Identification in English Auctions with Shill Bidding*

Guillermo Marshall†

January 13, 2022

Abstract

This paper investigates the impact of shill bidding on identification in English auctions in the independent private values framework. Shill bidding occurs when sellers bid in their auctions to drive up the price and is a common problem in auction platforms. I show that the distribution of valuations is partially identified in the presence of shill bidding, and I provide bounds for the distribution of valuations that hold even in the absence of shill bidding. I use these bounds to derive the identification region of the optimal reserve price. I apply these results to a sample of eBay auctions.

Keywords: Auctions, partial identification, shill bidding, eBay

1 Introduction

The use of online auctions exploded in the early days of Internet commerce (Einav, Farronato, Levin and Sundaresan 2018). Online auction platforms have reduced matching frictions between buyers and sellers, creating new trade opportunities, but these online marketplaces are far from frictionless. In a survey by the UK's Office of Fair Trading, online auction users reported that shill bidding—i.e., when the seller bids in their own auction to inflate the price—was a common problem on online auction platforms

*I thank Jeff Ely and Rob Porter as well as workshop participants at Northwestern University for helpful comments and suggestions. I am supported in part by funding from the Social Sciences and Humanities Research Council (Canada). All errors are my own.

†University of British Columbia, Sauder School of Business; guillermo.marshall@sauder.ubc.ca

(OFT 2007).[1] A number of eBay users in the UK and US have been prosecuted for using fake identities to drive up auction sales prices, and the penny auctions platform PennyBiddr was even shut down when the platform itself faced allegations of shill bidding.[2]

Why would a seller engage in shill bidding? There are a number of explanations. A first one is that some online auction platforms charge sellers a fee for posting a reserve price (e.g., eBay), which is a fee that the seller can avoid if they use shill bids in lieu of a reserve price (Kauffman and Wood 2005). A second one is that platforms such as eBay have a "second chance offer" feature, which allows a seller to allocate the object to the second highest bidder in the event that the winner fails to pay. This feature of the platform substantially reduces a seller's cost of engaging in shill bidder, as the seller can allocate the good to the second highest bidder in the event that the "winning bid" is a shill bid. A third explanation is that sellers may face uncertainty about primitives that impact the optimal reserve price. For example, a seller may face uncertainty about the exact composition of buyers (if bidders are asymmetric) or the number of bidders that will participate in the auction. As the seller observes the bidding process, the seller can learn about the composition of buyers or the number of bidders and use shill bids to "adjust" the reserve price (Graham, Marshall and Richard 1990, Wang, Hidvégi and Whinston 2001, Andreyanov and Caoui 2020).[3]

I study identification in an English auction with shill bidding in the symmetric independent private values framework.[4] Specifically, I investigate what an econometrician

[1] Shill bidding is also considered a criminal fraud in the UK and US (Kauffman and Wood 2005).

[2] See, for example, "3 Men Are Charged With Fraud In 1,100 Art Auctions on EBay," *The New York Times*, March 9, 2001, "Phony Bids Pose Difficulties, Putting eBay on the Defensive," *The Wall Street Journal*, May 24, 2000, "Officials Accuse Three in Scam To Drive Up Prices in eBay Bids," *The Wall Street Journal*, February 8, 2002, and "How do you catch online auction cheats?," *BBC News*, July 5, 2010. .

[3] The optimal reserve price can vary with the number of bidders in a number of cases. With independent private values, this may occur whenever the distribution of valuations $F(v)$ is such that $v - (1 - F(v))/F'(v)$ is not monotone increasing (see the discussion in Wang et al. (2001)). With affiliated private values, the valuations may depend on a common factor (e.g., market conditions), which may also affect the seller's valuation for the object (i.e., the value from a future sale if the object does not sell in the auction).

[4] See, for example, Hasker and Sickles (2010) for a survey of the use of eBay data in the economic literature.

can learn about the distribution of valuations and the optimal reserve price when using auction data that may be contaminated with shill bidding. The model of shill bidding is incomplete in that I make weak assumptions about the behavior of sellers.

I find that the distribution of valuations is partially identified in the presence of shill bidding. Consider an auction with $n + 1$ bidders, where n are legitimate bidders and one is a shill bidder. Assume that the game ends when there is only one bidder left and define the winning bid as the lowest price at which only one bidder remains active. The presence of a shill bidder implies that the winning bid may not necessarily be the second highest valuation among the n legitimate bidders. If the shill bidder places the highest bid, then the winning bid would be the highest valuation among the n legitimate bidders, whereas if a legitimate bidder places the highest bid, then the winning bid is the greater between the shill bid and the second highest valuation among the legitimate bidders. As a result, the winning bid is bounded between the second and first highest valuations among the n legitimate bidders. This inequality is the basis of the identification region for the distribution of valuations when the econometrician only observes the winning bid. I show how a dataset including data on all bids can be used to make the identification region tighter.

I show that even if there is no shill bidding in the data (i.e., all $n + 1$ bidders are legitimate), the true distribution of valuations will still be contained in this identification region, provided that the assumptions of the symmetric independent private values setting hold. This result allows the econometrician to test whether the symmetric independent private values assumptions are rejected by the data: under the null hypothesis of symmetric independent private values (i.e., all $n+1$ bidders are legitimate and draw their valuations independently from a common distribution), an estimate of the distribution of valuations must be contained in the bounds that I derive. If this condition fails to hold, the data rejects the symmetric independent private values assumptions. Rejection can be driven by bidder asymmetries (e.g., a shill bidder) or correlation in valuations.

I also use the bounds on the distribution of valuations to investigate what an econometrician can learn about the optimal reserve price of an auction when the data may be contaminated with shill bidding. I show that the optimal reserve price is also partially identified in this case, and I provide bounds for the optimal reserve price that become tighter as the econometrician observes more data about the bidding process.

Lastly, I apply these results to a sample of eBay auctions for 3.4-oz bottles of Armani Acqua di Gio perfume (mint condition), which took place between the years 2008 and 2010. I estimate the identification regions for the distribution of valuations and optimal reserve price, and I show that these estimates reject the symmetric independent private values assumptions, suggesting shill bidding (or some form of bidder asymmetry) or correlation in valuations.

This paper contributes to the literature on identification in auctions. Athey and Haile (2002) study identification in standard auctions. Haile and Tamer (2003) present identification results in an English auction that deviates from the "button auction" abstraction imposing weak conditions on bidder behavior. Song (2004) proposes a method to identify the distribution of valuations in English auctions with an unknown number of bidders. Tang (2011) bounds the revenue distributions of an auction under counterfactual formats. The results are derived without imposing parametric restrictions on the model structure and allow for affiliated values and signals. In related work, Coey, Larsen and Sweeney (2019) propose a test of independence of valuations and the number of valuations in ascending button auctions with symmetric independent private values, which can be used to bound counterfactual revenue distributions. More broadly, see Athey and Haile (2007) and Hendricks and Porter (2007) for literature surveys. My contribution is to derive an identification region for the distribution of valuations that is robust to shill bidding in the context of English auctions.

This paper is also related to the literature studying collusion in auctions, which is another type of auction fraud. The empirical literature studying collusion in auctions has mostly focused on detection and testing competitive versus collusive bidding rather than on the identification of objects of interest when the data may be contaminated with fraudulent behavior. See for example Feinstein, Block and Nold (1985), Porter and Zona (1993), Baldwin, Marshall and Richard (1997), Porter and Zona (1999), and Marmer, Shneyerov and Kaplan (2016). My contribution is in providing identification results in the presence of fraudulent behavior rather than the provide methods to detect such behavior.

The paper is organized as follows. Section 2 presents results on the identification of the distribution of valuations and discusses estimation. Section 3 investigates whether the optimal reserve price can be identified from auction data with shill bidding. Section 4 presents the empirical application using eBay auction data, and Section 5 concludes.

2 Identification of the distribution of valuations

2.1 Identification when observing only the winning bid

Consider an English auction with $n+1$ bidders, where n of these bidders have valuations for the object that are independent and drawn from a distribution F_v (i.e., the *legitimate* bidders) and one is a shill bidder.[5] Assume that the game ends when there is only one bidder left and define the winning bid as the lowest price at which only one bidder remains active.

From the perspective of the legitimate bidders, a shill bidder is equivalent to having a flexible reserve price. The presence of a shill bidder does not affect the legitimate bidders' strategy, as it is weakly dominant for a legitimate bidder with valuation v to remain active in the auction until the winning bid reaches v, as the game ends when there is only one active bidder.

The presence of a shill bidder in an English auction implies that the winning bid might not be equal to the second highest valuation among the legitimate bidders. If the shill bidder wins the object, then the winning bid would be the highest valuation for the object among all the legitimate bidders (i.e., the valuation of the last legitimate bidder to drop out). If the legitimate bidder with the highest valuation wins the object, the winning bid will be the maximum between the second highest valuation among legitimate bidders and the shill bid. That is, the winning bid in an auction with $n+1$ bidders is bounded between the highest and second highest valuations among all n legitimate bidders. This observation allows the econometrician to bound the distribution of valuations, as indicated in the following proposition.

Proposition 1 *Consider the environment described above, and suppose that the econometrician observes only the winning bid and the total number of bidders $(n_i+1) \in \Omega$ of every auction i, where n_i denotes the number of legitimate bidders and Ω is the set of unique values of $n+1$ that are observed by the econometrician. Suppose that the econometrician cannot distinguish between the shill bidder and the legitimate bidders. Then,*

[5]Throughout the paper, I assume that the number of bidders is the number of observed bidders. Song (2004) discusses identification when relaxing this assumption.

the identification region for $F_v(t)$ is given by

$$\begin{aligned} H[F_v(t)] &= \left[\max_{(n+1)\in\Omega} \phi_2^{-1}(F_{w,(n+1)}(t)|n), \min_{(n+1)\in\Omega} \phi_1^{-1}(F_{w,(n+1)}(t)|n)\right] \\ &\equiv [L_w(t), U_w(t)], \end{aligned}$$

where $F_{w,(n+1)}(t)$ is the distribution of the winning bid (i.e., transaction price) when the total number of bidders is $n+1$, and $\phi_1(\cdot|n)$ and $\phi_2(\cdot|n)$ are defined as

$$\phi_1(x|n) = x^n \quad \text{and} \quad \phi_2(x|n) = n(n-1)\int_0^x u^{n-2}(1-u)du.$$

Proof. The result follows from the observation that the winning bid with $(n+1)$ bidders, $W_{(n+1)}$,[6] is bounded above by the highest valuation among the n legitimate bidders, $V_{1:n}$, and is bounded below by the second highest valuation among the n legitimate bidders, $V_{2:n}$. That is,

$$V_{2:n} \leq W_{(n+1)} \leq V_{1:n}.$$

Hence,

$$F_{1:n}(t) \leq F_{w,(n+1)}(t) \leq F_{2:n}(t),\ \forall t, \forall (n+1). \tag{1}$$

Using the distribution of the first- and second-order statistics,

$$F_{1:n}(t) = F_v(t)^n \equiv \phi_1[F_v(t)|n] \tag{2}$$

and

$$F_{2:n}(t) = n(n-1)\int_0^{F_v(t)} u^{n-2}(1-u)du \equiv \phi_2[F_v(t)|n], \tag{3}$$

and that ϕ_1 and ϕ_2 are strictly monotonic functions in $F_v(t)$, inequality (1) implies

$$F_v(t) \leq \phi_1^{-1}(F_w(t)|n),\ \forall t, \forall (n+1).$$

and

$$\phi_2^{-1}(F_w(t)|n) \leq F_v(t),\ \forall t, \forall (n+1).$$

Putting these two inequalities together, we obtain

$$\phi_2^{-1}(F_w(t)|n) \leq F_v(t) \leq \phi_1^{-1}(F_w(t)|n),\ \forall t, \forall (n+1).$$

[6] I denote random variables in upper case and their realizations in lower case.

Intersecting these inequalities over all $(n+1) \in \Omega$ yields the result.

■

As I show in the next proposition, the bounds presented in Proposition 1 can be viewed as an identification region for the distribution of valuations that is robust to the presence of shill bidding. That is, the true distribution of valuations will be contained in the bounds derived in Proposition 1 even in the absence of shill bidding (i.e., the case where all $n+1$ bidders are legitimate bidders), provided that all $n+1$ bidders are legitimate and draw their valuations independently from a common distribution (i.e., symmetric independent private values). This makes the bounds in Proposition 1 informative about the distribution of valuations in any event.

Proposition 2 *If instead of having n legitimate bidders and 1 shill bidder, we have $n+1$ legitimate bidders with valuations drawn from F_v, then*

$$F_v(t) \equiv \phi_2^{-1}(F_{w,(n+1)}(t)|n+1) \in [\phi_2^{-1}(F_{w,(n+1)}(t)|n), \phi_1^{-1}(F_{w,(n+1)}(t)|n)].$$

Since this holds for all $n+1 \in \Omega$, Proposition 1 provides bounds to F_v that are robust to shill bidding.

Proof. See Appendix A for all omitted proofs. ■

Proposition 2 can also be used as the basis of a specification test, as the result in the proposition only holds if all bidders are drawing their valuations independently from F_v (i.e., a symmetric independent private values). Under the null hypothesis of symmetric independent private values, one can estimate the distribution of valuations using standard methods, i.e., the winning bid equals the second highest valuation among all $n+1$ bidders (see, for example, the identification results in Athey and Haile (2002)). The estimate $\hat{F}_v$ should then lie within the identification region for F_v, as indicated by Proposition 2. If this condition fails to hold, one can conclude that the data rejects the symmetric independent private values assumptions. Rejection can come from bidder asymmetries (e.g., a shill bidder drawing "valuations" from a distribution that is not F_v or asymmetric bidders more broadly) or correlation in the valuations. I note that this specification test does not rely on variation in the number of bidders, as does the specification test in Athey and Haile (2002); the test relies on properties of order statistics, which is a novelty.

Corollary 1 *A test for the symmetric independent values framework is given by checking whether*

$$\hat{F}_v(t) \in \hat{H}[F_v(t)]$$

holds for all t, where $\hat{F}_v(t)$ is an estimate of $F_v(t)$ assuming that the valuations of all $n+1$ bidders come from the same distribution $F_v(\cdot)$) (i.e., no shill bidding), and $\hat{H}[F_v(t)]$ is an estimate of the identification region given in Proposition 1.

2.2 Identification when observing all bids

I next present the identification region for the distribution of valuations when the econometrician observes all bids. As mentioned above, the presence of a shill bidder does not affect the strategy of a legitimate bidder, that is, it is still weakly dominant for a bidder with valuation v to remain active in the auction until the winning bid reaches v.

Consider an auction in particular and consider the bidder with the k-th highest valuation among the n legitimate bidders, $v_{k:n}$, with $k \in \{2, ..., n\}$. The bidder's valuation can be bounded as follows

$$b_{(k+1):(n+1)} \leq v_{k:n} \leq b_{k:(n+1)},$$

where $b_{k:(n+1)}$ and $b_{(k+1):(n+1)}$ denote the k-th and $(k+1)$-th highest bids, respectively, out of all $n+1$ bidders. These inequalities can be understood analyzing two separate cases. If the shill bidder exits the auction with a bid that is less than $v_{k:n}$, then the bidder with the k-th highest valuation will submit the k-th highest bid among all $n+1$ bidders, $b_{k:(n+1)}$, and it will be exactly equal to $v_{k:n}$ since $k \geq 2$. If instead, the shill bidder exits the auction with a bid above $v_{k:n}$, then the bidder with the k-th highest valuation will submit the $(k+1)$-th highest bid among all $n+1$ bidders, $b_{(k+1):(n+1)}$, and it will be exactly equal to $v_{k:n}$. These two observations imply the inequalities. The proposition below incorporates this extra information when constructing the identification region for $F_v(t)$, which results in bounds that are tighter than the ones derived in Proposition 1.

Proposition 3 *Consider the environment described above, and suppose that the econometrician observes all bids and the total number of bidders $(n_t+1) \in \Omega$ of each auction*

t. Suppose that the econometrician cannot distinguish between the shill bidder and the legitimate bidders. Then the identification region for $F_v(t)$ is given by

$$\begin{aligned} H[F_v(t)] &= \left[\max_{(n+1)\in\Omega} \left\{ \max_{k\in\{2,...,n\}} \phi_k^{-1}(G_{k:(n+1)}(t)|n) \right\}, \right. \\ &\qquad \left. \min_{(n+1)\in\Omega} \left\{ \min_{k\in\{1,...,n\}} \phi_k^{-1}(G_{(k+1):(n+1)}(t)|n) \right\} \right] \\ &\equiv [L_{all}(t), U_{all}(t)], \end{aligned}$$

where $G_{k:(n+1)}(t)$ is the distribution of the kth highest bid among $(n+1)$ bidders, where $\phi_k(s|n)$ is defined as

$$\phi_k(s|n) = \frac{n!}{(n-k)!(k-1)!} \int_0^s x^{n-k}(1-x)^{k-1}dx.$$

2.3 Estimation

Consider a sequence of T independent auctions. Each auction i has n_i+1 bidders, where n_i of them draw their valuation independently from F_v, while a shill bidder draws his "exit point" from some distribution H. Let Ω be the set of all values of n_i+1.

The estimator for the distribution function of the kth highest bid among all $n+1$ bids (which include legitimate bids and that of the shill bidder), $G_{k:(n+1)}(t)$ with $k \in \{1, ..., n+1\}$, is given by

$$G_{k:(n+1),T}(t) = \frac{1}{T_{n+1}} \sum_{i=1}^{T} 1\{m_i = (n+1); b^i_{k:(n+1)} \le t\},$$

where $T_{n+1} = \sum_{i=1}^T 1\{m_i = (n+1)\}$ and m_t is total number of bidders (including the shill bidder). Using these estimates, and the definition of ϕ_k, $k \in \{1, ..., n+1\}$,

$$\phi_k(s|n) = \frac{n!}{(n-k)!(k-1)!} \int_0^s x^{n-k}(1-x)^{k-1}dx,$$

an estimate for the identification region of $F_v(t)$ in the case where the econometrician observes all bids is given by

$$\begin{aligned} H_T[F_v(t)] &= \left[\max_{(n+1)\in\Omega} \left\{ \max_{k\in\{2,...,n\}} \phi_k^{-1}(G_{k:(n+1),T}(t)|n) \right\}, \right. \\ &\qquad \left. \min_{(n+1)\in\Omega} \left\{ \min_{k\in\{1,...,n\}} \phi_k^{-1}(G_{(k+1):(n+1),T}(t)|n) \right\} \right] \\ &\equiv [L_{all,T}(t), U_{all,T}(t)]. \end{aligned}$$

Similarly, an estimator for $H[F_v(t)]$ when the econometrician only observes the winning bid is given by

$$\begin{aligned} H_T[F_v(t)] &= \left[\max_{(n+1)\in\Omega} \phi_2^{-1}(G_{1:(n+1),T}(t)|n), \min_{(n+1)\in\Omega} \phi_1^{-1}(G_{1:(n+1),T}(t)|n)\right] \\ &\equiv [L_{w,T}(t), U_{w,T}(t)]. \end{aligned}$$

The following proposition establishes the consistency of these estimators.

Proposition 4 (Consistency) *Let us consider a sequence of T independent auctions, where in each auction k with total number of bidders $(n_k+1) \in \Omega$, n_k of them draw their valuation independently from $F_v : [\underline{v}, \bar{v}] \to [0,1]$, while a shill bidder draws his "exit point" from some distribution $H : [\underline{v}, \bar{v}] \to [0,1]$. Suppose that for each $m \in \Omega$, $T_m \to \infty$ as $T \to \infty$. Then, as $T \to \infty$,*

a) $L_{w,T}(t) \overset{a.s.}{\to} L_w(t)$ uniformly in t, and $U_{w,T}(t) \overset{a.s.}{\to} U_w(t)$ uniformly in t;

b) $L_{all,T}(t) \overset{a.s.}{\to} L_{all}(t)$ uniformly in t, and $U_{all,T}(t) \overset{a.s.}{\to} U_{all}(t)$ uniformly in t.

While these estimators are consistent, the estimators may be biased in small samples because of the concavity (convexity) of the min (max) function, as discussed in Haile and Tamer (2003). To see the problem, consider the estimate for the lower bound of the identification region, which amounts to taking the point-wise maximum of a number of cumulative distribution functions. In small samples, taking the maximum of these estimated cumulative distribution functions will tend to select an estimate with upward estimation error, which will lead to an upward bias of the lower bound. A similar problem arises for the upper bound of the identification region, but with a downward bias.

To alleviate the problem, Haile and Tamer (2003) replace the min (max) function in their estimators with a smooth weighted average of the estimated cumulative distribution functions that approximates the min (max). Specifically, they define the function

$$\mu(\hat{y}_1, \cdots, \hat{y}_J; \rho_T) = \sum_{j=1}^{J} \hat{y}_j \left[\frac{\exp(\hat{y}_j \rho_T)}{\sum_{k=1}^{J} \exp(\hat{y}_k \rho_T)}\right] \tag{4}$$

for $\rho_T \in \mathbb{R}$. When $\rho_T \to -\infty$, $\mu(\hat{y}_1, \cdots, \hat{y}_J; \rho_T)$ converges to $\min(\hat{y}_1, \cdots, \hat{y}_J)$. Likewise, when $\rho_T \to \infty$, $\mu(\hat{y}_1, \cdots, \hat{y}_J; \rho_T)$ converges to $\max(\hat{y}_1, \cdots, \hat{y}_J)$. For estimation, the authors replace the min (max) functions with the function in equation (4) choosing values of ρ_T that decrease (increase) to minus infinity (infinity) at an appropriate rate as $T \to \infty$ to ensure consistency. Following Haile and Tamer (2003), I make use of these smooth weighted averages in my empirical application to alleviate small sample bias.

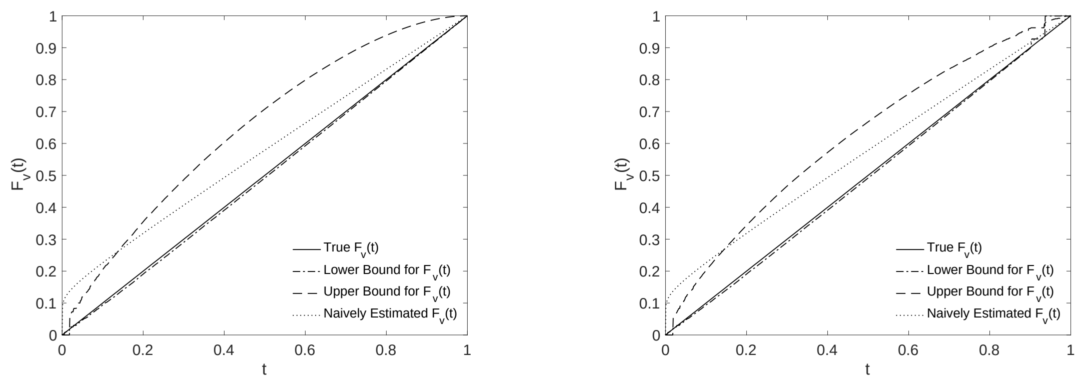
2.4 Monte Carlo simulations

I consider a sequence of T=1 million auctions, half of them with $n = 5$ legitimate bidders and the other half with $n = 4$ legitimate bidders. The valuations of all legitimate bidders are independent draws from a uniform distribution on the interval [0,1]. All auctions have one shill bidder. That is, the total number of players in an each auction i is $n_i + 1 \in \{5, 6\}$. The shill bidder bids independently up to a value S drawn from a distribution with cumulative distribution function H and support [0,1], and where $H(s) = s^{1/10}$. Note that the identification results above hold regardless of whether the shill bids are independent of the behavior of bidders—I make the independence assumption here for simplicity.

Figure 1 displays the actual distribution of valuations and the bounds derived using data on the winning bid only (panel A) and when using data on all bids (panel B). As discussed above, the figure shows how the bounds on the distribution of valuations become tighter when using more information, i.e., data on all bids.

Finally, I estimate the distribution of valuations under the assumption of no shill bidding (i.e., the naively-estimated distribution of valuations).[7] As expected, Figure 1 shows that the naive estimates fail to match the true distribution of valuations. Further, the figure shows that the naive estimates fall outside the identification region for the distribution of valuations, implying that the test proposed in Corollary 1 successfully detects the asymmetry created by shill bidding, which violates the symmetric independent private values assumptions.

[7]Using the prediction that bidders bid up to their valuation, the distribution of bids is an estimate of the distribution of valuations.



A) Winning bid only

B) All bids

Figure 1: Monte Carlo simulations: Value Distribution and Estimated Bounds using results in Propositions 1 and 3.

3 Identification of the optimal (fixed) reserve price

I next turn to studying identification of the optimal reserve price when using data of auctions with shill bidding. I assume that the seller wishes to set a fixed rather than a flexible reserve price (i.e., the seller does not wish to engage in shill bidding), making use of auction data in which sellers may have submitted shill bids.

In this section, I additionally make the following regularity assumptions.

Assumption 1

a) The distribution of valuations, F_v, is continuously differentiable.

b) The support of the distribution F_v is a compact interval, $[\underline{v}, \bar{v}]$.

Under these assumptions, the existence of an optimal reserve price is guaranteed for a number of bidders n. The optimal reserve price is given by the solution to the problem of the seller who wishes to maximize expected revenue:

$$\max_{r \in [\underline{v}, \bar{v}]} u_n(r|v_0) = \max_{r \in [\underline{v}, \bar{v}]} v_0 F_v(r)^n + n \int_r^{\bar{v}} (F_v(v) + vF_v'(v) - 1) F_v^{n-1}(v) dv,$$

where v_0 is the sellers valuation for the object (Riley and Samuelson 1981). I do not assume that the distribution of valuations satisfies the Myerson condition (i.e., $x - (1 - F_v(x))/F'(x)$ is monotone increasing), which may lead to multiple solutions or solutions that depend on the number of bidders (Wang et al. 2001).

Assumption 2 *The shill bidder's distribution of "exit points" is continuously differentiable.*

I also assume that the shill bidder's distribution of "exit points" is continuously differentiable. These assumptions together imply that the upper and lower bounds of $H[F_v(\cdot)]$ form continuously differentiable distribution functions, which is a testable implication.

3.1 Identification

Define the following bounds for the seller's expected revenue when the seller values the object at v_0 and faces n bidders,

$$u_n^U(r|v_0) = v_0 F_L(r)^n + n\int_r^{\bar{v}} (F_L(v) + vF_L'(v) - 1)F_L^{n-1}(v)dv \tag{5}$$

and

$$u_n^L(r|v_0) = v_0 F_H(r)^n + n\int_r^{\bar{v}} (F_H(v) + vF_H'(v) - 1)F_H^{n-1}(v)dv, \tag{6}$$

where $F_L(\cdot)$ and $F_U(\cdot)$ are given by the lower and upper bounds, respectively, of $H[F_v(\cdot)]$. Here I restrict to reserve prices that lie above of the seller's valuation for the object, and to distributions $G(\cdot) \in H[F_v(\cdot)]$ that are consistent with assumptions 1 and 2.

To see that $u_n^L(r|v_0)$ and $u_n^U(r|v_0)$ are in fact bounds for the seller's expected revenue, one can show that for $r \in [v_0, \bar{v}]$,

$$u_n(r|v_0, F) \geq u_n(r|v_0, G)$$

if $F(t) \leq G(t)$, $\forall t$. Further, since for any distribution function $G(\cdot) \in H[F_v(\cdot)]$ that is consistent with assumptions 1 and 2, $F_L(t) \leq G(t) \leq F_U(t)$ for all $t \in [\underline{v}, \bar{v}]$, the result follows.

The argument behind the identification approach can be illustrated using Figure 2. The dotted line in the figure is the constant function that takes the value given by

$$\sup_{r\in[\underline{v},\bar{v}]} u_n^L(r).$$

Since u_n^L is a lower bound for the true expected revenue function, u_n, we know that the true optimal reserve price(s), r^*, must satisfy

$$u_n(r^*) \geq \sup_{r\in[\underline{v},\bar{v}]} u_n^L(r).$$

At the same time, it must be that

$$u_n(r^*) \leq u_n^U(r^*),$$

since u_n^U is an upper bound for u_n. In this sense, the peak(s) of the function u_n, that give(s) us the optimal reserve price(s), can be achieved at any point r such that

$$\sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a) \leq u_n^U(r).$$

This set of points defines the identification region for the optimal reserve price when the seller faces n bidders. When the seller faces uncertainty about the number of bidders that they will face in the auction to be run, this set can be computed for each plausible value of the number of bidders, and the identification region for the optimal reserve price will be the union of these sets.[8]

Proposition 5 *Assume that Assumptions 1 and 2 hold. Given v_0, $F_L(\cdot)$ and $F_U(\cdot)$ (defined by the lower and upper bounds, respectively, of $H[F_v(\cdot)]$), $u_n^U(\cdot|v_0)$ and $u_n^L(\cdot|v_0)$ (defined in (5) and (6), respectively), the identification region of the optimal reserve price is given by*

$$H[r^*] = \bigcup_{n \in \aleph} \left\{ r \in [v_0, \bar{v}] : u_n^U(r) \geq \sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a) \right\},$$

where $\aleph$ is the set of possible number of bidders (potentially a singleton).

3.2 Estimation

Note that one can rewrite $u_n^j(r|v_0)$, for $j \in \{L, H\}$, as

$$\begin{aligned} u_n^j(r|v_0) &= v_0 F_{-j}(r)^n + n \int_r^{\bar{v}} (F_{-j}(v) + vF'_{-j}(v) - 1) F_{-j}^{n-1}(v) dv \\ &= \bar{v} + (v_0 - r) F_{-j}(r)^n - \int_r^{\bar{v}} F_{-j}^{n-1}(v)[n(1 - F_{-j}(v)) + F_{-j}(v)] dv, \end{aligned}$$

where $F_{-U}(\cdot) = F_L(\cdot)$ and $F_{-L}(\cdot) = F_U(\cdot)$. This expression is convenient for estimation, as it saves the econometrician from estimating $F'_{-j}(\cdot)$.

[8] Uncertainty about the number of bidders does not refer to uncertainty about the number of bidders in the auctions in the data.

Proposition 4 shows that the upper and lower bounds of the identification region for the distribution of valuations are uniformly consistent. Depending on the nature of the data used in the estimation, i.e., all bids (all) or winning bids only (w), the functions $F_{L,T}$ and $F_{U,T}$ are given by $F_{L,T}(t) \equiv L_{j,T}(t)$ and $F_{U,T}(t) \equiv U_{j,T}(t)$, for $j \in \{w, all\}$.

Having estimates of $F_U(\cdot)$ and $F_L(\cdot)$, and given v_0, the econometrician can estimate $u_n^L(\cdot|v_0)$ and $u_n^U(\cdot|v_0)$ using:

$$\begin{aligned} u_{n,T}^L(r|v_0) &= \bar{v} + (v_0 - r)F_{U,T}(r)^n - \int_r^{\bar{v}} F_{U,T}^{n-1}(v)[n(1 - F_{U,T}(v)) + F_{U,T}(v)]dv \\ u_{n,T}^U(r|v_0) &= \bar{v} + (v_0 - r)F_{L,T}(r)^n - \int_r^{\bar{v}} F_{L,T}^{n-1}(v)[n(1 - F_{L,T}(v)) + F_{L,T}(v)]dv, \end{aligned}$$

which I prove are uniformly consistent (i.e., $u_{n,T}^j(r|v_0) \overset{a.s.}{\to} u_n^j(r|v_0)$ uniformly in r, for $j \in \{L, U\}$) in Lemma 3 in Appendix A.

In order to estimate the optimal reserve price, I define the following function,

$$Q_n(t) = \max\left\{0, \sup_a u_n^L(a) - u_n^U(t)\right\}, \tag{7}$$

which is defined for a given value of the number of bidders, n. By the previous discussion, it follows that the identification region for the optimal reserve price is given by

$$\Xi_n \equiv \arg\min_{t \in [\underline{v}, \bar{v}]} Q_n(t).$$

The sample analogue of $Q_n(t)$ can be defined as

$$Q_{n,T}(t) = \max\left\{0, \sup_a u_{T,n}^L(a) - u_{T,n}^U(t)\right\}, \tag{8}$$

which in turn leads to the sample analogue of $Q_n(t)$, which can be defined as

$$\Xi_{n,T} \equiv \left\{t \in [\underline{v}, \bar{v}] : Q_{n,T}(t) \leq \inf_s Q_{n,T}(s) + \varepsilon_T\right\},$$

where $\varepsilon_T \to 0$ as $T \to \infty$.

In order to discuss consistency of the estimator $\Xi_{n,T}$, a notion of distance between two sets must be used. Consider two non-empty sets $A, B \subset \mathbb{R}^K$ and define $\rho(A, B) = \sup_{a \in A} \inf_{b \in B} |a - b|$. The Hausdorff distance between both sets is given by

$$d_H(A, B) = \max\{\rho(A, B), \rho(B, A)\}.$$

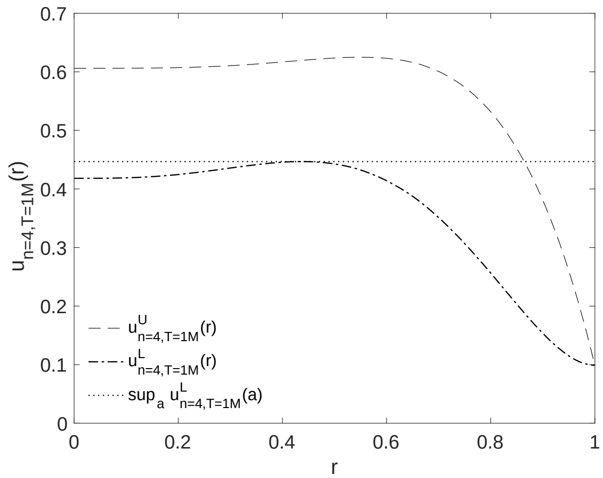


Figure 2: English Auction with Shilling. Estimated Identification Region for Optimal Reserve Price when Number of Bidders $n = 4$. T=1 million. $\Xi_{n=4,T=1M} = [0.1, 0.8642]$.

Proposition 6 (Consistency) *Suppose the conditions in Proposition 4 hold. Let the set of possible reserve prices be $[\underline{v}, \bar{v}]$. Let $Q_n(t)$ and $Q_{n,T}(t)$ be defined as in (7) and (8), respectively. Let $T \to \infty$, and $\varepsilon_T \overset{a.s.}{\to} 0$.*

a) Then $\rho(\Xi_{n,T}, \Xi_n) \overset{a.s.}{\to} 0$.

b) Let $\sup_{t \in R} |Q_{n,T}(t) - Q_n(t)|/\varepsilon_T \overset{a.s.}{\to} 0$. Then $\rho(\Xi_n, \Xi_{n,T}) \overset{a.s.}{\to} 0$.

3.3 Monte Carlo simulations

Following the example in Section 2.4, I estimate the identification region for the optimal reserve price using the bounds for $F_v(t)$ associated to the case when the econometrician only observes winning bids. Further, I consider the case in which the number of bidders that the seller expects to face is exactly equal to four, $\aleph = \{4\}$, and that seller's valuation for the object is $v_0 = 0.1$. Under this environment, the optimal reserve price is given by $r^* = 0.55$. Figure 2 depicts $u^L_{n=4,T=1M}(r)$ and $u^U_{n=4,T=1M}(r)$. Using the results in Proposition 5, the estimated identification region for the optimal reserve price in this example is given by

$$\Xi_{n=4,T=1M} = [0.1, 0.8642].$$

Table 1: Summary Statistics: Armani Acqua di Gio (3.4 oz) eBay auctions

Number of bidders	Number of auctions	Mean price (USD)	St. Dev. price	Min price	Max price
3	160	35.21	6.95	16.49	56.00
4	238	36.05	6.46	15.50	61.00
5	275	36.60	6.14	18.50	56.00
All	673	36.07	6.46	15.50	61.00

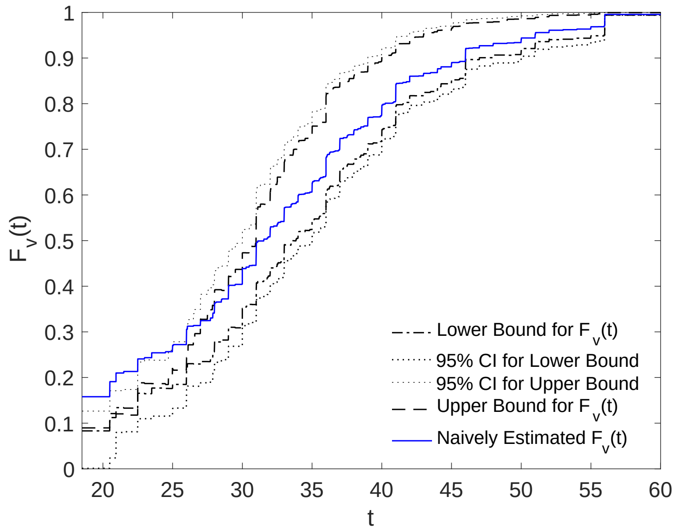
4 An application to eBay auctions

I apply my results on a sample of eBay auctions. The data include 673 auctions of sealed containers (mint condition) of Armani Acqua di Gio perfume (3.4 oz), which took place between the years 2008 and 2010. The auction format is the standard eBay format which resembles an English auction. The auctions in the sample do not feature the buy-it-now or reserve-price options that are available to sellers on eBay.

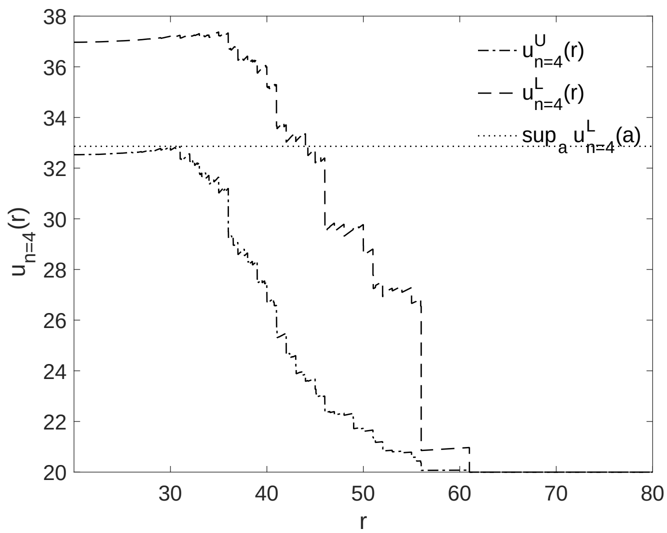
For each auction in the sample, I observe the number of active bidders (i.e., those who placed a bid) and the transaction price (i.e., the price paid by the winner of the auction). Throughout the analysis, I assume that the number of active bidders equals the number of bidders in the auction.[9] In the notation above, the number of active bidders equals $n + 1$, i.e., n legitimate bidders and 1 shill bidder. Table 1 displays some descriptive statistics. The table shows that the number of bidders range between 3 and 5, with 41 percent of the auctions featuring 5 bidders. The average transaction price across all auctions is \$36.07 (USD), and the average transaction price increases by between 55 and 85 cents with each additional bidder.

Figure 3 (panel A) displays the estimates for the identification region for the distribution of valuations using the results in Proposition 1. Given the sample size, I replace the min (max) functions in the estimator with the smooth weighted averages proposed by Haile and Tamer (2003), as discussing in Section 2. Specifically, I set $\rho_T = -\sqrt{673}$ and $\rho_T = \sqrt{673}$ for the lower and upper bounds, respectively, where 673 is the sample

[9]See footnote 5.



A. Distribution of valuations



B. Reserve price

Figure 3: Identification region for distribution of valuations of Armani Acqua di Gio perfume

Notes: These estimates are based on the results in Proposition 1 and Proposition 5. Confidence intervals in Panel A are one-sided and were computed using the bootstrap (2,500 replicates).

size. The figure also reports 95% confidence intervals for the bounds of the identification region. These confidence intervals are one-sided and were computed using the bootstrap with 2,500 replicates.[10]

As discussed in Proposition 2, even if there is no shill bidding in the data, the identification region for the distribution of valuations derived above will still contain the true value of F_v, provided that the symmetric independent private values assumptions hold. Making use of this result, I estimate the distribution of valuations under the null hypothesis of no shill bidding (or, more broadly, under the hypothesis of symmetric independent private values). Under the null hypothesis, the estimate of F_v must be contained in the region of identification of the distribution of valuations, as discussed in Corollary 1. If this condition fails to hold, we reject the null hypothesis of symmetric independent private values.

To estimate the distribution of valuations under the hypothesis of symmetric independent private values (i.e., all bids are legitimate), I make use of the identification results in Athey and Haile (2002). Following Athey and Haile (2007), I estimate the distribution of valuations separately for the subsamples with 3, 4, and 5 bidders, and then compute an optimally weighted average of these estimators to minimize variance of the estimated distribution of valuations. Under symmetric independent private values (i.e., no shill bidding), the transaction price equals the second highest valuation among all $n + 1$ bidders. Figure 3 (panel A) also displays the naively-estimated distribution of valuations assuming no shill bidding. As one can see in the figure, the estimated distribution of valuations fails to be fully contained in the estimated identification region for F_v, which suggests that the data rejects the symmetric independent private values assumptions. The rejection of the null hypothesis can be driven by shill bidding (or bidder asymmetries in general) or correlation in valuations. As discussed, this specification test does not rely on variation in the number of bidders, it instead depends on properties of order statistics.

Lastly, Figure 3 (panel B) plots the bounds for the seller's expected revenue when the seller values the good at $v_0 = \$20$ and expects an auction with 4 bidders. Using the results in Proposition 5, I find that the identification region for the optimal reserve price is $\hat{\Xi}_{n=4} = [20, 44]$, which provides some information to the seller, as the support of

[10]Haile and Tamer (2003) discuss consistency of bootstrapped confidence intervals in a similar setting.

the distribution includes values upwards of \$60. As discussed, the expected revenue of the seller under the optimal reserve price r^* must satisfy $u_n(r^*) \geq \sup_r u_n^L(r)$, implying that the optimal reserve price would guarantee the seller an expected revenue of at least \$32.86.

5 Conclusion

This paper studies identification in an English auction with shill bidding in an independent private values setting. I show that the distribution of valuations and the optimal reserve price are partially identified when shill bids may be present in the data. Partial identification stems from the fact that the winning bid no longer equals the second highest valuation among the legitimate buyers, as the shill bidder can win the auction. I show that the winning bid will be bounded between the second and first highest valuations among the legitimate buyers when a shill bidder is present. This observation can be used to bound the distribution of valuations and optimal reserve price. I apply these results on a sample of eBay auctions.

References

Andreyanov, Pasha and El Hadi Caoui (2020) "Secret reserve prices by uninformed sellers,"Technical report, Working paper, University of Toronto.

Athey, Susan and Philip A. Haile (2002) "Identification of Standard Auction Models," *Econometrica*, Vol. 70, No. 6, pp. 2107–2140.

Athey, Susan and Philip A Haile (2007) "Nonparametric approaches to auctions," *Handbook of econometrics*, Vol. 6, pp. 3847–3965.

Baldwin, L, R Marshall, and J.F. Richard (1997) "Bidder collusion at forest service timber sales," *Journal of Political Economy*, Vol. 105, No. 4, pp. 657–699.

Coey, Dominic, Bradley Larsen, and Kane Sweeney (2019) "The bidder exclusion effect," *The RAND Journal of Economics*, Vol. 50, No. 1, pp. 93–120.

Einav, Liran, Chiara Farronato, Jonathan Levin, and Neel Sundaresan (2018) "Auctions versus posted prices in online markets," *Journal of Political Economy*, Vol. 126, No. 1, pp. 178–215.

Feinstein, J., M. Block, and F. Nold (1985) "Asymmetric information and collusive behavior in auction markets," *American Economic Review*, Vol. 75, No. 3, pp. 441–460.

Graham, Daniel A., Robert C. Marshall, and Jean-Francois Richard (1990) "Phantom Bidding Against Heterogeneous Bidders," *Economic Letters*, Vol. 32, pp. 13–17.

Haile, Philip A. and Elie Tamer (2003) "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, Vol. 111, No. 1, pp. 1–51.

Hasker, Kevin and Robin Sickles (2010) "eBay in the Economic Literature: Analysis of an Auction Marketplace," *Review of Industrial Organization*, Vol. 37, pp. 3–42.

Hendricks, Ken and Robert H Porter (2007) "An empirical perspective on auctions," *Handbook of Industrial Organization*, Vol. 3, pp. 2073–2143.

Kauffman, Robert J. and Charles A. Wood (2005) "The effects of shilling on final bid prices in online auctions," *Electronic Commerce Research and Applications*, Vol. 4, pp. 21–34.

Manski, Charles F. and Elie Tamer (2002) "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, Vol. 70, No. 2, pp. 519–546.

Marmer, Vadim, Artyom A Shneyerov, and Uma Kaplan (2016) "Identifying collusion in english auctions," *Available at SSRN 2738789*.

OFT (2007) *Internet Shopping. An OFT market study.*: `http://www.oft.gov.uk/shared_oft/reports/consumer_protection/oft921.pdf`.

Porter, Robert and D. Zona (1993) "Detection of bid rigging in procurement auctions," *Journal of Political Economy*, Vol. 101, No. 3, pp. 518–538.

——— (1999) "Ohio school milk markets: An analysis of bidding," *RAND Journal of Economics*, Vol. 30, No. 2, pp. 263–288.

Riley, John G and William F Samuelson (1981) “Optimal auctions,” *The American Economic Review*, Vol. 71, No. 3, pp. 381–392.

Song, Unjy (2004) “Nonparametric estimation of an eBay auction model with an unknown number of bidders.”

Tang, Xun (2011) “Bounds on revenue distributions in counterfactual auctions with reserve prices,” *The RAND Journal of Economics*, Vol. 42, No. 1, pp. 175–203.

Wang, Wenli, Zoltan Hidvégi, and Andrew B. Whinston (2001) “Shill Bidding in English Auctions,” *Working Paper*.

Appendix A: Omitted proofs

Proof of Proposition 2

First, I show that $F_v(t) = \phi_2^{-1}(F_{w,(n+1)}(t)|n+1) \leq \phi_1^{-1}(F_{w,(n+1)}(t)|n) \equiv F_{vU}(t)$ for every t. From equations (2) and (3), we have that,

$$\begin{aligned} F_{w,(n+1)}(t) &= F_{vU}(t)^n \\ F_{w,(n+1)}(t) &= (n+1)F_v(t)^n - nF_v(t)^{n+1}. \end{aligned}$$

By setting these equal and rearranging, we obtain

$$F_{vU}(t)^n - F_v(t)^n = n[F_v^n(t) - F_v(t)^{n+1}] \geq 0.$$

The inequality, that follows from the fact that $F_v(t) \in [0,1]$, implies the result.

Lastly, I show that $F_v(t) = \phi_2^{-1}(F_{w,(n+1)}(t)|n+1) \geq \phi_2^{-1}(F_{w,(n+1)}(t)|n) \equiv F_{vL}(t)$ for every t. Again, from equation (3), we have that,

$$\begin{aligned} F_{w,(n+1)}(t) &= n(F_{vL}(t)^{n-1} - F_{vL}(t)^n) + F_{vL}(t)^n \\ F_{w,(n+1)}(t) &= nF_v(t)(F_v(t)^{n-1} - F_v(t)^n) + F_v(t)^n. \end{aligned}$$

By setting these expressions equal, and by using the fact that $F_v(t) \in [0,1]$, we obtain the following inequality

$$\begin{aligned} n(F_{vL}(t)^{n-1} - F_{vL}(t)^n) + F_{vL}(t)^n &= nF_v(t)(F_v(t)^{n-1} - F_v(t)^n) + F_v(t)^n \\ &\leq n(F_v(t)^{n-1} - F_v(t)^n) + F_v(t)^n. \end{aligned}$$

Note that the inequality can be rewritten as

$$\phi_2(F_{vL}(t)|n) \leq \phi_2(F_v(t)|n).$$

Since $\phi_2(\cdot|n)$ is a strictly increasing function, the result follows.

Proof of Proposition 3

The result follows from the observation that the highest valuation, $V_{1:n}$, among the n legitimate bidders, is bounded below by

$$V_{1:n} \geq B_{2:(n+1)},$$

since it can always be that the shill bidder is the highest bidder. This implies that

$$F_{1:n}(t) \leq G_{2:(n+1)}(t),\ \forall t, \forall (n+1).$$

By applying the inverse to the monotonic increasing function $\phi_1(\cdot|n)$ on both sides of the inequality, we obtain

$$F_v(t) \leq \phi_1^{-1}(G_{2:(n+1)}(t)|n),\ \forall t, \forall (n+1).$$

In general, the kth highest valuation ($k \in \{2, ..., n\}$), $V_{k:n}$, among the n legitimate bidders, is bounded by

$$B_{(k+1):(n+1)} \leq V_{k:n} \leq B_{k:(n+1)},$$

since it can always be that the shill bidder bids above or below $v_{k:n}$. This implies that

$$G_{k:(n+1)}(t) \leq F_{k:n}(t) \leq G_{(k+1):(n+1)}(t),\ \forall t, \forall (n+1).$$

By applying the inverse to the monotonic increasing function $\phi_k(\cdot|n)$ on the inequality, we obtain

$$\phi_k^{-1}(G_{k:(n+1)}(t)|n) \leq F_v(t) \leq \phi_k^{-1}(G_{(k+1):(n+1)}(t)|n),\ \forall t, \forall (n+1).$$

The result follows by intersecting all these inequalities, and over all $(n+1) \in \Omega$, and also with the identification region in Proposition 1.

Proof of Proposition 4

For part a), consider $L_{w,T}(t)$. We first have that by the Glivenko-Cantelli theorem,

$$G_{1:(n+1),T}(t) = \frac{1}{T_{n+1}} \sum_{i=1}^{T} 1\{m_i = (n+1); b^i_{1:(n+1)} \leq t\} \overset{a.s.}{\to} G_{1:(n+1)}(t) \equiv F_w(t)$$

uniformly in t, for all $(n+1) \in \Omega$. Since $\phi_2^{-1} : [0,1] \to [0,1]$ is a uniformly continuous function for all n, it follows from Lemma 2 that

$$\phi_2^{-1}(G_{1:(n+1),T}(t)|n) \overset{a.s.}{\to} \phi_2^{-1}(G_{1:(n+1)}(t)|n)$$

uniformly in t, for all $(n+1) \in \Omega$. Since the max function is continuous, it follows from the continuous mapping theorem that

$$L_{w,T}(t) \overset{a.s.}{\to} L_w(t),\ \forall t.$$

Finally, that the convergence of $L_{w,T}(t)$ to $L_w(t)$ is a.s. uniformly in t, follows from the following inequality

$$\sup_t |L_{w,T}(t) - L_w(t)| \leq \sum_n \sup_t |\phi_2^{-1}(G_{1:(n+1),T}(t)|n) - \phi_2^{-1}(G_{1:(n+1)}(t)|n)|.$$

The rest of the proof follows by applying analogous arguments.

Proof of Proposition 5

Fix $n \in \aleph$. Define

$$u_n(r|v_0) = v_0 F_v(r)^n + n \int_r^{\bar{v}} (F_v(v) + vF_v'(v) - 1) F_v^{n-1}(v) dv,$$

where $F_v(\cdot)$ is the true but unobserved distribution of valuations, and take

$$r_n^* \in \arg\max_r u_n(r|v_0).$$

It is true that

$$u_n(r_n^*) \geq \sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a) \tag{9}$$

$$u_n^U(r_n^*) \geq u_n(r_n^*) \tag{10}$$

since $u_n^U(t) \geq u_n(t) \geq u_n^L(t), \forall t$.

Suppose $r_n^* \notin H[r^*]$. That implies, in particular, that $r_n^* \notin \left\{r : u_n^H(r) \geq \sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a)\right\}$. If $r_n^* \notin \{r : u_n^U(r) \geq u_n^L(r_n^L)\}$, then

$$\sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a) > u_n^U(r_n^*).$$

But then by making use of (9) and (10), we reach the following contradiction

$$\sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a) > u_n^U(r_n^*) \geq \sup_{a \in [\underline{v}, \bar{v}]} u_n^L(a).$$

Proof of Proposition 6

Part a) follows from Lemma 4 in Appendix A. Part b) follows from Proposition 5b in Manski and Tamer (2002).

Appendix B: Additional results

Lemma 2 *Take a sequence of functions $\{g_T(\omega,\theta)\}$, $g_T : X \to Y$, that converges to $g(\theta)$ a.s. uniformly in $\theta \in \Theta$, that is,*

$$\Pr\left[\lim_{T\to\infty}\sup_{\theta\in\Theta}|g_n(\theta) - g(\theta)| = 0\right] = 1.$$

Take a uniformly continuous function $\psi : Y \to Y$. Then $\{\psi(g_T(\omega,\theta))\}$ converges to $\psi(g(\theta))$ a.s. uniformly in $\theta \in \Theta$.

Proof.

Fix any $\varepsilon > 0$. By uniform continuity of ψ, $\exists\delta > 0$ such that for any $x, y \in X$, $|x - y| < \delta$ implies $|\psi(x) - \psi(y)| < \varepsilon$.

By convergence a.s. uniformly of g_T,

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}|g_T(\theta) - g(\theta)| = 0 \quad \text{a.e. },$$

that is, $\exists\, T_\delta$ such that $\forall\, m \geq T_\delta$

$$\sup_{\theta}|g_m(\theta) - g(\theta)| < \delta \quad \text{a.e. }.$$

By uniform continuity of ψ, we conclude that $\forall m \geq T_\delta$

$$\sup_{\theta}|\psi(g_m(\theta)) - \psi(g(\theta))| < \varepsilon \quad \text{a.e. }.$$

Since this holds for any $\varepsilon > 0$,

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}|g_T(\theta) - g(\theta)| = 0 \quad \text{a.e.} \quad \Rightarrow \lim_{T\to\infty}\sup_{\theta\in\Theta}|\psi(g_T(\theta)) - \psi(g(\theta))| = 0 \quad \text{a.e. }.$$

The result follows since

$$1 = \Pr\left[\lim_{T\to\infty}\sup_{\theta\in\Theta}|g_T(\theta) - g(\theta)| = 0\right] \leq \Pr\left[\lim_{T\to\infty}\sup_{\theta\in\Theta}|\psi(g_T(\theta)) - \psi(g(\theta))| = 0\right].$$

■

Lemma 3 $u_{T,n}(r) \overset{a.s.}{\to} u_n(r)$ *uniformly in r.*

Proof.

Note that

$$
\begin{aligned}
\sup_r |u_{T,n}(r) - u_n(r)| &= \sup_r \Big|(v_0 - r)(F_T(r)^n - F(r)^n) \\
&\quad - \int_r^{\bar{v}} \left(F_T(v)^{n-1}[n(1-F_T(v)) + F_T(v)] - F(v)^{n-1}[n(1-F(v)) + F(v)]\right) dv\Big| \\
&\leq |K_1| \cdot \sup_r |F_T(r)^n - F(r)^n| \\
&\quad + \sup_r \left|\int_r^{\bar{v}} \left(F_T(v)^{n-1}[n(1-F_T(v)) + F_T(v)] - F(v)^{n-1}[n(1-F(v)) + F(v)]\right) dv\right| \\
&\leq |K_1| \cdot \sup_r |F_T(r)^n - F(r)^n| \\
&\quad + |K_2| \cdot \sup_v \left|F_T(v)^{n-1}[n(1-F_T(v)) + F_T(v)] - F(v)^{n-1}[n(1-F(v)) + F(v)]\right| \\
&= |K_1| \cdot \sup_r |\psi_1(F_T(r)) - \psi_1(F(r))| + |K_2| \cdot \sup_v |\psi_2(F_T(r)) - \psi_2(F(r))|,
\end{aligned}
$$

where K_1 and K_2 are constants, and $\psi_1 : [0,1] \to [0,1]$ and $\psi_2 : [0,1] \to [0,1]$ are uniformly continuous functions. Since $F_T(x) \stackrel{a.s.}{\to} F(x)$ uniformly in x, the result follows from Lemma 2. ■

Lemma 4 *$Q_T(t) \stackrel{a.s.}{\to} Q(t)$ uniformly in t.*

Proof.

Note that

$$
\begin{aligned}
\sup_t |Q_{n,T}(t) - Q_n(t)| &= \sup_t \Big|1\{\sup_a u^L_{T,n}(a) - u^H_{T,n}(t) > 0\}(\sup_a u^L_{T,n}(a) - u^H_{T,n}(t)) \\
&\quad -1\{\sup_a u^L_n(a) - u^H_n(t) > 0\}(\sup_a u^L_n(a) - u^H_n(t))\Big| \\
&\leq \sup_t \left|(\sup_a u^L_{T,n}(a) - u^H_{T,n}(t)) - (\sup_a u^L_n(a) - u^H_n(t))\right| \\
&\leq \left|\sup_a u^L_{T,n}(a) - \sup_a u^L_n(a)\right| + \sup_t \left|u^H_n(t) - u^H_{T,n}(t)\right| \\
&\leq \sup_a \left|u^L_{T,n}(a) - u^L_n(a)\right| + \sup_t \left|u^H_n(t) - u^H_{T,n}(t)\right|.
\end{aligned}
$$

Since $u_{T,n}(r) \stackrel{a.s.}{\to} u_n(r)$ uniformly in r, the result follows from Lemma 2. ■