Identification in First-Price and Dutch Auctions when the Number of Potential Bidders is Unobservable

Artyom Shneyerov,[a,*] Adam Chi Leung Wong[b,†]

[a]CIREQ, CIRANO and Department of Economics, Concordia University,
1455 de Maisonneuve Blvd. West, Montreal, Quebec H3G 1M8, Canada
[b]School of International Business Administration,
Shanghai University of Finance and Economics,
777 Guoding Road, Shanghai, China 200433

September, 2008
This Version: May, 2010

Abstract

Within the IPV paradigm, we show nonparametric identification of model primitives for first-price and Dutch auctions with a binding reserve price and auction-specific, unobservable sets of potential bidders.

Keywords: Nonparametric identification, Auctions

JEL Classification Numbers: C14, D44.

1 Introduction

Identification in auctions has been an active area of recent research in industrial organization. Beginning with the seminal contributions of Guerre et al. (2000) and Athey and Haile (2002), the literature has explored nonparametric identification of a variety of auction models under progressively weaker assumptions on observables.[1]

We contribute to this literature by showing nonparametric identification for first-price auctions with a binding reserve price r where the set of potential bidders varies from auction to auction and is unobservable. Those potential bidders whose valuations are lower than the reserve price r do not bid (enter). We assume independent private values (IPV). The model allows for ex-ante asymmetries among bidders. Specifically, we assume that bidders may belong to different groups.[2] We assume that only auctions that have attracted at least

*Corresponding author. Tel.: +1 514 848 2424 ext 5288. Fax: +1 514 848 4536. E-mail addresses: achneero@alcor.concordia.ca (A. Shneyerov), wongchileung@gmail.com (A.C.L. Wong)

†The authors are grateful to two anonymous referees and the associate editor for their comments.

[1]See also a recent book by Paarsch et al. (2006).

[2]This approach is adopted in Athey et al. (2004), Flambard and Perrigne (2006), Krasnokutskaya and Seim (2009) and Hubbard and Paarsch (2008).

one actual bidder are observable.[3] The objects we seek to identify are (a) the distribution of valuations $F_i(\cdot)$ for each bidder i, over and above the reserve price, and (b) $p(\cdot)$, the distribution of the sets of potential bidders. We show that these objects are identifiable under conditions that are standard in the theoretical analyses of asymmetric auctions.

As in Paarsch (1997), Athey et al. (2004), Song (2005), Li and Zheng (2009), Adams (2007) and Krasnokutskaya and Seim (2009), our basic identifying assumption is that bidders' valuations do not depend on the set of potential bidders. This is a plausible assumption in many applications. For example, in highway procurement auctions, bidders must be prequalified to participate in the auction based on the ability to perform the work rather than on their costs.[4] Another good example is the procurement of services and materials by the US Government Printing Office (GPO), where bidders are invited to participate through rotating lists.[5]

To illustrate the idea of our identification method, consider a symmetric setting. (We allow asymmetry in our analysis.) Since only the bidders with valuations over and above the reserve price r actually submit bids, the *entry probability* is $1-F(r)$. The identification of this probability is crucial as it is necessary for the identification of both primitive objects in (a) and (b) above.

Assume that the number of potential bidders N has support $\{\underline{N}, \underline{N}+1, ..., \bar{N}\}$. Then the number of actual bidders n has support $\{0, 1, ..., \bar{N}\}$. Since the support of n is observable, $\bar{N}$ is identifiable (in fact, can be consistently estimated as the sample maximum of n). When n takes the maximal possible value $\bar{N}$, the number of potential bidders is observable and also equal to $\bar{N}$.

To identify the entry probability, we use the following trick: when the number of actual bidders is $n = \bar{N} - 1$, the distribution of bids $G(\cdot|n = \bar{N} - 1)$ is a mixture of two components. The first component is the distribution of bids conditional on the number of *potential* bidders $N = \bar{N}$, and the second is the distribution of bids conditional on $N = \bar{N} - 1$. The mixture weights are the probabilities of $N = \bar{N}$ and $N = \bar{N} - 1$, conditional on the number of *actual* bidders $n = \bar{N} - 1$. Using a theoretical result that the upper bounds of bid supports are ordered (also proved in the paper), we show that these mixture weights are identified. They in turn identify the entry probability for every bidder.

We can now identify the distribution of valuations above the reserve price from the distribution of bids $G(\cdot|n = \bar{N})$ using standard methods, as in Guerre et al. (2000). Also, we can exploit the fact that the distribution of the number of n given N is Binomial with parameter $1 - F(r)$, and the marginal distribution of n is directly observable, to identify the distribution of the number of potential bidders N. (This is despite the fact that only the auctions that have attracted at least one actual bidder are observable.)

The above discussion presumes that all submitted bids are observable. In (strategically equivalent) Dutch auctions, only the winning bids can be observed by the econometrician. Still, if all bidder identities are observable, we prove that our identification results extend to Dutch auctions. (The proof uses the results in Berman (1963) and Athey and Haile (2002).)

Hu and Shum (forthcoming 2010), in a paper that is closely related and was concur-

[3] See Hendricks and Porter (2007) for a discussion of the empirical relevance of this assumption.
[4] Krasnokutskaya and Seim (2009).
[5] See `http://www.gpo.gov:80/pdfs/vendors/sfas/ppr.pdf` for a description of GPO auction rules.

rently written, consider identification and estimation of a model similar to ours. The main difference is that they allow the distribution of valuations to depend on the number of potential bidders. (Another difference is that they restrict attention to a symmetric model.) They show that identification nevertheless obtains provided an instrument is available that exogenously determines the number of potential bidders.[6] Their methods are based on recent results in the literature on misclassified regressors and are different from ours.

Several other papers in the empirical auction literature are related to our paper. Paarsch (1997), in his study of the Small Business Forest Enterprise Program (SBFEP) in British Columbia, estimates that the average number of actual bidders is about 3.29. Due to non-participation caused by a binding reserve price, the number of potential bidders exceeds the number of actual bidders. But if one uses a crude measure of the number of potential bidders such as the number of firms registered in the district of the auction, the number of potential bidders could be as high as 185. Clearly, with this measure, one would substantially overestimate the level of potential competition in the majority of auctions. Paarsch (1997) adopts a clever parametric estimation strategy that is based on conditional likelihood and eliminates the need to estimate the number of potential bidders. However, his approach is limited to ascending-bid (English) auctions.

Song (2005) and Adams (2007) consider identification and estimation of eBay auctions with an unknown number of potential bidders. Their methods are tailored for eBay auctions and are entirely different from ours. Song (2005) shows that the joint distribution of any two order statistics identifies the parent distribution. She then applies this result to eBay auctions, by arguing that in equilibrium, the second and third highest bidders bid truthfully. She develops a nonparametric estimator based on her identification result. Adams (2007) shows that, under certain additional assumptions, observing just the transaction price is sufficient for identification.

Most of the papers that estimated first-price auctions approached the measurement of potential competition empirically. In some cases, such a measure is readily available. For example, in highway procurement auctions conducted by state departments of transportation, the list of eligible firms is sometimes publicly released and can serve as a good proxy for potential competition (e.g. Li and Zheng (2009), Krasnokutskaya and Seim (2009) and Marmer et al. (2007)). In other cases, researchers have used geographic proximity as a basis for firm inclusion in the set of potential bidders (Athey et al. (2004), Hendricks et al. (2003)).

Since the structural auction estimates are sensitive to the measure of potential competition (Hendricks and Porter (2007)), another approach is to treat the number of potential bidders as a parameter to be estimated, as in Laffont et al. (1995). Ideally, this parameter would be auction specific, so a model for potential competition would be estimated jointly with the model of bidding. Nonparametric identification of the entire model is necessary as a foundation for such an approach, and our results provide such a foundation.

[6] After the revision work on this paper was completed, we have become aware of a new version of Hu and Shum (forthcoming 2010) where identification is also shown without the instrument. More exactly, a second bid in the auction may serve this purpose.

2 The model

We consider IPV first-price auctions. Bidders are ex-ante asymmetric: we assume that there are m groups of bidders. Within each group the bidders draw valuations from the same distribution F_i, but the distributions F_i may be different across the groups. The set of groups is denoted as $\mathcal{M} \equiv \{1, 2, \cdots, m\}$. The number of potential bidders in group i is denoted as N_i, and we write $N \equiv (N_1, \cdots, N_m)$. We refer to such an auction as N-*auction.* Our most important identifying assumption is that the distribution of valuations does not depend on the composition of bidder groups. (In the symmetric case, this is equivalent to the requirement that the distribution of bidders' valuations does not depend on the number of potential bidders.)

Assumption 1 *The distributions of bidders' valuations do not depend on* N*, i.e.* $\forall N, N' \in \mathbb{Z}_+^m$ *with* $N_i, N_i' > 0$ *we have* $F_i(v|N) = F_i(v|N') \equiv F_i(v)$.

This assumption rules out cases when the decision to become a potential bidder is correlated with the would-be bidder's valuation, for example. We assume that each distribution F_i has the same support, denoted as $[\underline{v}, \bar{v}]$, is differentiable on the support, and has density f_i which is bounded away from zero on its support.[7] The vector N, the distributions $F_i(\cdot|N)$, and the reserve price r are assumed to be commonly known to the bidders. In this setting, Maskin and Riley (2000) and Lebrun (1999) have shown existence and uniqueness of Bayesian-Nash equilibrium bidding strategies $B_i(\cdot|N)$.[8] These results imply that bidders from the same group must use identical bidding strategies.

Nonparticipation in an auction is due to the existence of a binding reserve price $r \in (\underline{v}, \bar{v})$. We assume that the numbers of potential bidders in each auction are unobservable (to the econometrician): only the bidders with valuations at least as high as the reserve price r submit serious bids. We treat non-serious bids as uninformative and ignore them. From now on, it will be assumed that every bidder submits a bid only if his valuation is at least r, thereby becoming an *actual* bidder. The number of actual bidders in group i is denoted as n_i, and we write $n \equiv (n_1, \cdots, n_m)$. The decision to become an actual bidder is called the *entry decision.* Only the auctions that have attracted at least one actual bidder are assumed to be registered in the dataset.

Assumption 2 *The identities of bidders and their bids in each auction are observable by the econometrician.*[9] *The reserve price is also observable and constant across auctions.*

This assumption implies that the vector n of the numbers of actual bidders in each group, is observable if $\sum_{i=1}^{m} n_i > 0$. Denote the C.D.F. of bids from a group i bidder, *conditional on entry* and the vector of potential bidders N, as $G_i^*(\cdot|N)$ $(N_i > 0)$. From the

[7] Identical supports is a standard assumption in the theoretical literature on asymmetric auctions. See e.g. Lebrun (1999). Little is known in general about the existence and uniqueness of equilibrium without this assumption. Also, if the supports are not identical, there exist reserve prices for which the low types are always screened and the identification argument would not go through in general.

[8] See also Bajari (2001).

[9] In fact, only the identities of bidders' groups, rather than that of bidders, need to be observed. It in particular implies that, in a symmetric setting (i.e. there is only one group), the identities are not required to be observed.

econometrician's point of view, N is randomly drawn from some probability distribution p and is unobservable. In other words, N is treated as an auction-specific effect. Since N is unobservable, the data do not reveal this C.D.F. They only reveal the C.D.F. of bids conditional on the numbers of actual bidders $G_i(b|n)$ $(n_i > 0)$.

The support of p is denoted as S, i.e. $p(N) > 0$ if and only if $N \in S$. Both p and S are unobservable.

Assumption 3 *For every group $i \in \mathcal{M}$, there exists some $N \in S$ such that $N_i \geq 2$. More succinctly, $\cup_{N \in S} \{i : N_i \geq 2\} = \mathcal{M}$.*

Without this assumption, we cannot guarantee that equilibrium bidding strategies are strictly increasing on $[r, \bar{v}]$, at least in some auction, for all groups, so that identification of $F_i(v)$ for $v \in [r, \bar{v}]$ might fail.[10]

A bidder from group i becomes active if $v \geq r$, i.e. with probability $1 - F_i(r)$. Since bidders draw their valuations independently, the distribution of n conditional on N is multinomial, with probabilities

$$\pi(n|N) = \prod_{i=1}^{m} \binom{N_i}{n_i} [1 - F_i(r)]^{n_i} [F_i(r)]^{N_i - n_i} \quad (n \leq N). \tag{1}$$

These probabilities are not observable. The marginal probabilities of n are

$$\rho(n) = \sum_{N \in S} p(N) \pi(n|N).$$

The support of ρ is denoted as s. Since the econometrician only observes the auctions with at least one active bidder, the marginal probabilities $\rho(n)$ are also unobservable; only the conditional probabilities

$$\rho^*(n) = \frac{\rho(n)}{1 - \rho(0)} \quad \left(\sum_{i=1}^{m} n_i > 0\right) \tag{2}$$

are observable.

3 Main results

The primitives that we seek to identify are $F_i(\cdot)$ for every $i \in \mathcal{M}$, and $p(N)$ for every $N \in S$. Before we turn to our results, consider the case when N is observable. Then the distribution $G_i^*(\cdot|N)$ and the $p(N)$ are also observable, and we can identify $F_i(r)$ from e.g.

$$\Pr\{n_i = 1|N\} = N_i (1 - F_i(r)) F_i(r)^{N_i - 1}.$$

The distributions $F_i(v|v \geq r)$ can be identified from first-order equilibrium conditions following the approach of Guerre et al. (2000).[11] Denote inverse bidding strategies as $\xi_i(b|N)$.

[10] See Lebrun (1999) and our Appendix for details.
[11] See also the discussion in Athey and Haile (2005).

If $b > r$ and $N_i > 0$, the inverse bidding strategies $\xi_i(b|N)$ can be found from the first-order conditions[12]

$$\xi_i(b|N) = b + \left\{ \sum_{j=1}^{m} \frac{N_j g_j^*(b|N)}{G_j^*(b|N) + \frac{F_j(r)}{1-F_j(r)}} - \frac{g_i^*(b|N)}{G_i^*(b|N) + \frac{F_i(r)}{1-F_i(r)}} \right\}^{-1}, \quad (3)$$

where $g_i^*(\cdot|N)$ is the density of $G_i^*(\cdot|N)$. Since $F_i(r)$ is identifiable, this leads to the identification of group i's bidding strategy $B_i(v|N)$ for $v > r$, and consequently of the distributions of valuations conditionally on entry, $F_i(v|v \geq r) = G_i^*(B_i(v|N)|N)$, and also unconditionally,

$$F_i(v) = [1 - F_i(r)] F_i(v|v \geq r) + F_i(r) \quad (v > r).$$

When N is unobservable, the distributions $G_i^*(\cdot|N)$ are in general also unobservable, but there are special cases in which they are observable. Let $\bar{S}$ and $\bar{s}$ be the maximal sets of the numbers of potential and actual bidders respectively:[13]

$$\bar{S} \equiv \left\{\bar{N} \in S : \nexists N \in S \text{ s.t. } \bar{N} < N\right\},$$
$$\bar{s} \equiv \{\bar{n} \in s : \nexists n \in s \text{ s.t. } \bar{n} < n\}.$$

Lemma 1 (Identification of the Maxiamal Set $\bar{S}$) *We have $\bar{S} = \bar{s}$. Since $\bar{s}$ is observable, the maximal set $\bar{S}$ is identifiable.*

Proof. For any $N \in S$, we have $\pi(n|N) > 0$ if and only if $n \leq N$. Therefore,

$$s = \{n : n \leq N \text{ for some } N \in S\}.$$

The result immediately follows. *Q.E.D.*

A typical element of $\bar{S}$ is denoted as $\bar{N}$.[14]

Remark 1 *Assumption 3 implies that all bidder types are represented in $\bar{S}$, i.e. $\forall i \in \mathcal{M}$ $\exists \bar{N} \in \bar{S}$ such that $\bar{N}_i > 0$ (in fact, $\bar{N}_i \geq 2$).*

When the number of actual bidders is maximal, i.e. $n = \bar{N}$ for some $\bar{N} \in \bar{S}$, obviously $G_i^*(\cdot|n) = G_i(\cdot|n)$. Since the latter distribution is observable, $G_i^*(\cdot|\bar{N})$ is identifiable for all $\bar{N} \in \bar{S}$ and i such that $\bar{N}_i > 0$. Our discussion of the observable N case then implies that, if the entry probabilities $\{1 - F_j(r)\}_{j=1}^m$ are identifiable, then $F_i(v|v \geq r)$ are also identifiable for all $i \in \mathcal{M}$.

[12] For the derivation of (3), see Appendix.

[13] We use the convention that: for any two vectors x_1 and x_2 of the same dimension, $x_1 < x_2$ means $x_1 \leq x_2$ and $x_1 \neq x_2$.

[14] An important issue is how to determine the maximal set $\bar{S}$ in practice when bidders are asymmetric. From the practical perspective, it may be convenient to make a stronger assumption that the support is rectangular: $S = \prod_{i=1}^m \{\underline{N}_i, ..., \bar{N}_i\}$. Then $\bar{S} = \{\bar{N}\}$ and each $\bar{N}_i$ can be consistently estimated as the sample maximum of n_i, essentially in the same way as is commonly done in the symmetric model.

Our main result shows that $F_i(r)$ and $p(N)$ are in fact identifiable. Denote the support of group i's bid distribution in the auction with the number of potential bidders N as $\left[r, \bar{b}(N)\right]$. (Recall that, even though bidders draw their valuations from distributions that may be different, the upper bounds of the supports are common for all bidders.) Our identification proof will rely on the following lemma.

Lemma 2 *$\bar{b}(N)$ is strictly increasing in N.*

It is well known that Lemma 2 always holds in a symmetric IPV model, i.e. when bidders draw their valuations from the same distribution. In the Appendix, we prove it in general. The bounds $\bar{b}\left(\bar{N}\right)$ for $\bar{N} \in \bar{S}$ are identifiable. It is because for $\bar{N} \in \bar{S}$, we observe $G\left(\cdot|\bar{N}\right)$ and the bound $\bar{b}\left(\bar{N}\right)$ is identified as the upper bound of the support of $G\left(\cdot|\bar{N}\right)$. Our main result is the following proposition.

Proposition 1 *$F_i(r)$ and $p(N)$ are identifiable.*

Proof. It is convenient to denote the conditional distribution of $N \geq n$ given n as $\lambda(N|n)$. By Bayes rule,

$$\lambda(N|n) = \frac{\pi(n|N)\, p(N)}{\rho(n)}. \tag{4}$$

Fix an arbitrary group $i \in \mathcal{M}$. Pick an $\bar{N} \in \bar{S}$ such that $\bar{N}_i > 0$. Remark 1 implies that such a choice is possible. We first show that $\lambda\left(\bar{N}|\bar{N}_{-i}\right)$, where

$$\bar{N}_{-i} \equiv \left(\bar{N}_1, \cdots, \bar{N}_{i-1}, \bar{N}_i - 1, \bar{N}_{i+1}, \cdots, \bar{N}_m\right),$$

is identifiable.

Notice that

$$1 - G_i(b|n) = \sum_{N: N \geq n} \lambda(N|n)\left[1 - G_i^*(b|N)\right]. \tag{5}$$

Lemma 2 implies $\bar{b}\left(\bar{N}_{-i}\right) < \bar{b}\left(\bar{N}_i\right)$. Thus if $b \in \left(\bar{b}\left(\bar{N}_{-i}\right), \bar{b}\left(\bar{N}_i\right)\right)$, we have

$$1 - G_i\left(b|\bar{N}_{-i}\right) = \lambda\left(\bar{N}|\bar{N}_{-i}\right)\left[1 - G_i^*\left(b|\bar{N}\right)\right].$$

On the other hand, when $n = \bar{N}$, the sum in (5) contains only one term, equal to $1 - G_i^*\left(b|\bar{N}\right)$. It follows that

$$\lambda\left(\bar{N}|\bar{N}_{-i}\right) = \lim_{b \uparrow \bar{b}(\bar{N})} \frac{1 - G_i\left(b|\bar{N}_{-i}\right)}{1 - G_i\left(b|\bar{N}\right)} \tag{6}$$

is identifiable.

We now show how to recover $F_i(r)$ from $\lambda\left(\bar{N}|\bar{N}_{-i}\right)$. First note that

$$\begin{aligned}
\pi\left(\bar{N}|\bar{N}\right) &= \prod_{i=1}^{m}\left[1-F_i(r)\right]^{\bar{N}_i}, \\
\pi\left(\bar{N}_{-i}|\bar{N}\right) &= \bar{N}_i\left(1-F_i(r)\right)^{\bar{N}_i-1}F_i(r)\cdot\prod_{j\neq i}\left[1-F_j(r)\right]^{\bar{N}_j}, \\
&= \bar{N}_i\frac{F_i(r)}{1-F_i(r)}\pi\left(\bar{N}|\bar{N}\right).
\end{aligned}$$

Then from (4), taking into account (1),

$$\begin{aligned}
\lambda\left(\bar{N}|\bar{N}_{-i}\right) &= \frac{\pi\left(\bar{N}_{-i}|\bar{N}\right)p\left(\bar{N}\right)}{\rho\left(\bar{N}_{-i}\right)} \\
&= \bar{N}_i\frac{F_i(r)}{1-F_i(r)}\pi\left(\bar{N}|\bar{N}\right)\frac{p\left(\bar{N}\right)}{\rho\left(\bar{N}_{-i}\right)}.
\end{aligned}$$

We can combine this equation with

$$\lambda\left(\bar{N}|\bar{N}\right) = \frac{p\left(\bar{N}\right)\pi\left(\bar{N}|\bar{N}\right)}{\rho\left(\bar{N}\right)} = 1$$

to eliminate $p(N)$. This yields

$$\frac{F_i(r)}{1-F_i(r)} = \frac{1}{\bar{N}_i}\lambda\left(\bar{N}|\bar{N}_{-i}\right)\frac{\rho\left(\bar{N}_{-i}\right)}{\rho\left(\bar{N}\right)}. \tag{7}$$

From (2),

$$\frac{\rho\left(\bar{N}_{-i}\right)}{\rho\left(\bar{N}\right)} = \frac{\rho^*\left(\bar{N}_{-i}\right)}{\rho^*\left(\bar{N}\right)},$$

and therefore (7) implies

$$\frac{F_i(r)}{1-F_i(r)} = \frac{1}{\bar{N}_i}\lambda\left(\bar{N}|\bar{N}_{-i}\right)\frac{\rho^*\left(\bar{N}_{-i}\right)}{\rho^*\left(\bar{N}\right)}. \tag{8}$$

Since the right-hand side of this equation contains only identifiable quantities, $F_i(r)$ is identifiable for each $i\in\mathcal{M}$.

Finally, we can recover $p(N)$ from the total probability equations. For $\alpha = 1-\rho(0)$, the law of total probability implies the following system of linear equations for $p(N)$:

$$\alpha\rho^*(n) - \sum_{N:N\geq n}\pi(n|N)\,p(N) = 0. \tag{9}$$

Since $F_i(r)$ are identifiable, $\pi(n|N)$ are also identifiable; see (1). Formally, consider the above system for any $\alpha\in(0,1)$. Write $p(N)$ as $P(N,\alpha)$ to make the dependency on α explicit. Since the probabilities $P(N,\alpha)$ enter the right-hand side of (9) only for $N\geq n$, the system has a recursive structure that allows one to uniquely determine $P(N, \alpha)$ for all N. To see this most easily, we can use an induction argument. Begin with those $N \in \bar{S}$, we have

$$P(N, \alpha) = \frac{\alpha \rho^* (N)}{\pi (N|N)}. \tag{10}$$

Next, for any given $N \notin \bar{S}$, if $p(N', \alpha)$ are known for all $N' > N$, and we can determine $P(N, \alpha)$ from (9) according to

$$P(N, \alpha) = \frac{1}{\pi (N|N)} \times \left[\alpha \rho^* (N) - \sum_{N' : N' > N} \pi \left(N|N'\right) P\left(N', \alpha\right) \right]. \tag{11}$$

To determine α, note that as a solution of a linear system, $P(N, \alpha)$ is homogeneous of degree 1 in α, so that $P(N, \alpha) = \alpha P(N, 1)$. For $\alpha = 1 - \rho(0)$, the law of total probability implies

$$(1 - \rho(0)) \sum_{N \in S} P(N, 1) = 1,$$

Since $P(N, 1)$ are now known, the above equation uniquely determines $\rho(0)$. Therefore $p(N)$ is identified: $p(N) = (1 - \rho(0)) P(N, 1)$. *Q.E.D.*

Remark 2 *We have chosen to abstract from observable auction heterogeneity, a feature almost always present in auction data. But we should stress that all our results are applicable under observable auction heterogeneity. The variation in reserve prices can also be considered as a form of observed heterogeneity. In such a model, one seeks to identify* $F(v|x)$ *and* $p(N|x)$*, where* x *is a vector of auction characteristics that may also include* r*. All our previous results go through if we use conditional distributions* $G_i^*(\cdot|n, x)$ *in place of* $G_i^*(\cdot|n)$*. In particular, the distribution* $F(v|x)$ *is identifiable for* $v \geq r$*.*

4 Extension to Dutch auctions

In this section, we show that our result generalizes to Dutch auctions, where only the winning bid is observable. We continue to assume that the identities of actual bidders are observable. Fix an $\bar{N} \in \bar{S}$. Restrict attention to auctions with $n = \bar{N}$ and groups with $\bar{N}_i > 0$. Let W_i be the highest bid submitted from group i (with $\bar{N}_i > 0$). Let $W \equiv \max_i W_i$ be the winning bid. And let I be the identity of the winning group, i.e. $W = W_I$.

Our data directly reveals the joint distribution of (I, W) conditional on $n = \bar{N}$ (which also implies $N = \bar{N}$):

$$H_i(w|\bar{N}) \equiv \Pr\left(I = i \;\&\; W \leq w | n = \bar{N}\right).$$

Begin by recovering $H_i^*(\cdot|\bar{N})$ the C.D.F. of W_i conditional on $n = N = \bar{N}$. The set of functions $\left\{H_i(\cdot|\bar{N})\right\}$ is related to the set $\left\{H_i^*(\cdot|\bar{N})\right\}$ via the functional equations

$$H_i(w|\bar{N}) = \int_r^w \prod_{j \neq i} H_j^*(t|\bar{N}) dH_i^*(t|\bar{N}).$$

One can verify (see Berman (1963) and Athey and Haile (2002)) that the solution for $\left\{H_i^*(\cdot|\bar{N})\right\}$ is given by

$$H_i^*(w|\bar{N}) = \exp\left\{-\int_w^\infty \left[\sum_j H_j(t|\bar{N})\right]^{-1} dH_i(t|\bar{N})\right\}. \tag{12}$$

Since the right-hand side of (12) contains only observable objects, $H_i^*(w|\bar{N})$ is identifiable. Now recall that $H_i^*(w|\bar{N})$ is the probability that all $\bar{N}_i$ bidders in group i submit bids below w, conditional on $n = N = \bar{N}$. We have

$$H_i^*(w|\bar{N}) = \left[G_i^*(w|\bar{N})\right]^{\bar{N}_i},$$

which proves that $G_i^*(w|\bar{N})$ is identifiable for every $\bar{N} \in \bar{S}$ and every i such that $\bar{N}_i > 0$. This implies that $\xi_i(b|N)$ and therefore $F_i(v|v \geq r)$ are identifiable provided the entry probabilities $\{1 - F_j(r)\}_{j=1}^m$ are identifiable. The rest of the identification proof follows exactly parallel to that of Proposition 1.

5 Concluding remarks

We have shown that a first-price IPV auction model where nonparticipation is due to a binding reserve price, and the set of potential bidders is unobservable, is nonparametrically identified under weak assumptions. We do not develop a nonparametric estimation method. In developing such a method, it would be interesting to consider a situation when variation in the reserve price is conditionally independent of N and V_i. Intuitively, this may lead to over identification which may also help improve efficiency in estimation. This may be an interesting direction for future research.

On the other hand, from an empirical perspective, parametric assumptions are always used in some form. Our results provide a foundation for parametric estimation methods such as in Laffont et al. (1995) or Donald and Paarsch (1996), but with auction-specific number of potential bidders. Generalization to other private value auction models, e.g. with unobserved heterogeneity, either assuming affiliated values as in Li et al. (2002) or within the IPV paradigm as in Krasnokutskaya (2003), is also left for future research.

6 Appendix

This appendix sketches the derivations of equilibrium conditions, and proves Lemma 2. In order to simplify notations, we do not divide bidders into groups like we do in the text. The set of bidders is $\mathcal{N}$ with $2 \leq |\mathcal{N}| < \infty$. Each bidder i draws his valuation v_i from the C.D.F. $F_i(\cdot)$.[15]

From here up to the proof of Lemma 3 below, we fix an $\mathcal{N}$-auction, and thus suppress the dependency of equilibrium objects on $\mathcal{N}$ in our notation, e.g. we write bidder i's inverse

[15] Clearly, from the theoretical point of view the setting here is equivalent to the one we use in the text, although they are different from the econometrician's point of view.

bidding strategy as $\xi_i(\cdot)$ rather than $\xi_i(\cdot|\mathcal{N})$. But when we prove Lemma 2, this dependency will become explicit.

For an $\mathcal{N}$-auction, bidder i solves

$$\max_b (v_i - b) \prod_{j \neq i} F_j(\xi_j(b)).$$

The first-order conditions are

$$\frac{1}{\xi_i(b) - b} = \sum_{j \neq i} \psi_j'(b) \tag{13}$$

where $\psi_j(b) \equiv \log F_j(\xi_j(b))$. These first-order conditions imply

$$\xi_i(b) = b + \left\{ \sum_j \frac{\frac{d}{db} F_j\left(\xi_j(b)\right)}{F_j\left(\xi_j(b)\right)} - \frac{\frac{d}{db} F_i\left(\xi_i(b)\right)}{F_i\left(\xi_i(b)\right)} \right\}^{-1} \tag{14}$$

Formula (3) in the text follows from (14).

Sum (13) over i and then divide through by $|\mathcal{N}| - 1$:

$$\frac{1}{|\mathcal{N}| - 1} \sum_j \frac{1}{\xi_j(b) - b} = \sum_j \psi_j'(b). \tag{15}$$

Subtract (13) from (15), we have

$$\psi_i'(b) = \frac{1}{|\mathcal{N}| - 1} \left[\sum_j \frac{1}{\xi_j(b) - b} - \frac{|\mathcal{N}| - 1}{\xi_i(b) - b} \right].$$

The above equation holds for $b \in (r, \bar{b}]$ where $\bar{b}$ is the equilibrium maximum bid. Therefore for all $b \in (r, \bar{b}]$

$$\xi_i'(b) = \frac{F_i\left(\xi_i(b)\right)}{\left(|\mathcal{N}| - 1\right) f_i\left(\xi_i(b)\right)} \left[\sum_{j \neq i} \frac{1}{\xi_j(b) - b} - \frac{|\mathcal{N}| - 2}{\xi_i(b) - b} \right]. \tag{16}$$

By Lebrun (1999) Theorem 1, the equilibrium is completely characterized by differential equations (16) and the following boundary conditions:

$$\xi_i(r+) \geq r \text{ for all } i, \text{ and } \xi_i(r+) = r \text{ except possibly one bidder}$$

$$\xi_i(\bar{b}) = \bar{v} \text{ for all } i.$$

Lebrun (1999) also shows existence (Theorem 2) and uniqueness (Corollary 1) of the equilibrium.

The proof of Lemma 2 will need the following result.

Lemma 3 *If $|\mathcal{N}| \geq 3$, $i \in \mathcal{N}$, $k \in \mathcal{N}$, and $i \neq k$, then for all $b \in (r, \bar{b}]$,*

$$\xi_i'(b) < \frac{F_i(\xi_i(b))}{(|\mathcal{N}|-2) f_i(\xi_i(b))} \left[\sum_{j \neq i,k} \frac{1}{\xi_j(b) - b} - \frac{|\mathcal{N}|-3}{\xi_i(b) - b} \right].$$

Proof. From $i \neq k$, we can rewrite (16) and get

$$\psi_i'(b) = \frac{1}{|\mathcal{N}|-1} \left[\sum_{j \neq i,k} \frac{1}{\xi_j(b) - b} - \frac{|\mathcal{N}|-3}{\xi_i(b) - b} + \left(\frac{1}{\xi_k(b) - b} - \frac{1}{\xi_i(b) - b} \right) \right]. \qquad (17)$$

From (13),

$$\frac{1}{\xi_k(b) - b} - \frac{1}{\xi_i(b) - b} = \psi_i'(b) - \psi_k'(b).$$

Substitute this into (17) and solve for $\psi_i'(b)$:

$$\psi_i'(b) = \frac{1}{|\mathcal{N}|-2} \left[\sum_{j \neq i,k} \frac{1}{\xi_j(b) - b} - \frac{|\mathcal{N}|-3}{\xi_i(b) - b} - \psi_k'(b) \right].$$

Since $\xi_k'(b) > 0$ for all $b \in (r, \bar{b}]$ and hence $\psi_i'(b) > 0$ as well, we get the result.[16] *Q.E.D.*

Now we can prove Lemma 2.

Proof of Lemma 2. It suffices to prove $\bar{b}(\mathcal{N}) > \bar{b}(\mathcal{N} \setminus \{k\})$ for all $\mathcal{N}$ with $2 \leq |\mathcal{N}| < \infty$. It is trivial if $|\mathcal{N}| = 2$, so suppose $|\mathcal{N}| \geq 3$. Suppose by the way of contradiction that $\bar{b}(\mathcal{N}) \leq \bar{b}(\mathcal{N} \setminus \{k\})$.

Step 1: We claim that, for small enough $\varepsilon > 0$, we have $\xi_i(b|\mathcal{N}) > \xi_i(b|\mathcal{N} \setminus \{k\})$ for all $b \in \left(\bar{b}(\mathcal{N}) - \varepsilon, \bar{b}(\mathcal{N})\right)$ and all $i \in \mathcal{N} \setminus \{k\}$.

This claim is obviously true if $\bar{b}(\mathcal{N}) < \bar{b}(\mathcal{N} \setminus \{k\})$. If $\bar{b}(\mathcal{N}) = \bar{b}(\mathcal{N} \setminus \{k\}) = \bar{b}$, it can be seen from

$$\xi_i'(\bar{b}|\mathcal{N}) = \frac{1}{(|\mathcal{N}|-1) f_i(\bar{v}) \left(\bar{v} - \bar{b}\right)} < \frac{1}{(|\mathcal{N}|-2) f_i(\bar{v}) \left(\bar{v} - \bar{b}\right)} = \xi_i'(\bar{b}|\mathcal{N} \setminus \{k\}).$$

Step 2: We claim that $\xi_i(b|\mathcal{N}) > \xi_i(b|\mathcal{N} \setminus \{k\})$ for all $b \in \left(r, \bar{b}(\mathcal{N})\right)$ and all $i \in \mathcal{N} \setminus \{k\}$.

Suppose not. Then going from $\bar{b}(\mathcal{N})$ downward, Step 1 implies that there is a first (largest) point $b^* \in \left(r, \bar{b}(\mathcal{N})\right)$ such that $\xi_{i^*}(b^*|\mathcal{N}) = \xi_{i^*}(b^*|\mathcal{N} \setminus \{k\})$ for some $i^* \in \mathcal{N} \setminus \{k\}$. Since b^* is the first point, we also have $\xi_j(b^*|\mathcal{N}) \geq \xi_j(b^*|\mathcal{N} \setminus \{k\})$ for all $j \in \mathcal{N} \setminus \{k\}$. Then it is easy to verify that Lemma 3 implies $\xi_{i^*}'(b^*|\mathcal{N}) < \xi_{i^*}'(b^*|\mathcal{N} \setminus \{k\})$. But then $\xi_{i^*}(b^* + \varepsilon|\mathcal{N}) < \xi_{i^*}(b^* + \varepsilon|\mathcal{N} \setminus \{k\})$ for small $\varepsilon > 0$, contradicting to the definition of b^*.

Step 3: It follows from Step 2 and (13) that for each $i \in \mathcal{N} \setminus \{k\}$ and each $b \in \left(r, \bar{b}(\mathcal{N})\right)$,

$$\sum_{j \in \mathcal{N} \setminus \{i,k\}} \psi_j'(b|\mathcal{N} \setminus \{k\}) > \sum_{j \in \mathcal{N} \setminus \{i\}} \psi_j'(b|\mathcal{N}) > \sum_{j \in \mathcal{N} \setminus \{i,k\}} \psi_j'(b|\mathcal{N}).$$

[16] The result that $\xi_k'(b) > 0$ is stronger than strict monotonicity of ξ_k (since $\xi_k'(b)$ might be 0 at isolated points). For its proof, see Lebrun (1997) Lemma A2-2.

Integrate over $\left(r, \bar{b}\left(\mathcal{N}\right)\right)$ and notice that $\psi_j(\bar{b}(\mathcal{N})|\mathcal{N}) = 0 \geq \psi_j(\bar{b}(\mathcal{N})|\mathcal{N}\setminus\{k\})$ for all j,

$$\sum_{j\in\mathcal{N}\setminus\{i,k\}} \log F_j\left(\xi_j(r+|\mathcal{N})\right) > \sum_{j\in\mathcal{N}\setminus\{i,k\}} \log F_j\left(\xi_j(r+|\mathcal{N}\setminus\{k\})\right) \quad \forall i\in\mathcal{N}\setminus\{k\}\,.$$

Therefore, for each $i \in \mathcal{N}\setminus\{k\}$, there is a $j \in \mathcal{N}\setminus\{i,k\}$ such that $\xi_j(r+|\mathcal{N}) > \xi_j(r+|\mathcal{N}\setminus\{k\})$. It follows that $\xi_j(r+|\mathcal{N}) > \xi_j(r+|\mathcal{N}\setminus\{k\}) \geq r$ holds for at least two distinct j's in $\mathcal{N}\setminus\{k\}$, contradicting the boundary condition. *Q.E.D.*

References

Adams, C., 2007. Estimating demand from eBay prices. International Journal of Industrial Organization 25 (6), 1213–1232.

Athey, S., Haile, P., 2002. Identification of Standard Auction Models. Econometrica 70 (6), 2107–2140.

Athey, S., Haile, P., 2005. Nonparametric Approaches to Auctions. Handbook of Econometrics 6.

Athey, S., Levin, J., Seira, E., 2004. Comparing Sealed Bid and Open Auctions: Theory and Evidence from Timber Auctions. Department of Economics, Stanford University.

Bajari, P., 2001. Comparing Competition and Collusion: A Numerical Approach. Economic Theory 18 (1), 187–205.

Berman, S., 1963. Note on Extreme Values, Competing Risks and Semi-Markov Processes. The Annals of Mathematical Statistics 34 (3), 1104–1106.

Donald, S., Paarsch, H., 1996. Identification, Estimation, and Testing in Parametric Empirical Models of Auctions within the Independent Private Values Paradigm. Econometric Theory 12 (3), 517–567.

Flambard, V., Perrigne, I., 2006. Asymmetry in Procurement Auctions: Evidence from Snow Removal Contracts*. The Economic Journal 116 (514), 1014–1036.

Guerre, E., Perrigne, I., Vuong, Q., 2000. Optimal Nonparametric Estimation of First-Price Auctions. Econometrica 68 (3), 525–574.

Hendricks, K., Pinkse, J., Porter, R. H., 2003. Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. Review of Economic Studies 70 (1), 115–145.

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

Hu, Y., Shum, M., forthcoming 2010. Estimating First-Price Auction Models with Unknown Number of Bidders: a Misclassification Approach. Journal of Econometrics.

Hubbard, T., Paarsch, H., 2008. Investigating bid preferences at low-price, sealed-bid auctions with endogenous participation. International Journal of Industrial Organization.

Krasnokutskaya, E., 2003. Identification and estimation in highway procurement auctions under unobserved auction heterogeneity, Working Paper, University of Pennsylvania.

Krasnokutskaya, E., Seim, K., 2009. Bid Preference Programs and Participation in Highway Procurement. Tech. rep., working paper.

Laffont, J., Ossard, H., Vuong, Q., 1995. Econometrics of First-Price Auctions. Econometrica 63 (4), 953–980.

Lebrun, B., 1997. First Price Auctions in the Asymmetric N Bidder Case. Working Paper, Les Cahiers de Recherche du GREEN.

Lebrun, B., 1999. First Price Auctions in the Asymmetric N Bidder Case. International Economic Review 40 (1), 125–142.

Li, T., Perrigne, I., Vuong, Q., 2002. Structural Estimation of the Affiliated Private Value Auction Model. Rand Journal of Economics 33 (2), 171–193.

Li, T., Zheng, X., 2009. Entry and competition effects in first-price auctions: Theory and evidence from procurement auctionsForthcoming, The Review of Economic Studies.

Marmer, V., Shneyerov, A., Xu, P., 2007. What Model for Entry in First-Price Auctions? A Nonparametric Approach. Working Paper.

Maskin, E., Riley, J., 2000. Asymmetric Auctions. Review of Economic Studies 67 (3), 413–438.

Paarsch, H., 1997. Deriving an Estimate of the Optimal Reserve Price: An Application to British Columbian Timber Sales. Journal of Econometrics 78 (2), 333–357.

Paarsch, H., Hong, H., Haley, M., 2006. An Introduction to the Structural Econometrics of Auction Data. MIT Press.

Song, U., 2005. Nonparametric Estimation of an eBay Auction Model with an Unknown Number of Bidders. Working Paper, University of British Columbia.