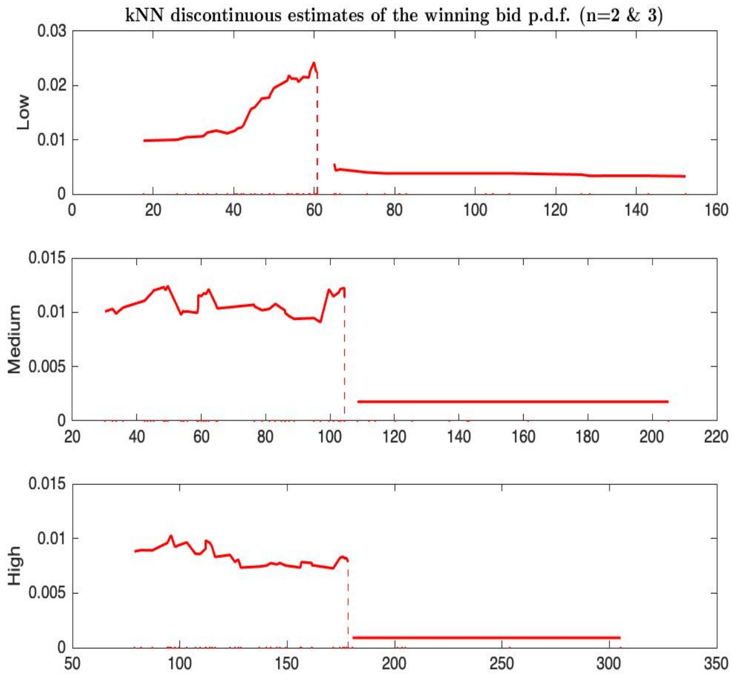


Figure 5: Conditional density estimation for 'Low', 'Medium' and 'High' subsamples

4.1.3 Analysis using the subsample of three-bid auctions

We provide further evidences on the lack of competition by repeating the above estimation procedure on the sample of three-bid auctions. First, we estimate the lowest number of bidders. Figure 6 displays the results of the Hill estimation procedure for this sample. Despite the maximum bids W_ℓ are now computed using only three-bid auctions, Figure 6 supports an estimate $\widehat{\underline{n}} = 2$ for the lowest number of bidders. This is a first evidence suggesting that one bidder may not contribute to the winning bid in this dataset. Second, we estimate the probabilities of $N = 2$ from pdf discontinuities obtained as in the previous section for each subsample 'Low', 'Medium' and 'High'. Table 2 reports the estimates. The results are qualitatively similar to the ones in Table 1, except for the 'Medium' sample for which the estimated probability

	Low	Medium	High
From W	.95	.90	.92
Observed	.51	.53	.45

Table 1: Estimated $\mathbb{P}(N = 2|A)$ using both two-bid and three-bid auctions

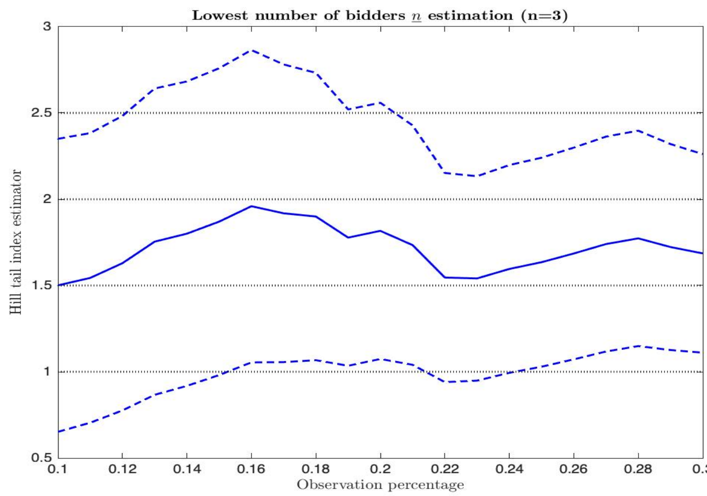


Figure 6: Hill estimators $\widetilde{\underline{n}}$: three-bid auctions

of having two bidders contributing to the winning bid is lower but still higher than 0.5. However, the variance of these estimations are likely to be very high due to their nonparametric nature and small sample sizes. Therefore, the estimation $\widehat{\underline{n}} = 2$ for the lowest number of bidders derived from Figure 6 may be a stronger evidence. This finding can be explained by the presence of a passive or dominated bidder in many auctions. More advanced explanations are however outside the scope of this paper.

	Low	Medium	High
From W	.98	.75	.91

Table 2: Estimated $\mathbb{P}(N = 2|A)$ using three-bid auctions

4.2 Bounds for risk aversion

Bidders' risk aversion behaviors in auctions have attracted much attention in the literature. A commonly adopted specification is the Constant Relative Risk Aversion (CRRA) $U(x) = x^{\theta}$. While Campo et al. (2011) showed that θ is not identified when participation is endogenous, Guerre et al. (2009) obtained risk-aversion identification when the exogenous participation is observed.[10] Using USFS timber auction data with both first-price and ascending auctions, Lu and Perrigne (2008) obtained estimates of θ in the range of 0.4. Using semiparametric restrictions for the bid distribution, Campo et al. (2011) obtained estimates of θ between 0.65 and 0.8 from a sample of USFS first-price auctions. All these works assume that all bids are equilibrium outcomes, whose failure leads to biased estimates of risk aversion parameters. Allowing for auction unobserved heterogeneity may however lead to conclude in favor of risk-neutral bidders, as in Grundl and Zhu (2019) who also consider USFS first-price auctions.

Hereafter, we assume the competition level N is exogenous and unobservable to the analyst. Following the same lines as in our benchmark case, we derive lower bounds for the CRRA parameter θ using the winning bid distribution. Suppose that

[10]Relaxing exogeneity is possible with entry models as seen from Gentry, Li and Lu (2017). These authors assume that all bids are equilibrium outcomes, which, as argued below, is an assumption that can be questioned for the considered samples. Indeed, under CRRA, the above estimation $\widehat{\underline{n}} = 2$ in the sample with three bids auctions remains valid, suggesting that one bid is not competitive in many auctions.

$N = n$ and observe that the expected utility generated by a bid $B_n(a)$ for a private value $V(\alpha)$ is

$$(V(\alpha) - B_n(a))^{\theta} a^{n-1} = \left[(V(\alpha) - B_n(a))\, a^{\frac{n-1}{\theta}} \right]^{\theta}$$

when the bidder's utility function is CRRA. The latter expression shows that, for equilibrium bids, α must maximize $a \mapsto (V(\alpha) - B_n(a))\, a^{\frac{n-1}{\theta}}$. Therefore, introducing risk-aversion here amounts to change $n-1$, i.e., the number of opponents of the risk-neutral case, into $(n-1)/\theta$. Arguing as for (3), it follows that the equilibrium bid quantile function now becomes

$$B_n(\alpha) = \frac{(n-1)/\theta}{\alpha^{\frac{n-1}{\theta}}} \int_0^{\alpha} t^{\frac{n-1}{\theta}-1} V(t)dt = V(\alpha) - \int_0^{\alpha} \left(\frac{t}{\alpha}\right)^{\frac{n-1}{\theta}} V^{(1)}(t)dt.$$

The last expression of $B_n(\cdot)$ shows that the maximum bid $B_n(1)$ increases with n. As (2) becomes $V(\alpha) = B_n(\alpha) + \theta\alpha B_n^{(1)}(\alpha)/(n-1)$, the conditional bid pdf $g_n(\cdot)$ is continuous and bounded away from 0. Arguing as for Corollary 2.1 yields that $g_n(\underline{v}) > 0$, implying that our Hill estimator is still consistent under risk-aversion. Hence the estimate $\widehat{\underline{n}} = 2$ obtained in the last subsection still applies under risk-aversion. In other words, one shall not treat all bids as equilibrium outcomes. Similarly, discontinuities of the winning bid identifies the largest number $\overline{n}$ of bidders, which amounts to 3 under risk-neutrality.

We now show how to derive bounds for the CRRA parameter θ, despite being agnostic about N. Accounting for risk-aversion in formula (7) of Corollary 2.1 and in the jump formula (9) yields that

$$g_n\left(\overline{b}_n\right) = \frac{\theta}{(n-1)\left(\overline{v} - \overline{b}_n\right)} \text{ and } \Delta_n = \frac{n\theta p_n}{(n-1)(\overline{v} - \overline{b}_n)} \, .$$

Updating the formulas (11) and (12) for $\overline{v}$ and p_n gives

$$\overline{v}(\theta) = \frac{\theta + \sum_{n=\underline{n}}^{\overline{n}} \frac{n-1}{n}\Delta_n \overline{b}_n}{\sum_{\underline{n}}^{n=\overline{n}} \frac{n-1}{n}\Delta_n},$$

$$p_n(\theta) = \frac{\frac{n-1}{n}\Delta_n}{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k} + \frac{1}{\theta}\frac{n-1}{n}\Delta_n \left(\frac{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k \overline{b}_k}{\sum_{k=\underline{n}}^{\overline{n}} \frac{k-1}{k}\Delta_k} - \overline{b}_n \right), \quad n = \underline{n}, \ldots, \overline{n}.$$

Using that $\overline{v}(\theta) \leq \overline{b}_n$ and $0 \leq p_n(\theta) \leq 1$ for $n = \underline{n}, \ldots, \overline{n}$ gives bounds for θ that can be estimated. For the estimated values of $\underline{n}$, $\overline{n}$, $\overline{b}_n$ and Δ_n, only the constraint $p_2(\theta) \leq 1$ is informative. This gives, for each considered subsamples, estimated lower bounds for θ which are reported in the next table.

	Low	Medium	High
Two and three bids	.78	.48	.44
Three bids	.90	.25	.65

Table 3: Lower bound for the CRRA parameter θ inferred from $\widehat{p}_2(\theta|A) \leq 1$

Assuming that the risk aversion parameter is constant across subsamples, we obtain the best lower bound for the CRRA coefficient θ 0.9, which is quite high relative to the estimates reported in the literature, such as Lu and Perrigne (2008) and Campo et al. (2011). This suggests that accounting for the possibility that some bids are not necessarily equilibrium outcomes may affect risk-aversion estimation.[11] Risk-neutrality becomes more plausible when assuming that N is not observed, ie

[11]Note also that the CRRA lower bounds of Table 3 could be, in theory, improved. Indeed, an estimated upper bound $\widehat{\overline{p}}_2(A)$ can be obtained by dividing the number of winning bids below the estimated discontinuity $\widehat{\overline{b}}_2$, as reported in Figure 5 by the considered subsample size and by using the inequality $\widehat{p}_2(\theta|A) \leq \widehat{\overline{p}}_2(A)$ instead of $\widehat{p}_2(\theta|A) \leq 1$ to obtain a lower bound for θ. This would lead to replace the lower bound 0.25 by 0.47, but, when providing a meaningful bound, would give a risk-aversion larger than 1 in many other cases. Such finding is not so surprising in view of the small sample sizes. Guerre and Gimenes (2019) also report that risk-aversion can be difficult to estimate.

that competition is a source of unobserved heterogeneity. This parallels Grundl and Zhu (2019), who however consider auction unobserved heterogeneity associated with the auctioned good.

5 Final remarks

This paper shows that, under the independent symmetric private value paradigm, the first-price auction winning bid is sufficient to identify model primitives when buyers observe competition. The case where buyers do not observe competition is more difficult, but still feasible with an instrument when participation is restricted by a reserve price or entry cost. Whether entry is constrained by a reserve price imposed by the seller, or by an entry cost can be tested.

An empirical application on USFS timber first-price auction winning bids has illustrated the usefulness of our theoretical results for competition analysis. A discontinuity detection procedure has found discontinuities in the winning bid pdf, suggesting that bidders observe competition according to our theoretical findings. Investigating further the distribution of the number of competitive bidders reveals that most auctions with three bids include a dominated bidder, who does not contribute to the winning bid. Relaxing risk-neutrality by assuming a Constant Relative Risk Aversion (CRRA) utility function leads to a similar conclusion. Deriving bounds for the CRRA parameter from the winning bid illustrates the impact of ignoring dominated bidding on risk-aversion estimation: risk-neutrality looks much more plausible using our approach which tackles with this issue than found in previous studies.

Our theoretical findings could be the source of many further developments. Discriminating the various possible explanations for dominated bidding, such as a poor understanding of the auction rules or collusion among others, would be very interesting. Econometric methods for modeling the participation decision based upon the winning bid can also be useful for data that can be analyzed through the auction machinery. The statistical methods used in this paper for competition analysis can probably be refined. Estimating the private value distribution from winning bids necessitates to extend estimation methods for irregular models as proposed in Hirano and Porter (2003) or Chernozhukov and Hong (2004) to allow for many density discontinuities.

Further, our theoretical results shed light on new identification arguments for discrete mixture models, which are widely used in economic applications, in particular when unobserved heterogeneity is plausible. In our model, the mixture components are generated by the same function. The components are ordered according first-order stochastic dominance and their supports are nested. These two features may appear in other relevant economic mixtures. See for instance An (2017) who studied non equilibrium bids from heterogeneous agents whose beliefs follow from level k thinking, and where k is unobserved. A key ingredient is that these support components can be identified, here through discontinuities of the mixture pdf but many other characteristics can be used for such a purpose.

6 Proof section

6.1 Proof of Lemma 2.3

Consider (i) first. Since $\alpha_k = G_{\overline{n}}\left[B_{\overline{n}-1}\left(\alpha_{k-1}\right)\right]$ with $B_{\overline{n}-1}\left(\alpha\right) \leq B_{\overline{n}}\left(\alpha\right)$,

$$\alpha_k = G_{\overline{n}}\left[B_{\overline{n}-1}\left(\alpha_{k-1}\right)\right] \leq G_{\overline{n}}\left[B_{\overline{n}}\left(\alpha_{k-1}\right)\right] = \alpha_{k-1},$$

which implies that α_k decreases. Moreover, $\beta_k = B_{\overline{n}-1}\left(\alpha_k\right)$ decreases because $B_{\overline{n}-1}\left(\cdot\right)$ is strictly increasing. Since $\alpha_k \geq 0$, α_k converges to a limit α which satisfies $\alpha = G_{\overline{n}}\left[B_{\overline{n}-1}\left(\alpha\right)\right]$ under Assumption IPV. In other words, the limit α satisfies $B_{\overline{n}}\left(\alpha\right) = B_{\overline{n}-1}\left(\alpha\right)$. This gives $\alpha = 0$ as $B_{\overline{n}}\left(\alpha\right) > B_{\overline{n}-1}\left(\alpha\right)$ except for $\alpha = 0$.

Now, consider (ii). That α_k is identified for all k will follow from an induction argument, observing α_1 is identified. Suppose then that α_k and $\{V(\alpha), \alpha \in [\alpha_k, 1]\}$ are identified. Recall

$$\alpha_{k+1} = G_{\overline{n}}(\beta_k) = G_{\overline{n}}[B_{\overline{n}-1}(\alpha_k)], \quad \beta_{k+1} = B_{\overline{n}-1}(\alpha_{k+1}).$$

Then (4) and Lemma 2.2 give that $\{B_n(\alpha); \alpha \in [\alpha_k, 1]\}$, for all $n = \underline{n}, \ldots, \overline{n}-1$ are identified, as β_k. Now (13) and Lemma 2.2 show that $G_{\overline{n}}(b)$ is identified for all $b \geq \beta_k$, and then $\alpha_{k+1} = G_{\overline{n}}(\beta_k)$ is identified. (2) then gives that $\{V(\alpha); \alpha \in [\alpha_{k+1}, 1]\}$ is identified. This ends the proof of the Lemma. □

6.2 Proof of Proposition 3.1

Suppose that the transaction price is generated from a first-price auction with a reserve price R and N_R known by the $\overline{n}$ buyers and such that Assumption IPV holds. Then q in (i) satisfies $q = F(R)$, R is the unique real number such that $\mathbb{G}(W = R) > 0$, and N_R has a Binomial distribution with parameter $(\overline{n}, F(R))$. This gives

$$\begin{aligned} \mathbb{G}(\text{Not sold}) &= \mathbb{P}(N_R = 0) = q^{\overline{n}}, \\ \mathbb{G}(\text{Object sold at price } R) &= \mathbb{P}(N_R = 1) = \overline{n} q^{\overline{n}-1}(1-q). \end{aligned}$$

Now, for $G_n(\cdot) = B_{n,R}^{-1}(\cdot)$, it holds for $W = \max_{i \in \mathcal{N}} B_i$ that

$$\begin{aligned} G(b) &= \mathbb{G}(W \leq b \,|\, N_R \geq 2) \\ &= \sum_{n=2}^{\overline{n}} \mathbb{G}\left(\max_{i=1,\ldots,n} B_i \leq b \,|\, N_R = n\right) \mathbb{P}(N_R = n \,|\, N_R \geq 2) \\ &= \sum_{n=2}^{\overline{n}} G_n^n(b) \frac{\overline{n}!}{n!(\overline{n}-n)!} \frac{F^{\overline{n}-n}(R)\,[1-F(R)]^n}{1 - F^{\overline{n}}(R) - \overline{n} F^{\overline{n}-1}(R)\,[1-F(R)]}, \end{aligned}$$

which is (ii). $G_n(\cdot) = B_{n,R}^{-1}(\cdot)$ and (15) gives (iii).

For the reverse implication, observe that (i)-(iii) identify a number of potential buyers $\overline{n}$, a screening level $F(R)$, a reserve price R which is uniquely defined when $0 < F(R) < 1$, and a private value quantile function $V(\alpha)$, where $\alpha \geq F(R)$, which can be extended to $[0,1]$ so that Assumption IPV holds. It can be easily seen from (16) that the distribution of the outcome distribution of this first-price auctions is $\mathbb{G}(\cdot)$.

6.3 Proof of Corollary 3.1

(i) follows from Theorem 2.1, which gives identification of $V_R(\cdot)$ and $\overline{n}$. For (ii), observe that the expression of $G(\cdot)$ in Proposition 3.1-(ii) holds with $q = F(R)$. Since $G(\cdot)$ is the winning bid distribution of a first-price auction with *i.i.d.* private values drawn from $F_R(\cdot)$ and independent from the number of participants N_R, Lemma 2.2 shows that

$$p_{\overline{n}}(R) = \frac{(1 - F(R))^{\overline{n}}}{1 - F(R)^{\overline{n}} - \overline{n}F(R)^{\overline{n}}(1 - F(R))},$$
$$p_{\overline{n}-1}(R) = \frac{\overline{n}(1 - F(R))^{\overline{n}-1}F(R)}{1 - F(R)^{\overline{n}} - \overline{n}F(R)^{\overline{n}}(1 - F(R))},$$

are identified if $\overline{n} \geq 3$. Thus the ratio $\overline{n}p_{\overline{n}}(R)/p_{\overline{n}-1}(R)$ is identified, and thereby $F(R)$, since

$$F(R) = \frac{1}{1 + \overline{n}p_{\overline{n}}(R)/p_{\overline{n}-1}(R)}.$$

This together with the expression of $F_R(\cdot)$ shows that $F(v)$ is identified for $v \geq R$. For (iii), $\mathbb{G}(\text{Not sold}) = F^2(R)$ identifies $F(R)$ and then $F(v)$ is also identified for $v \geq R$ by (i). When $\overline{n} = 2$, $F(R)$ cannot be recovered from $G(\cdot)$. Since $\underline{v} < R$, it can be easily seen that a given $F_R(\cdot)$ can be generated by two distinct $\{F(v), v \geq R\}$, which generates the same $G(\cdot)$. Hence $\{F(v), v \geq R\}$ cannot be identified when $\overline{n} = 2$. $\square$

6.4 Proof of Proposition 3.2

Suppose that $\mathbb{G}(\cdot)$ has been generated by a first-price auction satisfying Assumption IPV, with a reserve price R but with the number of participants N_R not observed by buyers. The probability that $N_R = 0$ is $F(R)^{\overline{n}} \in (0,1)$ if $\underline{v} < R < \overline{v}$, which is also equal to $\mathbb{G}$ (Not sold). Since the $B_i \geq R$ are given by $B_R(A_{i,R})$, W has a continuous distribution with support $\left[R, \overline{b}_R\right]$, $\overline{b}_R = B_R(1)$. Since all $A_{i,R}$ and N_R are independent and because N_R has a binomial distribution with parameter $(\overline{n}, 1 - F(R))$, the c.d.f. $G(b) = \mathbb{G}(W \leq b\,|\text{Sold})$ of W is given by

$$
\begin{aligned}
G(b) &= \frac{\sum_{n=1}^{\overline{n}} \mathbb{P}\left(\max_{i=1,\ldots,N_R} B_R(A_{i,R}) \leq b | N_R = n\right) \mathbb{P}(N_R = n)}{1 - \mathbb{P}(N_R = 0)} \\
&= \sum_{n=1}^{\overline{n}} \frac{\frac{\overline{n}!}{n!(\overline{n}-n)!} (1 - F(R))^n F(R)^{\overline{n}-n}}{1 - F(R)^{\overline{n}}} G_R^n(b) \\
&= \frac{(F(R) + (1 - F(R)) G_R(b))^{\overline{n}} - F(R)^{\overline{n}}}{1 - F(R)^{\overline{n}}},
\end{aligned}
$$

where $G_R(\cdot) = B_R^{-1}(\cdot)$ has support $\left[R, \overline{b}_{\overline{n}}\right]$. Setting $q = F(R)$ gives (i) and (ii) follows from (18) and (19).

For the converse, observe that the reserve price is identified as $B_R(0)$, set $F(R) = q$ gives

$$
V_R(\alpha) = B_R(\alpha) + \left(\alpha + \frac{q}{1-q}\right) \frac{B_R^{(1)}(\alpha)}{\overline{n} - 1},
$$

$F_R(\cdot) = V_R^{-1}(\cdot)$ and $F(v) = (1 - F(R))F_R(v) + F_R(v)$ for $v \geq R$. Choose $\underline{v} < R$ and note that $F(\cdot)$ can be extended to $[\underline{v}, R]$ to obtain a continuously differentiable c.d.f. with bounded support $[\underline{v}, \overline{v}]$. It can be easily seen that these primitives generate first-price auction best response bids compatible with the outcome distribution $\mathbb{G}(\cdot)$.□

6.5 Proof of Lemma 3.1

It holds that

$$g_R(b) = \frac{1}{B_R^{(1)}\left[G_R(b)\right]} \quad \text{for all } b \text{ in } (R, \overline{b}_R)$$

which diverges when b goes to R as $B_R^{(1)}(0) = 0$ and $G_R(R) = 0$. Let $V_R(\cdot)$ be the continuously differentiable quantile function of Proposition 3.2:

$$V_R(\alpha) = B_R(\alpha) + \left(\alpha + \frac{q}{1-q}\right) \frac{B_R^{(1)}(\alpha)}{\overline{n}-1}.$$

Hence, for $\alpha > 0$,

$$\begin{aligned} B_R^{(2)}(\alpha) &= (\overline{n}-1)\left(\frac{V_R^{(1)}(\alpha) - B_R^{(1)}(\alpha)}{\alpha + \frac{q}{1-q}} - \frac{V_R(\alpha) - B_R(\alpha)}{\left(\alpha + \frac{q}{1-q}\right)^2}\right) \\ &= (\overline{n}-1)\frac{1-q}{q} V_R^{(1)}(0) + o(1) \text{ when } \alpha \to 0. \end{aligned}$$

It follows that, when α goes to 0 and b to R

$$\begin{aligned} B_R^{(1)}(\alpha) &= (\overline{n}-1)\frac{(1-q)^2}{q} V^{(1)}(R)\,\alpha + o(\alpha), \\ B_R(\alpha) &= R + \frac{\overline{n}-1}{2}\frac{(1-q)^2}{q} V^{(1)}(R)\,\alpha^2 + o\left(\alpha^2\right), \\ G_R(b) &= \left(\frac{2}{\overline{n}-1}\frac{q}{(1-q)^2 V^{(1)}(R)}\right)^{1/2} (b-R)^{1/2} + o\left((b-R)^{1/2}\right), \\ g_R(b) &= \left(\frac{1}{2(\overline{n}-1)}\frac{q}{(1-q)^2 V^{(1)}(R)}\right)^{1/2} (b-R)^{-1/2} + o\left((b-R)^{-1/2}\right), \end{aligned}$$

which shows $0 < \lim_{b \downarrow R} (b-R)^{1/2} g_R(b) < \infty$. Now, for b in $(R, \overline{b}_R]$,

$$\begin{aligned} G(b) &= \frac{\left[q + (1-q)\,G_R(b)\right]^{\overline{n}} - q^{\overline{n}}}{1-q^{\overline{n}}}, \\ g(b) &= \frac{\overline{n}(1-q)\left[q + (1-q)\,G_R(b)\right]^{\overline{n}} g_R(b)}{1-q^{\overline{n}}}, \end{aligned}$$

which gives the divergence result for $g(\cdot)$. $\square$

6.6 Proof of Proposition 3.3

For (i), note that the lower bound of the support of $G(\cdot)$ is R. For (ii), consider $\overline{n} = 2$, reserve price $R = 1/2$ and private values uniform over $[0, 1]$. The next computation shows that the outcome distribution can also be rationalized with $\overline{n} = 4$, $R = 1/2$, and a private value distribution distinct from uniform.

We first compute the distribution $\mathbb{G}(\cdot)$ generated by $\overline{n} = 2$, $R = 1/2$, private values uniform over $[0, 1]$ and a reserve price $R = 1/2$, so that $F(R) = 1/2$ and $F^{\overline{n}}(R) = 1/4$. The private value quantile function is $V(\tau) = \tau$, $V_R(\alpha) = (1 + \alpha)/2$ and the best-response bid quantile function is

$$B_{R,2}(\alpha) = \frac{1+\alpha}{4} + \frac{1}{4(1+\alpha)} = \frac{(1+\alpha)^2 + 1}{4(1+\alpha)}, \quad \alpha \in [0, 1].$$

The corresponding continuous component $G(b) = \mathbb{G}(W \leq b|\text{Sold})$ of $\mathbb{G}(\cdot)$ is

$$G(b) = \frac{\left(\frac{1}{2} + \frac{1}{2}B_{R,2}^{-1}(b)\right)^2 - \frac{1}{4}}{1 - \frac{1}{4}},$$

setting $B_{2,R}^{-1}(b) = 0$ for $b \leq 1/2$ and $B_2^{-1}(b) = 1$ for $b \geq 5/8$. The outcome distribution is

$$\mathbb{G}(\text{not sold}) = \frac{1}{4}, \quad \mathbb{G}(W \leq b|\,\text{sold}) = G(b).$$

If $\mathbb{G}(\cdot)$ can be rationalized by first-price auctions with reserve price $1/2$, $\overline{n} = 4$ and $F(R)^{\overline{n}} = 1/4$, which gives $F(1/2) = 1/\sqrt{2}$, there must exist a bid quantile function $B_{4,R}(\cdot)$ satisfying $B_{4,R}(0) = 1/2$, such that

$$G(b) = \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right) B_{4,R}^{-1}(b)\right)^4 - \frac{1}{4}}{1 - \frac{1}{4}}$$

and a private value quantile function as in Proposition 3.2-(ii). Solving

$$\frac{\left(\frac{1}{2} + \frac{1}{2}B_{2,R}^{-1}(b)\right)^2 - \frac{1}{4}}{1 - \frac{1}{4}} = \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right) B_{4,R}^{-1}(b)\right)^4 - \frac{1}{4}}{1 - \frac{1}{4}}$$

shows that it must hold that

$$B_{4,R}(\alpha) = B_{2,R}\left[2\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^2 - 1\right]$$
$$= \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^2}{2} + \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^{-2}}{8},$$

which is such that $B_{4,R}(0) = R = 1/2$. This also gives

$$B^{(1)}_{4,R}(\alpha) = \left(1 - \frac{1}{\sqrt{2}}\right) \times \left\{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right) - \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^{-3}}{4}\right\},$$

which satisfies $B^{(1)}_{4,R}(0) = 0$. To check that $B_{4,R}(\cdot)$ is compatible with a private value quantile function, it remains to be checked that

$$v(\alpha) = B_{4,R}(\alpha) + \frac{\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha}{1 - \frac{1}{\sqrt{2}}}\frac{B^{(1)}_{4,R}(\alpha)}{3}$$
$$= \frac{5}{6}\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^2 + \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^{-2}}{24}$$

increases with α. Note that

$$\frac{dv(\alpha)}{d\alpha} = 2\left(1 - \frac{1}{\sqrt{2}}\right) \times \left\{\frac{5}{6}\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right) - \frac{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^{-3}}{24}\right\}$$
$$= \frac{2\left(1 - \frac{1}{\sqrt{2}}\right)}{\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^{-3}}\left\{\frac{5}{6}\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^4 - \frac{1}{24}\right\}$$

with, for all $\alpha \in [0,1]$,

$$\frac{5}{6}\left(\frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right)\alpha\right)^4 - \frac{1}{24} \geq \frac{5}{6} \times \frac{1}{4} - \frac{1}{24} = \frac{1}{6}.$$

It follows that $v^{(1)}(\cdot) > 0$. Recall $v(0) = 1/2$. Consider a c.d.f. $F(\cdot)$ over $[0,1]$ such that, for $\alpha \in \left[1/\sqrt{2}, 1\right]$,

$$F^{-1}(\alpha) = v\left(\frac{\alpha - \frac{1}{\sqrt{2}}}{1 - \frac{1}{\sqrt{2}}}\right), \quad \text{so that } F\left(\frac{1}{2}\right) = \frac{1}{\sqrt{2}}.$$

A proper definition of $F(\cdot)$ on $[0,1/2]$ ensures that $F(\cdot)$ satisfies Assumption IPV. By Proposition 3.2, $\mathbb{G}(\cdot)$ is rationalized by a first price auction with a reserve price $1/2$ and: 1) uniform private values with $\overline{n}=2$; 2) non-uniform private values with c.d.f. $F(\cdot)$, $\overline{n}=4$, and $F(R)=1/\sqrt{2}$. This establishes the non-identifiability statement in (ii).

For (iii), note that $0<F(R)<1$. It is sufficient to show that

$$\mathbb{P}(N_R=0)=F(R)^{\overline{n}} \text{ and } \mathbb{P}(N_R=1)=\overline{n}F(R)^{\overline{n}-1}(1-F(R))$$

identify $\overline{n}$ and $F(R)$, as $G(\cdot)$ will in turn identify $G_R(\cdot)$ and $F(v)$ for all $v\geq R$. Since

$$\overline{n}=\frac{\log\mathbb{P}(N_R=0)}{\log F(R)} \text{ and } \frac{\mathbb{P}(N_R=1)}{\mathbb{P}(N_R=0)}=\overline{n}\frac{1-F(R)}{F(R)},$$

it follows that $F(R)\in(0,1)$ solves

$$\frac{1-F(R)}{F(R)\log F(R)}=\frac{\mathbb{P}(N_R=1)}{\mathbb{P}(N_R=0)\log\mathbb{P}(N_R=0)};$$

in other words, $\varphi(F(R))$ is identified, where $\varphi(x)=(1-x)/(x\log x)$. Since,

$$\frac{\partial\varphi(x)}{\partial x}=-\frac{1}{x\log x}-\frac{(1-x)(1+\log x)}{(x\log x)^2}=-\frac{\log x-(x-1)}{(x\log x)^2}\geq 0,$$

where the inequality is strict for all x in $[0,1]$[12] except for $x=1$, $\varphi(\cdot)$ is a one-to-one mapping over, $(0,1)$ and $F(R)$ is identified so that $\overline{n}=\frac{\log\mathbb{P}(N_R=0)}{\log F(R)}$ is identified. □

6.7 Proof of Proposition 3.4

As $R(z)$ is the lower bound of $G(\cdot|z)$, $R(\cdot)$ is identified. Since

$$\mathbb{G}(\text{not sold}|z)=F(R(z)),^{\overline{n}}$$

[12]Using $\frac{\partial}{\partial x}[\log x-(x-1)]=\frac{1}{x}-1\geq 0$ gives $\log x-(1-x)\leq\log 1=0$ over $[0,1]$.

$F\left(R\left(\cdot\right)\right)^{\overline{n}}$ is identified. The upper bound $\overline{b}_R\left(\cdot\right) = B_R(1|z)$ of $G\left(\cdot|\cdot\right)$ is identified and, for all b and all $z \in \mathcal{Z}$,

$$\left(F\left(R\left(z\right)\right) + \left(1 - F\left(R\left(z\right)\right)\right) G_R\left(b|z\right)\right)^{\overline{n}}$$

is also identified. Since, when $t > 0$ goes to 0,

$$\begin{aligned}
&\left(F\left(R\left(z\right)\right) + \left(1 - F\left(R\left(z\right)\right)\right) G_R\left(\overline{b}_R\left(z\right) - t|z\right)\right)^{\overline{n}} \\
&= \left[1 + \left(1 - F\left(R\left(z\right)\right)\right)\left\{G_R\left(\overline{b}_R\left(z\right) - t|z\right) - 1\right\}\right]^{\overline{n}} \\
&= 1 - \overline{n}\left(1 - F\left(R\left(z\right)\right)\right) g_R\left(\overline{b}_R\left(z\right)|z\right) t + o\left(t\right),
\end{aligned}$$

the function $\gamma\left(z\right) = \overline{n}_R\left(1 - F\left(R\left(z\right)\right)\right) g_R\left(\overline{b}_R\left(z\right)|z\right)$ is identified over $\mathcal{Z}$. Since $g_R\left(\overline{b}_R\left(z\right)|z\right) = 1/B_R^{(1)}\left(1|z\right)$, taking $\alpha = 1$ in (18) gives

$$g_R\left(\overline{b}_R\left(z\right)|z\right) = \frac{1}{\left(\overline{n} - 1\right)\left(1 - F\left(R\left(z\right)\right)\right)\left(\overline{v} - \overline{b}_R\left(z\right)\right)},$$

and thus

$$\gamma\left(z\right) = \frac{\overline{n}}{\overline{n} - 1}\frac{1}{\overline{v} - \overline{b}_R\left(z\right)}.$$

Since, by (19),

$$\begin{aligned}
\overline{b}_R\left(z\right) &= \overline{v} - \int_0^1 \left[F(R(z)) + (1 - F(R(z)))t\right]^{\overline{n}-1} V_R^{(1)}\left(t\right) dt \\
&= \overline{v} - \int_0^1 \left[F(R(z)) + (1 - F(R(z)))t\right]^{\overline{n}-1} \\
&\qquad \times V\left[F(R(z)) + (1 - F(R(z)))t\right](1 - F(R(z)))dt \\
&= \overline{v} - \int_{F(R(Z))}^1 u^{\overline{n}-1} V^{(1)}(u) du,
\end{aligned}$$

with $V^{(1)}\left(\cdot\right) > 0$, $R\left(z_1\right) \neq R\left(z_2\right)$ implies that $\overline{b}_n\left(z_1\right) \neq \overline{b}_n\left(z_2\right)$ and then $\gamma\left(z_1\right) \neq \gamma\left(z_2\right)$. This identifies $\overline{v}$ and $\overline{n}$ as

$$\overline{v} = \frac{\gamma\left(z_2\right)\overline{b}_R\left(z_2\right) - \gamma\left(z_1\right)\overline{b}_R\left(z_1\right)}{\gamma\left(z_2\right) - \gamma\left(z_1\right)},$$

$$\overline{n} = \frac{\gamma\left(z\right)\left(\overline{v} - \overline{b}_R\left(z\right)\right)}{\gamma\left(z\right)\left(\overline{v} - \overline{b}_R\left(z\right)\right) - 1}.$$

Since $F^{\overline{n}}(R(\cdot))$ is identified, $F(R(\cdot))$ is identified. Hence $\{G_R(b|z), b \geq R(z), z \in \mathcal{Z}\}$ is identified, since $\{B_R(\alpha), \alpha \geq F(R(z)), z \in \mathcal{Z}\}$ is identified. (18) implies that $F(v)$ is identified for any v such that there is a z satisfying $v \geq R(z)$. Then $\{F(v), v \geq \inf_{z\in\mathcal{Z}} R(z)\}$ is identified by continuity of $F(\cdot)$. □

6.8 Proof of Proposition 3.5

Suppose that an entry model with entry cost satisfying Assumption E generates $\mathbb{G}(\cdot|z)$. That (i)-(iii) hold can be established as in Proposition 3.1. Suppose now that (i)-(iii) hold. Let $\overline{n}$ and $s(z) = q(z)$ be as in (i). For s in $q(\mathcal{Z})$, let $V_c(\alpha|s)$ be as in (iii). Set $F_c(\cdot|s) = V_c^{-1}(\cdot|s)$ and

$$F(v|s) = -\frac{\partial}{\partial s}\left[(1-s) F_c(v|s)\right]$$

which is a conditional c.d.f. over $[\underline{v}, \overline{v}] \times q(\mathcal{Z})$, that can be extended over $[\underline{v}, \overline{v}] \times [0,1]$. Defining $c(\cdot)$ through (21) gives a cost function which rationalizes $\mathbb{G}(\cdot|\cdot)$. □

6.9 Proof of Corollary 3.2

(i) follows from Theorem 2.1, observing that $\overline{n}$ is identified by the number of discontinuities of the p.d.f. $g(\cdot|z)$. The proof of (ii) starts as the proof of Corollary 3.1-(ii), using the probabilities

$$p_{\overline{n}}(z) = \frac{(1-s(z))^{\overline{n}}}{1-s(z)^{\overline{n}} - \overline{n}s(z)^{\overline{n}}(1-s(z))},$$
$$p_{\overline{n}-1}(z) = \frac{\overline{n}(1-s(z))^{\overline{n}-1} s_{\overline{n}}(z)}{1-s(z)^{\overline{n}} - \overline{n}s(z)^{\overline{n}}(1-s(z))},$$

which are identified from $G(\cdot|\cdot)$ if $\overline{n} \geq 3$ to identify $s(z)$. Then $F(\cdot|s)$ can be identified over the identified interval $s(\mathcal{Z})$ using $F(v|s) = -\frac{\partial}{\partial s}\left[(1-s) F_c(v|s)\right]$ and the identification of $c(\cdot)$ over $\mathcal{Z}$ follows from (21). The proof of (iii) is similar to the proof of Corollary 3.1-(iii), identifying $s(z)$ from $\mathbb{G}(\text{not sold}) = s(z)^2$. □

6.10 Proof of Proposition 3.7

The proof proceeds as for Proposition 3.4. Repeating the arguments of Section 6.7 gives that $s^{\overline{n}}(z)$ and

$$\gamma(z) = \frac{\overline{n}}{\overline{n}-1}\frac{1}{\overline{v}-\overline{b}_c(z)}$$

are identified over $\mathcal{Z}$. As in the proof of Proposition 3.4, this holds provided that $\overline{b}_c(z) = B_c(1|s(z))$ is a non-constant function to identify $s(\cdot)$ and $\overline{n}$ as assumed in Proposition 3.7. □

Appendix: discontinuity detection algorithm

Consider a subsample A = Low, Medium, or High. Hereafter, we omit A for convenience. Let W_ℓ, $\ell = 1, \ldots, L$ be winning bids. First, we estimate a tentative discontinuity at each data point as the difference between the density estimates on its left and right sides. Consider a first "small" bandwidth h_0, set to 0.2 in the applications. The tentative discontinuity at $W_{(\ell)}$ is estimated using the difference of left and right k-NN density estimators

$$\widehat{\delta}_{h_0}(W_{(\ell)}) = \frac{\ell - \max\left(\ell - h_0 L/2, 1\right)}{L\left(W_{(\ell)} - W_{(\ell - h_0 L/2)}\right)} - \frac{\min\left(\ell + h_0 L/2, L\right) - \ell}{L\left(W_{(\ell + h_0 L/2)} - W_{(\ell)}\right)},$$

where $\ell - h_0 L/2$ is truncated to 1 if negative and $\ell + h_0 L/2$ to L if larger than L. Second, we estimate the magnitude of the density at each point and calculate a threshold. Define also the k-NN pdf estimator and the critical value[13]

$$\widehat{g}_{h_0}(W_{(\ell)}) = \frac{\min(\ell + h_0 L/2, L) - \max(\ell - h_0 L/2, 1)}{L\left(W_{(\ell+h_0L/2)} - W_{(\ell-h_0L/2)}\right)},$$

$$C_{(\ell)}(\epsilon; h_0) = \widehat{g}_{h_0}(W_{(\ell)})c(\epsilon; h_0) \text{ with}$$

$$c(\epsilon; h_0) = \sqrt{\ln(1/h_0)} + \frac{\ln\ln(1/h_0) - \ln(\pi) + 2\epsilon}{2\sqrt{\ln(1/h_0)}}$$

and where $\epsilon = \epsilon_L$ goes to 0 with L, and is set to 0.01 here.

We now use the tentative discontinuity estimates and thresholds to estimate locations of jump points and jump sizes. If

$$\widehat{\delta}_{h_0}(W_{(\ell_1^\star)}) = \max_{1\leq\ell\leq L} \widehat{\delta}_{h_0}(W_{(\ell)})$$

is smaller than the critical value $C_{(\ell_1^\star)}(\epsilon; h_0)$ then the conditional winning bid pdf $g(\cdot)$ has no discontinuities. Otherwise a discontinuity is found at $W_{(\ell_1^\star)}$, with an estimated jump $\widehat{\delta}_{h_0}(W_{(\ell_1^\star)})$. The next jump will be searched using the same procedure but excluding the indexes ℓ between $\ell_1^\star - h_0L/2$ and $\ell_1^\star + h_0L/2$. The procedure is then iterated until iteration $\widehat{q}$, such that the potential jump is smaller than $C_{(\ell_{\widehat{q}}^\star)}(\epsilon; h_0)$. The number of jumps is then $\widehat{q} - 1$, so that the estimation of the largest number $\overline{n}$ of bidders is $\widehat{\overline{n}} = \widehat{\underline{n}} + \widehat{q}$. Ordering the jumps locations $W_{(\ell_q^\star)}$ gives an estimation of the conditional bid support boundary $\widehat{\overline{b}}_{\underline{n}+q}$ and of the discontinuity jumps $\widehat{\Delta}_{\widehat{\underline{n}}+q}$. A conditional winning bid density estimator incorporating discontinuities is then, for $\ell_q^\star < \ell \leq \ell_{q+1}^\star$ with $\ell_0^\star = 1$,

$$\widehat{g}_{h_1}^d(W_{(\ell)}) = \frac{\min\left(\ell + h_1 L/2, \ell_{q+1}^\star\right) - \max\left(\ell - h_1 L/2, \ell_q^\star + 1\right)}{L\left(W_{\left(\min\left(\ell+h_1L/2, \ell_{q+1}^\star\right)\right)} - W_{\left(\left(max\left(\ell-h_1L/2, \ell_q^\star+1\right)\right)\right)}\right)},$$

where the bandwidth $h_1 > h_0$ is set to 0.5 in our application. Other values of $\widehat{g}_{h_1}^d(\cdot)$ are then obtained by linear interpolation.

[13]Note that the critical value $C_{(\ell)}(\epsilon; h_0)$ is proportional to the estimated pdf, as standard for k-NN estimation. This contrasts with the square root estimated pdf used in Chu and Cheng (1996) for their kernel approach.

References

ABRANTES-METZ, R. & P. BAJARI (2009). Screens for conspiracies and their multiple applications. *Antitrust* **24**, 66–71.

AN, Y., Y. HU & M. SHUM (2010). Estimating first-price auctions with an unknown number of bidders: a misclassification approach. *Journal of Econometrics* **157**, 328–341.

ATHEY, S. & P.A. HAILE (2007). Nonparametric approaches to auctions. *Handbook of Econometrics* **6A**, 3847–3965.

BULOW, J. & AND P. KLEMPERER (1996). Auctions versus negotiations. *American Economic Review* **86**, 180–194.

CAMPO, S., E. GUERRE, I. PERRIGNE & Q. VUONG (2011). Semiparametric estimation of first-price auctions with risk-averse bidders. *Review of Economic Studies* **78**, 112–147.

CHERNOZHUKOV, V. & H. HONG (2004). Likelihood estimation and inference in a class of nonregular econometric models. *Econometrica* **72**, 1445-1480.

CHOI, S., L. NESHEIM & I. RASUL (2016). Reserve price effects in auctions: estimates from multiple regression-discontinuity designs. *Economic Inquiry* **54**, 294–314.

CHU, C.K. & P.E. CHENG (1996). Estimation of a jump points and jump values of a density function. *Statistica Sinica* **6**, 79–95.

COMPIANI, G. & Y. KITAMURA (2016). Using mixtures in econometrics models: a brief review and some new results. *Econometrics Journal* **19**, C95–C127.

Coviello, D. & M. Mariniello (2014). Public requirements in public procurement: evidence from a regression discontinuity design. *Journal of Public Economics* **109**, 76–100.

De Silva, D.G., T., Li & J. Zhao (2019). Efficacy of a bidder training program in procurement auctions with selection and entry. Work in progress. Vanderbilt University.

D'Haultfœuille, X. & P. Février Identification of mixture models using support variations. *Journal of Econometrics* **189**, 70–82.

Freyberger, J. & B.L. Larsen (2017). Identification in ascending auctions with application to digital rights management. Document paper, University of Wisconsin and Stanford University.

Gardes, L. & G. Struffler (2014). Estimation of the conditional tail index using a smoothed local Hill estimator. *Extremes* **17**, 45-–75.

Gayraud, G. (2002). Minimax estimation of a discontinuity for the density. *Journal of Nonparametric Statistics* **14**, 59–66.

Gentry, M. & T. Li (2014). Identification in auctions with selective entry. *Econometrica* **82**, 315–344.

Gentry, M., T. Li & J. Lu (2017). Identification and inference in first-price auctions with risk-averse bidders and selective entry. Working Paper, LSE, Vanderbilt University and NUS.

Gimenes, N. & E. Guerre (2019). Quantile regression methods for first-price auctions. Working paper, PUC Rio and University of Kent.

Guerre, E., I. Perrigne & Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica* **68**, 525–574.

Guerre, E., I. Perrigne & Q. Vuong (2009). Nonparametric identification of risk-aversion in first-price auctions under exclusion restrictions. *Econometrica* **77**, 1193–1227.

Grundl, S. & Y. Zhu (2019). Identification and estimation of risk aversion in first-price auctions with unobserved auction heterogeneity. *Journal of Econometrics* **210**, 363–378.

Haile, P.A., H. Hong & M. Shum (2003). Nonparametric tests for common values in First-price sealed-bid auctions. Document paper, Cowles Foundation.

Han, L. & W.C. Strange (2015). The microstructure of housing markets: search, bargaining, and brokerage. *Handbook of Urban and Regional Economics* **5B**, 813-886.

Hendricks, K. & R.H. Porter (2007). An empirical perspective on auctions. *Handbook of Industrial Organization* **3**, 2073–2143.

Hill, J.B. & A. Shneyerov (2013). Are there common values in first-price auctions? A tail-index nonparametric test. *Journal of Econometrics* **174**, 144–164.

Hirano, K. & J.R. Porter (2003). Asymptotic efficiency in parametric structural models with parameter-dependent support. *Econometrica* **71**, 1307–1338.

Hu, Y. & Y. Sasaki (2016). Identification of paired nonseparable measurement error models. *Econometric Theory* **33**, 955–979.

Ibragimov, I.A. & R.Z. Has'minskii (1981). *Statistical estimation. Asymptotic theory.* Springer.

IMBENS, G.W. & T. LEMIEUX (2008). Regression discontinuity designs: a guide to practice. *Journal of Econometrics* **142**, 615–635.

JALES , H. & Z. YU (2017). Identification and estimation using a density discontinuity approach. *Advances in Econometrics* **38**, 29–72, special issue on “Regression Discontinuity Designs”, Matias D. Cattaneo & Juan Carlos Escanciano Eds.

KIM, D. (2018). Governnment incentives and firm location choices. Job market paper, Yale University.

KLEVEN, H.J. (2016). Bunching. *Annual Review of Economics* **8**, 435–464.

LAFFONT, J.J., H. OSSARD & Q. VUONG (1995). Econometrics of first-price auctions. *Econometrica* **63**, 953–980.

LAMY, L. (2012). The econometrics of auctions with asymmetric anonymous bidders. *Journal of Econometrics* **167**, 113–132.

LI, T. (2005). Econometrics of first-price auctions with entry and binding reservation prices. *Journal of Econometrics* **126**, 173–200.

LI, T. & X. ZHENG (2009). Entry and competition effects in first-price auctions: theory and evidence from procurement auctions. *Review of Economic Studies* **76**, 1397–1429.

LI, T. & X. ZHENG (2012). Information acquisition and/or bid preparation: a structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics* **168**, 29–46.

LIU, N. & Y. LUO (2016). A Nonparametric Test for Comparing Valuation Distributions in First-Price Auctions. Document paper, University of Toronto.

Liu, N., Q. Vuong & H. Xu (2017). Rationalization and identification of binary games with correlated types. *Journal of Econometrics* **201**, 249-–268.

Lu, J. & I. Perrigne (2008). Estimating risk aversion from ascending and sealed-bid auctions: the case of timber auction data. *Journal of Applied Econometrics* **23**, 871-896.

Marmer, V., A. Shneyerov & P. Xu (2013). What model for entry in first-price auctions? A nonparametric approach. *Journal of Econometrics* **176**, 46–58.

Maskin, E. & J. Riley (1984). Optimal auction with risk averse buyers. *Econometrica* **52**, 1473–1518.

McAfee, R.P. & J. McMillan (1987). Auctions with a stochastic number of bidders. *Journal of Economic Theory* **43**, 1–19.

Milgrom, P. (2001). *Putting auction theory to work.* Cambridge University Press.

Oudshoorn, C.G.M. (1998). Asymptotically minimax estimation of a function with jumps. *Bernoulli* **4**, 15–33.

Paarsch, H.J. (1997). Deriving an estimate of the optimal reserve price: an application to British Columbian timber sales. *Journal of Econometrics* **78**, 333–357.

Quint, D. (2015). Identification in symmetric English auctions with additively separable unobserved heterogeneity. Document Paper, University of Wisconsin.

Riley, J.G. & W.F. Samuelson (1981). Optimal auctions. *The American Economic Review* **71**, 381–392.

Slattery, C.R. (2019). Bidding for firms: subsidy competition in the U.S.. Job market paper, University of Virginia.

Shneyerov, A. & A.C.L. Wong (2011). Identification in first-price and Dutch auctions when the number of potential bidders is unobservable. *Games and Economic Behavior* 72, 574–582.

Song, U. (2004). Nonparametric estimation of an eBay auction model with an unknown number of bidders. Working paper, University of British Columbia

Xia, Z. & P.H. Qiu (2015). Jump information criterion for statistical inference in estimating discontinuous curves. *Biometrika* **102**, 397–408.

Ye, L. (2007). Incentive bidding and a theory of two-stage auctions. *Games and Economic Behavior* **58**, 181–207.