INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2007-24

The Econometrics of Auctions with Asymmetric Anonymous Bidders*

L. LAMY [1]

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

---

* I would like to thank Philippe Février and Bernard Sanalié for very helpful discussions.
[1] Laboratoire d'Economie Industrielle, CREST-INSEE, 28 Rue des Saints-Pères, 75007 PARIS, France. laurent.lamy@ensae.fr

The Econometrics of Auctions with Asymmetric Anonymous Bidders *

Laurent Lamy[†]

*I would like to thank Philippe Février for very helpful discussions.

[†]Laboratoire d'Economie Industrielle, CREST-INSEE, 28 rue des Saints-Pères 75007 Paris. e-mail: laurent.lamy@ensae.fr

Abstract

We consider standard auction models when bidders' identities are not observed by the econometrician. First, we adapt the definition of identifiability to a framework with anonymous bids and we explore the extent to which anonymity reduces the possibility to identify private value auction models. Second, in the asymmetric independent private value model which is nonparametrically identified, we adapt Guerre, Perrigne and Vuong [10]'s two-stage estimation procedure. Our multi-stage kernel-based estimator achieves the optimal uniform convergence rate when identities are observed.

Keywords: Auctions, asymmetric, nonparametric identification, nonparametric estimation, anonymous bids, uniform convergence rate
JEL classification: D44, C14

Abstract

Nous considérons les modèles classiques d'enchères lorsque les identités des enchérisseurs ne sont pas observées par l'économètre. D'une part, nous adaptons la notion d'identifiabilité à un cadre avec des enchères anonymes et nous explorons dans quelle mesure l'anonymat réduit les possibilités d'identification des modèles à valeurs privées. D'autre part, dans le modèle asymétrique à valeurs privées et indépendentes, qui est identifiable nonparamétriquement, nous adaptons la procedure dévelopée par Guerre, Perrigne and Vuong [10]. Notre procedure d'estimation en plusieurs étapes est optimale au regard du taux de convergence suivant la norme uniforme.

Mots-clés: Enchères, Asymétries, Identification nonparamétrique, Estimation nonparamétrique, anonymat, taux de convergence uniforme
Classification JEL: D44, C14

1 Introduction

Motivated by the fact that the identities of the bidders are lacking to the econometrician in some auction data either because this information is confidential or have been lost, or because submissions are structurally anonymous as in internet auctions, we consider a setup where bidders' identities are not observed by the econometrician.[1] At first glance, anonymity reduces considerably the scope of the economic analysis and invites the econometrician to assume that bidders are ex ante symmetric.[2] On the other hand, the presence of asymmetries has been the key determinant of many empirical studies of auction data. In Porter and Zona [23, 24] and Pesendorfer [22], the bidding behavior of alleged cartel participants is compared to the ones of non-cartel bidders through reduced form approaches. In Hendricks and Porter [11], neighbor firms are shown to be better informed in auctions for drainage leases. The aim of this paper is to lay the foundations of the econometric of auctions under anonymous data and to show how we can deal with asymmetric models. We adopt the so-called structural approach (see Paarsch and Hong [21]) and focus on the private value single-unit auction model.

First, we adapt the definition of identifiability to a framework with anonymous bids by requiring the unique characterization of bidders' primitives up to a permutation of bidders' identities. Then, in the spirit of Laffont and Vuong [13] we explore the extent to which anonymity reduces the possibility to identify private value models in standard auctions with risk neutral buyers.[3] We show in Proposition 3.1 that anonymity prevents the identification of the asymmetric affiliated private value model, contrary to Campo, Perrigne and Vuong [7]'s analysis when bidders' identities are observed by the econometrician. When the identities of the bidders are not observed, the method that is currently implemented is to assume symmetry as an identifying restriction and to develop Guerre, Perrigne and Vuong [10]'s nonparametric methodology (henceforth GPV), which reaches the best rate of uniform convergence for the symmetric independent private value model. However, the validity of this method relies on the assumption that bidders are symmetric, an assumption that can not be rejected on any testable restriction without further restrictions if bids are anonymous. Furthermore, for auction models that explicitly involve asymmetry, e.g. models with collusion, with shill bidding or if the underlying market is intrinsically asymmetric, this

[1]The same motivation is the starting point of Yokoo et al. [30]'s analysis of combinatorial auctions when bidders have the possibility to submit false-name bids.

[2]See Song [27] and Sailer [26] for eBay auction models with symmetric bidders. Thus those models exclude any shill bidding activity from the seller, a pervasive phenomenon that is analyzed in Lamy [14].

[3]Risk aversion adds new caveats in the first price auction. See Campo, Guerre, Perrigne and Vuong [6].

identification route is not appropriate. We propose another identification route. We show in Proposition 3.1 that the asymmetric independent private value (IPV) model is identified. One crucial step in the resolution of this inverse problem is to recover the underlying cumulative distribution functions (CDFs) $(F_{\mathbf{B}_{\mathbf{i}}^*})_{i=1,\cdots,N}$ of each buyers' bids from the CDFs $(F_{\mathbf{B}_p})_{p=1,\cdots,N}$ of the order statistics of the anonymous bids. By exploiting independence, the vector of the N bidders CDFs $(F_{\mathbf{B}_{\mathbf{i}}^*})_{i=1,\cdots,N}$ corresponds to the roots of a polynomial of degree N whose coefficients are linear combinations of the CDFs $(F_{\mathbf{B}_p})_{p=1,\cdots,N}$.

Second, we propose a multi-stage kernel-based estimation procedure to recover the underlying distributions of bidders' private values. We mainly adapt GPV's two-stage estimation procedure. We establish the uniform consistency of our estimator and show that it attains the best uniform convergence rate for estimating the latent density of private values from observed bids. Indeed, we obtain the same rate as with nonanonymous bids. Our estimation procedure also fits the setup where the econometrician may benefit from some additional information as the identity of the winner, e.g. in Li and Perrigne [17], or the identities of the second-highest and highest bidders, e.g. in Baldwin et al [3]. In those latter cases, we know from Athey and Haile [2] that the asymmetric IPV model is identified. Nevertheless, in this framework, the existing nonparametric methodology from GPV may not perform very well in small data sets because it only uses the highest bidding statistics. In particular, in the second stage of GPV's estimation procedure, the pseudo-values are computed only for those bids that are not anonymous. On the contrary, our estimation procedure uses the complete vector of bids at both stages. In particular, we obtain for each bid a pseudo private value according to each possible identities of the bidder. Then, to estimate the distribution of private values, we should estimate for each bid the probability that it comes from a given bidder.

In a nutshell, we face two identification routes with anonymous bids: either to assume symmetry and to apply GPV's method allowing for correlated signals as in Li, Perrigne and Vuong [18] or to assume independence but not symmetric and to apply ours. Furthermore, with partially anonymous data, our estimation procedure may be more suitable for small data sets since it exploits all bids.

The paper is organized as follows. In Section 2, we introduce the model and the definition of identification under anonymity. In Section 3, we consider nonparametric identification of private value models under anonymity. In section 4, for the asymmetric independent private value model which is identified, we propose a multi-stage kernel-based estimator that corresponds to the natural extension of GPV's procedure. We establish its asymptotic properties allowing for heterogeneity across auctions and variations in the set of participants. In section 5, we conclude by indicating some future lines of research. Two Appendices contain the proofs of our results.

2 The Model

Consider an auction of a single indivisible good with $n \geq 2$ risk-neutral bidders. We consider the first price and second price sealed-bid auctions with no reserve price and when all bids are collected by the econometrician. We mention in section 5 how to extend our methodology with bidding reserve prices and with incomplete sets of bids. Nevertheless, if the econometrician can observe the amounts submitted by all bidders, we assume that bids are anonymous, i.e. she can not observe their corresponding identities. Hence, she observes the ordered vector of bids $B = (B_1, \cdots, B_p, \cdots, B_n)$, where B_p denotes the pth order statistic of the vector of bids B. But she does not observe $B^* = (B_1^*, \cdots, B_i^*, \cdots, B_n^*)$, where B_i^* denotes the amount submitted by bidder i. Subsequently, we use the indices i, j for bidders' identities and p, r for bidding order statistics.

We consider the private value paradigm: each participant $i = 1, \cdots, n$ is assumed to have a private value x_i for the auctioned object. Hence, bidder i would receive utility $x_i - p$ from winning the object at price p. In the first price and second price auctions, the price p is equal to B_n and B_{n-1}, respectively. Let $F_{\mathbf{X_i}}(.)$ and $F_{\mathbf{X}}(.)$ denote the cumulative distribution functions of X_i and $X = (X_1, \cdots, X_n)$, respectively, which are assumed to be absolutely continuous with probability density functions (PDF) $f_{\mathbf{X_i}}(.)$ and $f_{\mathbf{X}}(.)$ and compact support $[\underline{x}, \overline{x}]$ and $[\underline{x}, \overline{x}]^n$, respectively.[4,5] Each bidder is privately informed about x_i, whereas the common distribution $F_{\mathbf{X}}(.)$ is assumed to be common knowledge among bidders. When we refer to models with *symmetric* bidders we assume that the joint distribution of $\mathbf{X}$ is exchangeable with respect to buyers' indices. On the other hand, when we treat models allowing *asymmetric* bidders we drop the exchangeability assumption. For a generic random variable $\mathbf{S}$ and a class of events $\mathbf{E}$, we denote respectively $F_{\mathbf{S}|\mathbf{E}}(.|e)$ and $f_{\mathbf{S}|\mathbf{E}}(.|e)$ the CDF and PDF of $\mathbf{S}$ conditionally on an event e in $\mathbf{E}$.

Our analysis falls into two classes of models:
Independent Private Values (IPV): $F_{\mathbf{X}}(x) = \prod_{i=1}^n F_{\mathbf{X_i}}(x_i)$.
Strictly Affiliated Private Value (APV): $\frac{\partial^2 \log f_{\mathbf{X}}}{\partial x_i \partial x_j} > 0$ for $i \neq j$

[4] Throughout, uppercase letters are used for distributions, while lowercase letters are used for densities. We also follow the standard notation by using an uppercase letter for a statistic and the corresponding lowercase letter for its realization.

[5] We restrict ourselves to the common-support case that guarantees that almost all bids are 'serious' bids, i.e. win with a strictly positive probability. Otherwise identification is obtained only for 'serious' types. See Lebrun [16] for the analysis of the first-price auction with different supports.

Assumption A 1 *The joint density $f_{\mathbf{X}}$ is bounded, atomless and strictly positive on $[\underline{x}, \overline{x}]^n$.*

We restrict attention to Bayesian Nash Equilibrium in weakly undominated pure strategies, denoted by $(\beta_1(.), \cdots, \beta_n(.))$, where $\beta_i(.)$ is the bidding function of bidder i. In the equilibrium of the second price auction, buyers are thus bidding their private value. Hence, the link between bids and private types is straightforward:

$$x_i = b_i \equiv \xi_i^{nd}(b_i, F_{\mathbf{B}}) \tag{1}$$

In the first price auction, under assumption (1), Athey [1] guarantees the existence of an increasing pure strategy equilibrium if private values are affiliated and thus in the IPV and APV models. The link between bids and types for each bidder i is made by a standard rewriting of the first order differential equation derived from bidder i's optimization program:

$$x_i = b_i + \frac{F_{\mathbf{B}^*_{-i}|\mathbf{B}^*_i}(b_i|b_i)}{f_{\mathbf{B}^*_{-i}|\mathbf{B}^*_i}(b_i|b_i)} \equiv \xi_i^{rst}(b_i, F_{\mathbf{B}}), \tag{2}$$

where, for bidder i, $\mathbf{B}^*_{-i}$ denotes the maximum of the bids from bidder i's opponents.

Following Laffont and Vuong [13], we extend the literature on identification of private value models to the case where bids are anonymous. On the one hand, if bidders' identities are observed, then identifiability corresponds to the condition that, if two possible underlying distributions $F_{\mathbf{X}}(.)$ and $F'_{\mathbf{X}}(.)$ of private signals lead to the same distribution of bids $F_{\mathbf{B}^*}(.)$, then it follows that $F_{\mathbf{X}}(.)$ and $F'_{\mathbf{X}}(.)$ are equal. On the other hand, the following definition introduces the notion of identifiability that makes sense under anonymity.

Definition 1 (Identifiability under anonymity) *Under anonymous bidding, an auction model is said to be identifiable if for two possible underlying distributions $F_{\mathbf{X}}(.)$ and $F'_{\mathbf{X}}(.)$ of private values leading to the same distribution of bids $F_{\mathbf{B}}(.)$, then it follows that $F_{\mathbf{X}}(.)$ and $F'_{\mathbf{X}}(.)$ are equal up to a permutation of the potential buyers, i.e. there exists a permutation $\pi : [1, n] \to [1, n]$ such that $F_{\mathbf{X}}(x_1, \cdots, x_n) = F'_{\mathbf{X}}(x_{\pi(1)}, \cdots, x_{\pi(n)})$ for almost any vector of types X.*

Our definition of identifiability corresponds to the possibility of recovering an anonymous joint distribution of buyers' private values. Note that this information is not sufficient with asymmetric PV models for the computation of the optimal mechanism à la Myerson [20] that requires the knowledge of bidders' identities. Nevertheless, it is sufficient for the computation of the optimal anonymous mechanism or the optimal reserve price in a standard auction.

3 Nonparametric Identification

Anonymity restricts the degree of information of the data and thus it can only reduce the identification possibilities. In particular we show that asymmetric affiliated private value models are not indentified on the contrary to Campo, Perrigne and Vuong [7]'s identification result in a framework where bidders' identities are observed. Nevertheless, we show in Proposition [3.1] that, for a complete set of bids, either symmetry or independence restores identification. The surprising result is that anonymity does not prevent the identification of asymmetric IPV models. Our proof is constructive as it gives $F_{\mathbf{X}}(.)$ as a function of $F_{\mathbf{B}}(.)$. The empirical counterparts of this construction will then be used in the section devoted to nonparametric estimation. The proof of this result is thus given in the body of the text. The solution of this inverse problem contains two steps. First we derive bidder's distribution from the distribution of B, the vector of the bidding order statistics. It is the innovative step which relies on the root-finding of a well chosen polynomial. The second step is the identification of bidders' signals and is well-known: it is straightforward in the second price auction, whereas the first price auction has been treated by GPV. Note that identification under anonymity can not be proved directly from local arguments as in Roehring [25] (see also Benkard and Berry [4]) as was the case with nonanonymous bids as in GPV. The reason is that anonymity breaks differentiability at some points: the function that maps the vector of bidders' private values X into the bidding order statistics B is not differentiable at any point x such that $\beta_i(x_i) = \beta_j(x_j)$ for $i \neq j$.

Proposition 3.1 *Under the full observation of any submitted bids and under anonymous bids, in the first price and second price auctions and for $n \geq 2$:*

- *The asymmetric APV model is not identified. For any distribution $F_{\mathbf{X}}(.)$ from the asymmetric APV model, there exists a continuum of local perturbations of $F_{\mathbf{X}}(.)$ that stay in the asymmetric APV model and that are observationally equivalent to $F_{\mathbf{X}}(.)$, i.e. that lead to the same distribution of bids.*
- *The symmetric APV model is identified.*
- *The asymmetric IPV model is identified.*

The second point is immediate since the identification result in Li, Perrigne and Vuong [18] does not rely on the observability of bidders' identities. For the first point, we can construct, as is done in the appendix, a continuum of local perturbations of the primitives that are observationally equivalent. For any IPV model, the local perturbations constructed in the proof of the first point of Proposition 3.1 break independence, which illustrates, in this framework, the more general point that any unordered (i.e. observable up to a permutation) vector of independent components is observationally equivalent to a model where the components are correlated. In other words, the econometrician has to assume independence in order to identify asymmetry. Independence can not be fully tested under anonymity.[6] Nevertheless, independence involves some testable restrictions under anonymity and some partial tests could be built. Such developments are left for further research.[7]

The rest of this section is devoted to the proof of the third point which will guide our estimation procedure. We observe the distributions $F_{\mathbf{B_p}}$ for any $p = 1, \cdots, n$. Independence implies exchangeability, then we can identify the CDF $F_{\mathbf{B}}^{(r:m)}(u)$, $r \leq m$, that corresponds to the rth order statistic among m bidders that would result by exogenous variation of the number of bidders, by recursive use of the formula (see Athey and Haile [2] p.2128)

$$\frac{m-r}{m} F_{\mathbf{B}}^{(r:m)}(u) + \frac{r}{m} F_{\mathbf{B}}^{(r+1:m)}(u) = F_{\mathbf{B}}^{(r:m-1)}(u), \quad \forall u, r, m, r \leq m-1, m \leq n. \tag{3}$$

Remark that the corresponding induction is initialized by noting that $F_{\mathbf{B}}^{(p:n)} = F_{\mathbf{B_p}}$. In particular, it implies the identification of the CDFs $F_{\mathbf{B}}^{(r:r)}$ for any $r \in [1, n]$. Indeed, the expression of $F_{\mathbf{B}}^{(r:r)}$ corresponds to a linear combination of the CDFs $F_{\mathbf{B_p}}$, for $p = 1, \cdots, n$. Moreover, independence allows us to express $F_{\mathbf{B}}^{(r:r)}(b)$ as a function of the distributions $F_{\mathbf{B_i^*}}(b), i = 1, \cdots, n$ in the following way.

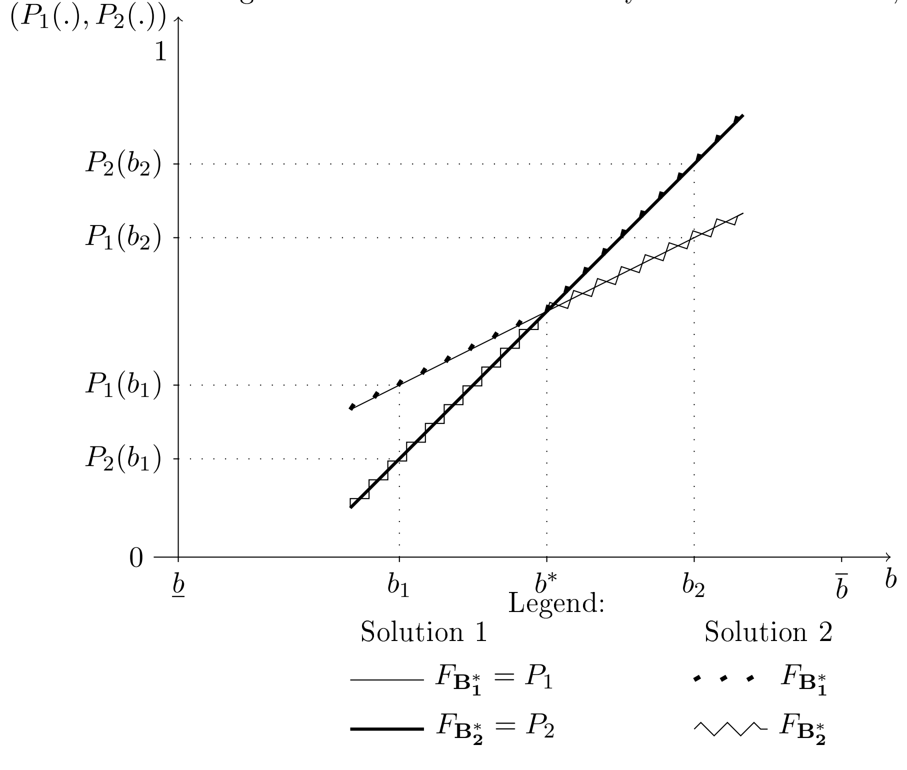
$$\begin{aligned}
F_{\mathbf{B}}^{(1:1)}(b) &= \frac{1}{n} \cdot \sum_{i=1}^{n} F_{\mathbf{B_i^*}}(b) \\
F_{\mathbf{B}}^{(2:2)}(b) &= \frac{1}{n(n-1)} \cdot \sum_{i_1, i_2, i_1 \neq i_2} F_{\mathbf{B_{i_1}^*}}(b) \cdot F_{\mathbf{B_{i_2}^*}}(b) \\
&\;\vdots \\
F_{\mathbf{B}}^{(r:r)}(b) &= \frac{1}{n(n-1)\cdots(n-r+1)} \cdot \sum_{i_1, \cdot, i_r, i_k \neq i_{k'}} \prod_{i_k \in \{i_1, \cdot, i_n\}} F_{\mathbf{B_{i_k}^*}}(b) \\
&\;\vdots \\
F_{\mathbf{B}}^{(n:n)}(b) &= \frac{1}{n!} \cdot \sum_{i_1, \cdot, i_n, i_k \neq i_{k'}} \prod_{i_k \in \{i_1, \cdot, i_n\}} F_{\mathbf{B_{i_k}^*}}(b)
\end{aligned} \tag{4}$$

[6] The nonparametric approaches in the literature that test whether the different components of a vector $X = (x_1, \cdots, x_m) \in \mathbb{R}^m$ are independent, e.g. the Blum, Kiefer and Rosenblatt [5] test, consider that the statistician observes ordered vectors, i.e. she can distinguish $X = (x_1, \cdots, x_m)$ from $X' = (x_{\pi(1)}, \cdots, x_{\pi(m)})$ where π is a permutation of bidders' indices. With respect to our setup, those tests are requiring nonanonymous bids.

[7] If the alternative to independence is not strict correlation but rather affiliation, then it is an open question whether independence can be tested.

The CDFs $(F_{\mathbf{B}_\mathbf{i}^*}(b))_{i=1,\cdots,n}$ in the above system of equations correspond exactly to the n roots of the polynomial of degree n: $u \to \sum_{i=0}^{n} a_i(b) \cdot (-1)^{n-i} \cdot u^i$, where $a_n(b) = 1$ and $a_i(b) = \frac{n(n-1)\cdots(i+1)}{(n-i)!} \cdot F_{\mathbf{B}}^{(n-i:n-i)}(b)$, for $i < n$. When b is fixed, such a solution is unique. By continuity of the coefficients of the polynomial as a function of b and since the roots of a polynomial depends continuously on its coefficients (see [29]), there exists a continuous function $b \to (P_1(b), \cdots, P_n(b))$ mapping the vector of solutions. What remains to show is the more restrictive condition that the CDFs $F_{\mathbf{B}_\mathbf{i}^*}(b)$, $i = 1, \cdots, n$, are unique up to a permutation. If the n roots of the above polynomial were always distinct for any b in the interior of the bidding support $(\underline{b}, \overline{b})$, then by continuity of $F_{\mathbf{B}^*}(.)$ the only candidate solution would be $(F_{\mathbf{B}_\mathbf{1}^*}(.), \cdots, F_{\mathbf{B}_\mathbf{n}^*}(.)) = (P_1(.), \cdots, P_n(.))$ (up to a permutation). On the contrary, if the maps $P_i(b)$ cross then the way we construct the continuous selection of the roots $(P_1(.), \cdots, P_n(.))$ is no more unique as it is illustrated in Figure 1 where two candidate solutions are depicted for $n = 2$ and when the roots cross at least once.

Figure 1: Identification of the asymmetric IPV model, $n = 2$



Indeed, the sole knowledge of the CDFs $F_{\mathbf{B}}^{(p:m)}$ for any p, m such that $p \leq m \leq n$ can not discriminate between these two possible solutions. Nevertheless, the knowledge of the joint distribution $F_{\mathbf{B}}$ of all order statistics selects a unique solution. For example, consider the case $n = 2$ and a point

b^* where $P_1(.)$ and $P_2(.)$ strictly cross as in Figure 1. We consider a point b_2 at the right of the intersection (respectively b_1 at the left of the intersection) such that the derivative of the upper root as a function of b, $P_2'(b_2)$ (resp. $P_1'(b_1)$), is strictly bigger (resp. strictly smaller) than the derivative of the lower root, $P_1'(b_2)$ (resp. $P_2'(b_1)$). Such a point exists in the right (resp. left) neighborhood of b^* since the intersection is strict. Then the two candidate solutions lead to different predictions in term of the joint density of the order statistics: $f_{\mathbf{B}}(b_1, b_2) = f_{\mathbf{B_1^*}}(b_1) \cdot f_{\mathbf{B_2^*}}(b_2) + f_{\mathbf{B_1^*}}(b_2) \cdot f_{\mathbf{B_2^*}}(b_1)$. The difference of the density $f_{\mathbf{B}}(b_1, b_2)$ between the two depicted solutions is equal to $(P_2'(b_2) - P_1'(b_2)) \cdot (P_2'(b_1) - P_1'(b_1)) \neq 0$. The argument remains valid for any number of bidders and also for more general intersections where the roots may coincide on an interval.

4 Nonparametric Estimation

In practice the auctioned objects can be heterogeneous and the number and the identities of the participants can vary across auctions. Consider a set of potential bidders, denoted $\mathbb{I}$, among which a subset I participates in an auction for a single and indivisible object. We assume that the number of participants, denoted by n_{I}, and their identities are common knowledge among bidders and are also observed by the econometrician.[8]

In this section, we adapt GPV's two step estimation procedure to recover the densities of bidders' private values in the first price auction.[9] Two caveats arise. First we can not directly estimate with kernel techniques the ratio $F_{\mathbf{B^*_{-i}}|\mathbf{b_i}}(.|.)/f_{\mathbf{B^*_{-i}}|\mathbf{b_i}}(.|.)$ since identities are not observed. An indirect procedure leading to the same uniform convergence rate in any inner closed subset of the bidding support is obtained. Second, if $F_{\mathbf{B^*_{-i}}|\mathbf{b_i}}(\cdot|\cdot)/f_{\mathbf{B^*_{-i}}|\mathbf{b_i}}(\cdot|\cdot)$ is suitably estimated, we can apply (2) to define pseudo private values in the first price auction. For each bid, a vector of pseudo private values, i.e. for each possible identities of the bidder. With anonymity, an additional step is needed: for a given vector of bid $b = (b_1, \cdots, b_p, \cdots, b_n)$, we have to estimate the probability that buyer i's bid b_i^* is equal to b_p for any $k \in [1, n]$. Then instead of a unique pseudo private value for a given bidder, we obtain a weighted vector of n pseudo private values that is used to estimated nonparametrically buyers' private values densities. When buyers' CDFs $F_{\mathbf{X_i}}(.|Z)$

[8] The observation of the identities of the participants by the econometrician may appear in contradiction with our paradigm of anonymous bids. If we could not observe participants identities, as on eBay, we can adapt our method if we are prepared to make specific assumptions about the identities of the fluctuating bidders (real bidder versus shill bidder). Anyway, in an asymmetric framework, the exogenous participation assumption that is often made for identification as in Athey and Haile [2] may not be suitable since the expected payoffs in the auction differ across bidders.

[9] See Flambard and Perrigne [9] for the the implementation of this procedure in the asymmetric private value model with nonanonymous bids.

have R bounded continuous derivatives and if d denotes the dimension of the (continuous) covariates Z, we obtain the same optimal uniform rate as in GPV: $(L/\log L)^{R/(2R+d+3)}$.

We also lead in parallel the analysis for the second price auction which is not straightforward as it was with nonanonymous bids. If bidders' identities were observed, private values would be directly observed by applying (1) and the optimal uniform rate of convergence for estimating private values densities is $(L/\log L)^{R/(2R+d+1)}$ (see Stone [28]). Under anonymous bids, our procedure for the second price auction reaches this optimal rate.

Denote Σ_{I} the set of the $n_{\mathrm{I}}!$ permutations between participants' identities and the order statistics of the bids. Such an assignment of the bids to the participants is denoted $\pi : \mathrm{I} \to [1, n_{\mathrm{I}}]$ where $\pi(i) = p$ means that the pth order statistic of the bids corresponds to bidder i, i.e. $b_i^* = b_p$. To cover both the case when bidders' identities remain fully anonymous with the common case when only the identity of the winner is disclosed, we consider the most general case when the econometrician may have some information linking some submitted bids with the identities of some participants. This information is modelled as a subset $\sigma_{\mathrm{I}} \subset \Sigma_{\mathrm{I}}$, e.g. $\sigma_{\mathrm{I}} = \Sigma_{\mathrm{I}}$ corresponds to the case when bids are fully anonymous. The opposite case when σ_{I} always is a singleton corresponds to nonanonymous bids, then GPV's procedure should be preferred. Our estimation procedure is flexible and imposes no restriction on the way σ_{I} varies across auctions.

4.1 Regularity Assumptions and Key Properties

Let Z_l denote the vector of relevant continuous characteristics for the lth auctioned object and I_l the set of participants in the lth auction. The vector (Z_l, I_l) is assumed to be common knowledge among bidders and is observed by the econometrician. Relative to our previous notation, in this section, we will work with conditional distributions and densities of private values and bids given (Z_l, I_l). E.g., $F_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(.|Z_l, I_l)$ denotes the CDF of bidder i's private value X_{il} for the lth auction. Thus (1) and (2) for the first and second price auction become respectively:

$$X_{il} = B_{il}^* + \frac{F_{\mathbf{B}^*_{-i}|\mathbf{B}^*_i,\mathbf{Z},\mathbf{I}}(B_{il}^*|B_{il}^*, Z_l, \mathrm{I}_l)}{f_{\mathbf{B}^*_{-i}|\mathbf{B}^*_i\mathbf{Z},\mathbf{I}}(B_{il}^*|B_{il}^*, Z_l, \mathrm{I}_l)}, \tag{5}$$

and

$$X_{il} = B_{il}^*. \tag{6}$$

The next assumptions concern the underlying generating process as well as the smoothness of the latent joint distribution of $(X_{il}, Z_l, \mathrm{I}_l)$ for any $i \in \mathrm{I}_l$.

Assumption A 2 (i) *The* $(d{+}1)$*-dimensional vectors* $(Z_l, \mathrm{I}_l), l = 1, 2, \cdots$, *are independently and identically distributed as* $F_{\mathbf{Z},\mathbf{I}}(.,.)$ *with density* $f_{\mathbf{Z},\mathbf{I}}(.,.)$.

(ii) *For each* l *the variables* $X_{il}, i \in \mathrm{I}_l$ *are independently distributed conditionally upon* (Z_l, I_l) *as* $F_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(.|.,.)$ *with density* $f_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(.|.,.)$*, for* $i \in \mathrm{I}_l$.

As in Campo et al. [6], we consider here that the support of buyers' private values does not depend on the (Z, I) to simplify the presentation, while the general case can be fully treated as in GPV. It implies that the lower bound of the support of buyers' bids does not depend on the variables I and Z. Throughout we denote by $S(*)$ and $S^o(*)$ the support of $*$ and its interior, respectively. Let $\mathcal{I} \subset \mathbb{I}$ be the set of possible values for I_l. Note that $\mathcal{I}$ is finite.

Assumption A 3 *For each bidder* $i \in I \subset \mathcal{I}$,

(i) $S(F_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}) = \{(x, z, I) : z \in [\underline{z}, \overline{z}], x \in [\underline{x}, \overline{x}], I \subset \mathcal{I}\}$*; with* $\underline{z} < \overline{z}$*;*

(ii) *for* $(x, z, I) \in S(F_{\mathbf{X_i},\mathbf{Z},\mathbf{I}})$*,* $f_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(x|z, I) \geq c_f > 0$*, and for* $(z, I) \in S(F_{\mathbf{Z},\mathbf{I}})$*,* $f_{\mathbf{Z},\mathbf{I}}(z, I) \geq c_f > 0$*;*

(iii) *for each* $I \subset \mathcal{I}$*,* $F_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(.|.,I)$ *and* $f_{\mathbf{Z},\mathbf{I}}(.,I)$ *admit up to* $R{+}1$ *continuous bounded partial derivatives on* $S(F_{\mathbf{X_i},\mathbf{Z},\mathbf{I}})$ *and* $S(F_{\mathbf{Z},\mathbf{I}})$*, with* $R \geq 1$.

The next assumption is not necessary for identification as established in Proposition 3.1 without heterogeneity across objects. Nevertheless, heterogeneity requires an additional structure to identify the model. Similar intersections as the one in Figure 1 when b varies may arise when the variable capturing heterogeneity Z varies. But the different solutions are observationally equivalent without some mild additional assumptions. Here to preserve identification, we make the (strong) assumption that bidding distributions can be ordered according to first order stochastic dominance. With two classes of bidders, Maskin and Riley [19] show that first order stochastic dominance for private values is sufficient for first order stochastic dominance for equilibrium bids.[10] Moreover, to simplify our estimation procedure, we also assume that the dominance is strict in the interior of the bidding support.

Assumption A 4 (Strict Stochastic Dominance) *The bid densities* $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}(.|z, I)$ *are strictly ordered according to first order stochastic dominance:*

[10] An alternative identification strategy with two classes of bidders is to make assumptions on the comparative statics of the bidding distribution according to Z. Another strategy would rely on the point that, generically, at an intersection, only one candidate solution is differentiable at this point.

$$F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}(b|z,I) > F_{\mathbf{B_{i+1}^*}|\mathbf{Z},\mathbf{I}}(b|z,I), \;\; if \;\; b \in S^0(f_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}})$$

for any $i \in I$ *and any* z, I*.*

A crucial step in deriving uniform rates of convergence in some inverse problem is to study the smoothness of the observables that is implied by the smoothness of the latent distributions of the primitives of the model. Here, relative to GPV, we do not observe the vector of bids B^* but only the vector of bidding order statistic B. Thus we are interested in the smoothness of the densities $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(.|.,I)$ for any $p \in I$. This is the purpose of the next proposition. It is the analog of proposition 1 in GPV which derives similar results for the bid densities $f_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}(.|.,.)$.

Proposition 4.1 *Given A3, the conditional distribution* $F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(.|.,I)$*,* $p \in I$ *and* $I \subset \mathbb{I}$*, satisfies for both the first and second price auctions (if not specified):*

(i) *its support* $S(F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}})$ *is such that* $S(F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}) = \{(b,z,I) : z \in [\underline{z},\overline{z}], b \in [\underline{b}(z,I,p),\overline{b}(z,I,p)], I \subset \mathcal{I}\}$ *with* $\overline{b}(z,I,p) > \underline{b}(z,I,p)$ *for any* I,p*. Moreover,* $(\underline{b}(.,I,p),\overline{b}(.,I,p))$ *admit up to* $R+1$ *continuous bounded derivatives on* $[\underline{z},\overline{z}]$ *for each* $I \subset \mathcal{I}$ *and* $p = 1,\cdots,n_I$*. We have* $\underline{b}(z,I,p) = \underline{x}$*. In the second price auction,* $\overline{b}(z,I,p) = \overline{x}$*. In the first price auction* $\overline{b}(z,I,n_I) = \overline{b}(z,I,n_I-1)$*.*

(ii) *for* $(b,z,I) \in \mathcal{C}(B_n)$*,* $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(b,z,I) \geq c_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}} > 0$*, where* $\mathcal{C}(B_n)$ *is a closed subset of* $S^0(F_{\mathbf{B_n}|\mathbf{Z},\mathbf{I}})$*;*

(iii) *for each* (I,p)*,* $p = 1,\cdots,n_I$*,* $F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(.|.,I)$ *admits up to* $R+1$ *continuous bounded partial derivatives on* $S(F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}) \setminus (\{\overline{b}(z,I,p)\}_{p=1,\cdots,n_I-1})$*;*

(iv) *in the first price auction, for each* (I,p)*,* $p = 1,\cdots,n_I$*, if* $\mathcal{C}(B_p)$ *is a closed subset of* $S^o(F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}) \setminus (\{\overline{b}(z,I,p)\}_{p=1,\cdots,n_I})$*, then* $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(.|.,I)$ *admits up to* $R+1$ *continuous bounded partial derivatives on* $\mathcal{C}(B_p)$*;*

(v) *in the second price auction, for each* (I,p)*,* $p = 1,\cdots,n_I$*,* $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(.|.,I)$ *admits up to* R *continuous bounded partial derivatives on* $S(F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}) \setminus (\{\overline{b}(z,I,p)\}_{p=1,\cdots,n_I-1})$*.*

Note that by comparing (iv) and (v), the bid densities in the first price auction are smoother than for the second price auction. Thus $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(.|.,I)$ can be estimated uniformly at a faster rate, namely $(L/\log L)^{(R+1)/(2R+d+3)}$, in the first price than in the second price auction, namely $(L/\log L)^{R/(2R+d+1)}$. In particular, the optimal bandwidths -that we specify later in assumption A6- are asymptotically smaller for the second price auction than for the first price auction. Nevertheless the optimal uniform convergence rate will be smaller in the first price auction than in the second price auction. This is due to the more indirect nature of the link between observables and latent distributions in the first price auction, see equation (5) versus (6).

Proposition 4.1 differs from the one appearing in GPV as irregularities of the CDF of the observed variables may appear in the interior of their support, more precisely at the upper bound of the bidding support of the (at most $n_I - 2$) bidders such that $\overline{b}(z, I, p) < \overline{b}(z, I, n)$. In the following, to alleviate notation, we make the simplifying assumption A5 that the bidding supports of all bidders coincide, i.e. $\overline{b}(z, I, p)$ does not depend on p. Our uniform consistency results extend provided that the neighborhood of the bidders' signals than make them bid $\overline{b}(z, I, p)$ are removed. In the same way as the support of bidders' private values is consistently estimated in GPV and that the neighborhoods of the lower and upper bounds of the support are removed with an suitable trimming, we can trim those inner neighborhoods.

Assumption A 5 (Common bidding support) *All bidders have the same bidding support: $\overline{b}(z, I, p)$ does not depend on p.*

4.2 Optimal Uniform Convergence Rate

In this section, we adopt a minmax approach to obtain bounds for the rate at which the latent density of private values can be estimated uniformly from observed bids. The next proposition gives an upper bound for the optimal uniform convergence rate for estimating $f_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(.|.,.)$ from observed (anonymous) bids. GPV derives the same bound for the symmetric IPV model and nonanonymous bids. Here we extend their result to the asymmetric IPV model. In the following, for a given density function f, denote by $||f||_r$ (resp. $||f||_{r,\mathcal{C}}$) the maximum of f and all its derivatives up to the rth order on $S(F)$ (resp. on $\mathcal{C}$).

Proposition 4.2 *Assume A2-A5 and $||f^o_{\mathbf{X},\mathbf{Z},\mathbf{I}}(x, z, I)||_R < M$. Let $\mathcal{C}(X)$ be an inner compact subset of $S(f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}})$ with nonempty interior. There exists a constant $\kappa > 0$ such that*

$$\lim_{\epsilon \to 0} \underline{\lim}_{L \to +\infty} \inf_{\widehat{f}_L} \sup_{f \in U_\epsilon(f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}})} Prob_f \left[\left(\frac{L}{\log L}\right)^{\frac{R}{(2R+d+3)}} \sup_{(x,z,I) \in \mathcal{C}(X)} ||\widehat{f}_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(x|z, I) - f_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(x|z, I)||_0 > \kappa \right] > 0$$

in the first price auction, and

$$\lim_{\epsilon \to 0} \underline{\lim}_{L \to +\infty} \inf_{\widehat{f}_L} \sup_{f \in U_\epsilon(f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}})} Prob_f \left[\left(\frac{L}{\log L}\right)^{\frac{R}{(2R+d+1)}} \sup_{(x,z,I) \in \mathcal{C}(X)} ||\widehat{f}_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(x|z, I) - f_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(x|z, I)||_0 > \kappa \right] > 0$$

in the second price auction, where the infimums are taken over all possible estimators $\widehat{f}_L$ of $f_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(.|.,.)$ based upon $(B_{pl}, Z_l, \mathrm{I}_l)$ for any $p = 1, \cdots, n_{I_l}$ and $l = 1, \cdots, L$ and where $U_\epsilon(f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}})$ is a neighborhood of $f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}}$ defined as

$$U_\epsilon(f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}}) \equiv \left\{ f; \sup_{(x,z,i) \in S(F^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}})} ||f(x, z, I) - f^o_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(x, z, I)||_0 < \epsilon, ||f(.,.,.)||_R < M \right\},$$

where $M > 0$.

The set of possible estimators based upon anonymous bids is tautologically smaller than those based upon $(B^*_{il}, Z_l, \mathrm{I}_l)$ for any $i \in I_l$ and $l = 1, \cdots, L$. Thus it is sufficient to prove the above proposition with this richer set of estimators. In this latter case, for the second price auction where observed bids correspond exactly to private values, the above result follows from Khas'minskii [12]. In the first price auction, the above proposition has been proved when the model is restricted to be symmetric among bidders by GPV who adapts Khas'minskii [12]'s arguments. It is intuitive that a faster local rate of uniform convergence is not available in the general case with asymmetric bidders. Nevertheless, due to the local nature of the above result, the argument is not tautologic. Indeed, since a general asymmetric model with n bidders involves n overlapped differential equations for bidders' distributions, the asymmetric structure may 'smooth' the link between observables and the latent private values. We show in the appendix how GPV's proof has to be adapted.

4.3 Definition of the Estimator

The purpose of this section is to adapt GPV's two step procedure to asymmetric auctions with anonymous bids. Using independence, (5) and (6) can be rewritten as

$$X_{il} = B^*_{il} + \psi_i(B^*_{il}, Z_l, \mathrm{I}_l), \tag{7}$$

where $\psi_i(.,.,.)$ is defined as

$$\psi_i(b, z, I) = \begin{cases} \left[\sum_{j \in I_l, j \neq i} \frac{f_{\mathbf{B}^*_j|\mathbf{Z},\mathrm{I}}(B^*_{il}|Z_l,\mathrm{I}_l)}{F_{\mathbf{B}^*_j|\mathbf{Z},\mathrm{I}}(B^*_{il}|Z_l,\mathrm{I}_l)}\right]^{-1}, & \text{in the first price auction} \\ 0, & \text{in the second price auction} \end{cases} \tag{8}$$

The first step in GPV's approach consists in estimating the maps $\psi_i(.,.,.)$. The main caveat is that we do not observe the variables B^*_{il} but only the order statistics B_{pl}. Thus we need to convert our estimations of the CDFs and PDFs of B_{pl}, that can be done with the standard kernel estimation techniques, into estimations for the CDFs and PDFs of B^*_{il}.

Using the observations $\{(B_{pl}, Z_l, \mathrm{I}_l); p \in \mathrm{I}_l, l = 1, \cdots, L\}$, our first step consists in estimating the CDFs and the PDFs of the pth ordered statistics of the bids for $p \in [1, n_I]$.

$$\widehat{F}_{\mathbf{B_p},\mathbf{Z},\mathrm{I}}(b, z, \mathrm{I}) = \min\left\{\frac{1}{Lh^d_{F_{\mathbf{B_p}|\mathbf{Z}}}} \sum_{l=1}^{L} \mathbf{1}(B_{pl} \leq b) K_{F_{\mathbf{B_p}|\mathbf{Z}}}\left(\frac{z - Z_l}{h_{F_{\mathbf{B_p}|\mathbf{Z}}}}\right) \mathbf{1}(\mathrm{I}_l = \mathrm{I}), 1\right\} \tag{9}$$

$$\widehat{f}_{\mathbf{B_p},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}) = \frac{1}{Lh^d_{f_{\mathbf{B_p}|\mathbf{Z}}}}\sum_{l=1}^{L}\mathbf{1}(B_{pl}\le b)K_{f_{\mathbf{B_p}|\mathbf{Z}}}\left(\frac{b-B_{pl}}{h_{f_{\mathbf{B_p}|\mathbf{Z}}}},\frac{z-Z_l}{h_{f_{\mathbf{B_p}|\mathbf{Z}}}}\right)\mathbf{1}(\mathrm{I}_l=\mathrm{I}) \quad (10)$$

Here $h_{F_{\mathbf{B_p}|\mathbf{Z}}}, h_{f_{\mathbf{B_p}|\mathbf{Z}}}$ are some bandwidths, and $K_{F_{\mathbf{B_p}|\mathbf{Z}}}(.)$ and $K_{F_{\mathbf{B_p}|\mathbf{Z}}}(.,.)$ are kernels with bounded supports. By recursive use of the empirical counterpart of the formula (3), we estimate $\widehat{F}^{(r:r)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)$ and $\widehat{f}^{(r:r)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)$ for $r=1,\cdots,n$, which respectively corresponds (up to a known multiplicative coefficient) to the coefficients and their derivatives of the polynomial whose roots is the vector of bidders' bidding distribution $\{F_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}\}_{1\le i\le n}$.

For $r\le m\le n$, we define $\widehat{F}^{(r:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)$ and $\widehat{f}^{(r:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)$ by recursive use of the formulas:

$$\frac{m-r}{m}\widehat{F}^{(r:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)+\frac{r}{m}\widehat{F}^{(r+1:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)=\widehat{F}^{(r:m-1)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I),\forall b,z,r\le m-1 \quad (11)$$

$$\frac{m-r}{m}\widehat{f}^{(r:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)+\frac{r}{m}\widehat{f}^{(r+1:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)=\widehat{f}^{(r:m-1)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I),\forall b,z,r\le m-1 \quad (12)$$

As a weighted sum of the estimators $\widehat{F}_{\mathbf{B_p},\mathbf{Z},\mathrm{I}}$ which are confined in the interval $[0,1]$, the estimators $\widehat{F}^{(r:m)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)$ are confined in the interval $[0,1]$.

Let $\Upsilon:[0,1]^n\to\mathbb{Z}^n$ be the function such that $(\omega_1,\cdots,\omega_n)=\Upsilon(a_0,\cdots,a_{n-1})$ (where $\omega_1\ge\cdots\ge\omega_n$) is the ordered vector of the roots (possibly complex number) counted with their order of multiplicity of the polynomial $Q(u)=u^n+\sum_{i=0}^{n-1}a_i\cdot(-1)^{n-i}u^i$, i.e. $Q(u)=\prod_{i=1}^{n}(u-\omega_i)$. Uherka and Sergott [29] show that Υ is continuous and hence uniformly continuous on the compact $[0,1]^n$.

In a second step, in view of (4), it would be natural to estimate the CDFs $\widehat{F}_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(.,.), i\in I$ by

$$(\widehat{F}_{\mathbf{B^*_{j_1}},\mathbf{Z},\mathbf{I}}(b,z,I),\cdots,\widehat{F}_{\mathbf{B^*_{j_{n_I}}},\mathbf{Z},\mathbf{I}}(b,z,I))=\mathcal{R}[\Upsilon(\widehat{F}^{(n:n)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I),\cdots,\widehat{F}^{(1:1)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I))], \quad (13)$$

where $\mathcal{R}[z]$ denotes the real part of the complex vector z and $I=(j_1,\cdots,j_{n_I})$, where $j_1<\cdots<j_{n_I}$.

The derivation of the polynomial relation with respect to b leads to:

$$\begin{aligned}\frac{\partial Q(u)}{\partial b} &= \sum_{i=0}^{n_I-1} f^{(n-i:n-i)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)\cdot(-1)^{n-i}\cdot u^i\\ &= -\sum_{i\in I}\prod_{j\in I,\ j\ne i}\Big(u-F_{\mathbf{B_j^*},\mathbf{Z},\mathbf{I}}(b,z,I)\Big)\cdot f_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(b,z,I),\forall u,b,z,I\end{aligned}$$

Thus under assumption (4) that there are no multiple roots, we have a natural estimator for bidders' densities.

$$\widehat{f}_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(b,z,I)=\frac{\sum_{k=0}^{n_I-1}\widehat{f}^{(n-k:n-k)}_{\mathbf{B},\mathbf{Z},\mathbf{I}}(b,z,I)\cdot(-1)^{n_I-k+1}\cdot\left[\widehat{F}_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(b,z,I)\right]^k}{\prod_{j\in I,\ j\ne i}\Big(\widehat{F}_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(b,z,I)-\widehat{F}_{\mathbf{B_j^*},\mathbf{Z},\mathbf{I}}(b,z,I)\Big)} \quad (14)$$

Note that we have assumed strict asymmetry to avoid singularity points in the estimation of $f_{\mathbf{B}_{\mathbf{i}}^*,\mathbf{Z},\mathbf{I}}$ in any closed subset of $S^o(F_{\mathbf{B}_{\mathbf{i}}^*,\mathbf{Z},\mathbf{I}})$. Now we have all the elements to estimate the function $\psi_i(.,.,.)$ in the first price auction.[11]

In view of (7) and similarly to GPV, it would be natural to construct pseudo private values for each order statistic $p = 1, \cdots, n_I$ and for each potential bidder $i \in I_l$:

$$\widetilde{X}_{ipl} = B_{pl}^* + \widetilde{\psi}_i(B_{pl}^*, Z_l, \mathrm{I}_l), \tag{15}$$

Unfortunately, as has been emphasized by GPV, the estimator of $\psi_i(.,.,.)$ in the first price auction is asymptotically biased at the boundaries of the support and trimming is recquired. The same trimming is also needed in the second price auction.

In this aim we first estimate the boundary of the support of the joint distribution of (B, Z, I), which is unknown. Since the support of (Z, I) can be assumed to be known, we focus on the estimation of the support $[\underline{b}(z, I), \overline{b}(z, I)]$ of the conditional distribution of B given (Z, I). By our simplifying assumption $\underline{b}(z, I)$ does not depend on (z, I) and is estimated by the minimum of all submitted bids. On the other hand, $\underline{b}(z, I)$ should be estimated as in GPV. Let $h_\delta > 0$. We consider the following partition of $\mathbb{R}^d$ with a generic hypercube of side h_δ:

$$\vartheta_{k_1,\cdots,k_d} = [k_1 h_\delta, (k_1+1)h_\delta) \times \cdots \times [k_d h_\delta, (k_d+1)h_\delta),$$

where $k_1, \cdots, k_d$ runs over $\mathbb{Z}^d$. This induces a partition of $[\underline{z}, \overline{z}]$. Given a set of participants I and a value z, the estimate of the upper boundary $\overline{b}(z, I)$ is the maximum of those bids for which $I_l = I$ and the corresponding value of X_l falls in the hypercube $\vartheta_{k_1,\cdots,k_d}(z)$ containing z. Formally, our estimators for the lower and upper boundaries are respectively given by:

$$\widehat{\overline{b}}(z, I) = \sup\{B_{n_I l}, l = 1, \cdots, L; X_l \in \vartheta_{k_1,\cdots,k_d}(z), I_l = I\} \tag{16}$$

$$\widehat{\underline{b}} = \inf\{B_{1l}, l = 1, \cdots, L\} \tag{17}$$

Our estimator for $S(F_{\mathbf{B_p},\mathbf{Z},\mathrm{I}})$ is $\widehat{S}(F_{\mathbf{B_p},\mathbf{Z},\mathrm{I}}) = \{(b, z, I) : b \in [\widehat{\underline{b}}, \widehat{\overline{b}}(z, I)], z \in [\underline{z}, \overline{z}], I \in \mathcal{I}\}$.

We now turn to the trimming. It is well known that kernel estimators are asymptotically biased at the boundaries of the support. Following GPV, we have to trim out observations that are close to the boundaries of the support. Because $\underline{b} \le \widehat{\underline{b}} \le \widehat{\overline{b}}(z, I) \le \overline{b}$, $\widehat{f}_{\mathbf{B}_p,Z,I}(B_{pl}, Z_l, I)(.,.,.)$ and thus $\widehat{f}_{\mathbf{B}_{\mathbf{j}}^*,\mathbf{Z},\mathbf{I}}(.,.,.)$ are asymptotically unbiased on $[\widehat{\underline{b}} + \frac{\rho_{f_{\mathbf{B_p}|\mathbf{Z}}} \cdot h_{f_{\mathbf{B_p}|\mathbf{Z}}}}{2}, \widehat{\overline{b}}(z, I) - \frac{\rho_{f_{\mathbf{B_p}|\mathbf{Z}}} \cdot h_{f_{\mathbf{B_p}|\mathbf{Z}}}}{2}]$. This leads to defining the sample of pseudo private values $\{\widehat{X}_{ipl}, i \in I_l; p =$

[11] Our procedure easily adapts if the multiplicity of the root $F_{\mathbf{B}_{\mathbf{j}}^*,\mathbf{Z},\mathbf{I}}$ is $k > 1$ by considering the polynomial $u \to \frac{\partial^k Q(u)}{\partial b(\partial u)^{k-1}}$ evaluated at $u = F_{\mathbf{B}_{\mathbf{j}}^*,\mathbf{Z},\mathbf{I}}(b, z, I)$.

$1, \cdots, n; l = 1, \cdots, L\}$ where $\widehat{X}_{ipl}$, the estimate of the private value of bidder i would it be the bidder of the pth order statistic of the vector of bids B_l, is defined by

$$\widehat{X}_{ipl} = \begin{cases} B_{pl} + & \left[\sum_{j \neq i} \frac{\widehat{f}_{\mathbf{B}_j^*}(B_{pl}, Z_l)}{\widehat{F}_{\mathbf{B}_j^*}(B_{pl}, Z_l)}\right]^{-1} \\ & \text{if } \widehat{\underline{b}} + h_{f_{\mathbf{B_p|Z}}} \leq B_{pl} \leq \widehat{\overline{b}}(Z_l, I_l), \\ +\infty & \text{otherwise} \end{cases} \tag{18}$$

in the first price auction and

$$\widehat{X}_{ipl} = \begin{cases} B_{pl} & \text{if } \widehat{\underline{b}} + h_{f_{\mathbf{B_p|Z}}} \leq B_{pl} \leq \widehat{\overline{b}}(Z_l, I_l), \\ +\infty & \text{otherwise} \end{cases} \tag{19}$$

in the second price auction.

Contrary to GPV, we should not use directly this pseudo sample of private values in a standard kernel estimation to estimate $f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x, z, I)$. Each pseudo values do not have the same weighted since for a given order statistic B_p the probability that it results from a given bidder i depends on the identity of this bidder. Thus we have to estimate the corresponding probability weights. Under anonymity, there are at most $n_I!$ vectors of private values that can rationalize a given vector of bids $(B_{1l}, \cdots, B_{n_I l})$. Denote by $\widetilde{\pi} \in \Sigma_I$ the permutation that matches a given vector of bidding order statistics $(B_{1l}, \cdots, B_{n_I l})$ with the unobserved vector of bids $(B^*_{1l}, \cdots, B^*_{nl})$.

The following expression gives the theoretical probability, denoted by $Prob(\widetilde{\pi} = \pi | (b_1, \cdots, b_{n_I}, z, I))$, that the assignment of bidders to the observed order statistics corresponds to the permutation π:

$$Prob(\widetilde{\pi} = \pi | (b_1, \cdots, b_{n_I}, z, I)) = \frac{\prod_{i \in I} f_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(b_{\pi(i)} | z, I)}{\sum_{\pi' \in \sigma_I} \prod_{i \in I} f_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(b_{\pi'(i)} | z, I)} \cdot \mathbf{1}\{\pi \in \sigma_I\}. \tag{20}$$

Note that we use the information set σ_I to refine our beliefs on $\widetilde{\pi}$. Then the probability, denoted by P_{ip}, that the pth order statistic results from bidder i equals to the sum of the above probabilities for all the permutations that assign i to the pth order statistic, i.e.

$$P_{ip} = \sum_{\pi \in \sigma_I \text{ s.t. } \pi(i) = p} Prob(\widetilde{\pi} = \pi | (b_1, \cdots, b_n, z, I)). \tag{21}$$

Their empirical counterparts, $\widehat{P}_l(\pi)$ and $\widehat{P}_{ipl}$ are given straightforwardly by means of our previous estimators and are thus asymptotically unbiased if order statistics belong to the interval $S^0(F_{\mathbf{B_p},\mathbf{Z},\mathbf{I}})$.

$$\widehat{Prob}(\widetilde{\pi}=\pi|(B_l,Z_l,I_l)) = \frac{\prod_{i\in I_l}\widehat{f}_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(B_{\pi(i)l}|Z_l,I_l)}{\sum_{\pi'\in\sigma_{I_l}}\prod_{i\in I_l}\widehat{f}_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(B_{\pi'(i)l}|Z_l,I_l)}\cdot \mathbf{1}\{\pi\in\sigma_{I_l}\} \tag{22}$$

$$\widehat{P}_{ipl} = \sum_{\pi\in\sigma_{I_l}\ s.t.\ \pi(i)=p} \widehat{Prob}(\widetilde{\pi}=\pi|(B_l,Z_l,I_l)) \tag{23}$$

In the final step we use the pseudo sample $\{(\widehat{X}_{ipl},\widehat{P}_{ipl},Z_l), i=1,\cdots,n, p=1,\cdots,n, l=1,\cdots,L\}$ to estimate nonparametrically the densities $f_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(x|z,I)$ by $\widehat{f}_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(x|z,I) = \widehat{f}_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(x,z,I)/\widehat{f}_{\mathbf{Z},\mathbf{I}}(z,I)$, where

$$\widehat{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) = \frac{1}{Lh^{d+1}_{f_{\mathbf{X_i},\mathbf{Z}}}}\sum_{l=1}^{L}\sum_{p\in \mathrm{I}_l}\widehat{P}_{ipl}\cdot K_{f_{\mathbf{X_i},\mathbf{Z}}}(\frac{x-\widehat{X}_{ipl}}{h_{f_{\mathbf{X_i},\mathbf{Z}}}},\frac{z-Z_l}{h_{f_{\mathbf{X_i},\mathbf{Z}}}})\cdot\mathbf{1}(I_l=I), \tag{24}$$

$$\widehat{f}_{\mathbf{Z},\mathbf{I}}(z,I) = \frac{1}{Lh^d_{\mathbf{Z}}}\sum_{l=1}^{L}\sum_{p=1}^{n}K_{f_{\mathbf{Z}}}(\frac{z-Z_l}{h_{f_{\mathbf{Z}}}})\cdot\mathbf{1}(I_l=I), \tag{25}$$

$h_{f_{\mathbf{X_i},\mathbf{Z}}}$ and $h_{f_{\mathbf{Z}}}$ are bandwidths, and $K_{f_{\mathbf{X_i},\mathbf{Z}}}$ and $K_{f_{\mathbf{Z}}}$ are kernels.

We now turn to the choice of kernels and bandwidths defining our multi-step estimators.

Assumption A 6 • ***KERNELS***

(i) *The kernels* $K_{F_{\mathbf{B_p}|\mathbf{Z}}}(.)$, $K_{f_{\mathbf{B_p}|\mathbf{Z}}}(.,.)$, $K_{f_{\mathbf{X_i},\mathbf{Z}}}(.,.)$ *and* $K_{f_{\mathbf{Z}}}(.)$ *are symmetric with bounded hypercube supports of length equal to* 2 *and continuous bounded first derivatives with respect to their continuous argument.*

(ii) $\int K_{F_{\mathbf{B_p}|\mathbf{Z}}}(z)dz=1$, $\int K_{f_{\mathbf{B_p}|\mathbf{Z}}}(b,z)dbdz=1$, *for any* $p=1,\cdots,n$, $\int K_{f_{\mathbf{X_i},\mathbf{Z}}}(x,z)dxdz=1$ *for any* $i=1,\cdots,n$ *and* $\int K_{f_{\mathbf{Z}}}(z)dz=1$.

(iii) $K_{F_{\mathbf{B_p}|\mathbf{Z}}}(.)$, $K_{f_{\mathbf{B_p}|\mathbf{Z}}}(.,.)$, $K_{f_{\mathbf{X_i},\mathbf{Z}}}(.,.)$ *and* $K_{f_{\mathbf{Z}}}(.)$ *are of order* $R+1$, $R+1$,R *and* $R+1$ *respectively, i.e. moments of order strictly smaller than the given order vanish.*

• ***BANDWIDTHS***

(i) *In the first price auction, the bandwidths* $h_{F_{\mathbf{B_p}|\mathbf{Z}}}, h_{f_{\mathbf{B_p}|\mathbf{Z}}}$, *for* $p=1,\cdots,n$, $h_{f_{\mathbf{X_i},\mathbf{Z}}}$ *for* $i=1,\cdots,n$ *and* $h_{f_{\mathbf{Z}}}$ *are of the form:*

$$h_{F_{\mathbf{B_p}|\mathbf{Z}}} = \lambda_{F_{\mathbf{B_p}|\mathbf{Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+2)}},\quad h_{f_{\mathbf{B_p}|\mathbf{Z}}} = \lambda_{f_{\mathbf{B_p}|\mathbf{Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+3)}},$$

$$h_{\mathbf{f_{X_i,Z}}} = \lambda_{\mathbf{f_{X_i,Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+3)}}, \quad h_{f_{\mathbf{Z}}} = \lambda_{f_{\mathbf{Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+2)}},$$

where the λ's are strictly positive constants.

(ii) *In the second price auction, the bandwidths $h_{F_{\mathbf{B_p|Z}}}, h_{f_{\mathbf{B_p|Z}}}$, for $p = 1, \cdots, n$, $h_{f_{\mathbf{X_i,Z}}}$ for $i = 1, \cdots, n$ and $h_{f_{\mathbf{Z}}}$ are of the form:*

$$h_{F_{\mathbf{B_p|Z}}} = \lambda_{F_{\mathbf{B_p|Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d)}}, \quad h_{f_{\mathbf{B_p|Z}}} = \lambda_{f_{\mathbf{B_p|Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+1)}},$$

$$h_{\mathbf{f_{X_i,Z}}} = \lambda_{\mathbf{f_{X_i,Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+1)}}, \quad h_{f_{\mathbf{Z}}} = \lambda_{f_{\mathbf{Z}}}(\frac{\log L}{L})^{\frac{1}{(2R+d+2)}},$$

(iii) *The "boundary" bandwidth is of the form $h_\delta = \lambda_\delta(\frac{\log L}{L})^{\frac{1}{d+1}}$ if $d > 0$*
where the λ's are strictly positive constants.

4.4 Uniform Consistency

Our main result establishes the uniform consistency of our multistage kernel-based estimators for the first and second price auctions and with their rates of convergence. As a preliminary step, next proposition, the analog of propositions 2 and 3 in GPV, establishes the uniform consistency with their rates of convergence of our nonparametric estimators of the upper and lower boundaries $\bar{b}(z, I)$ and $\underline{b}$ and also the rates at which the pseudo private values $\widehat{X}_{ipl}$ and the pseudo probabilities $\widehat{P}_{ipl}$ converge uniformly to their true values.

Proposition 4.3 *Under A1-A6, for any closed subset $\mathcal{C}$ of $S^o(F_{\mathbf{X},\mathbf{Z},\mathrm{I}})$, we have almost surely:*

$$\sup_{(z,I)\in[\underline{z},\overline{z}]\times\mathcal{I}} |\widehat{\bar{b}}(z, I) - \bar{b}(z, I)| = O(\frac{\log L}{L})^{\frac{1}{d+1}},$$

$$|\widehat{\underline{b}} - \underline{b}| = O(\frac{\log L}{L})^{\frac{1}{d+1}}$$

for both the first and second price auctions.

The pseudo values and pseudo probabilities satisfy almost surely:

(i)

$$\sup_{i,p,l} \mathbf{1}_{\mathcal{C}}(X_{ipl}, Z_l, \mathrm{I}_l)|\widehat{X}_{ipl} - X_{ipl}| = O\left((\frac{\log L}{L})^{\frac{R+1}{(2R+d+3)}}\right)$$

(ii)

$$\sup_{i,p,l} \mathbf{1}_{\mathcal{C}}(X_{ipl}, Z_l, \mathrm{I}_l)|\widehat{P}_{ipl} - P_{ipl}| = O\left((\frac{\log L}{L})^{\frac{R+1}{(2R+d+3)}} \right)$$

in the first price auction and

(i)

$$\sup_{i,p,l} \mathbf{1}_{\mathcal{C}}(X_{ipl}, Z_l, \mathrm{I}_l)|\widehat{X}_{ipl} - X_{ipl}| = 0$$

(ii)

$$\sup_{i,p,l} \mathbf{1}_{\mathcal{C}}(X_{ipl}, Z_l, \mathrm{I}_l)|\widehat{P}_{ipl} - P_{ipl}| = O\left((\frac{\log L}{L})^{\frac{R}{(2R+d+1)}} \right)$$

in the second price auction.

In the same way as the vector of pseudo private values are not sufficient to estimate the CDFs of each bidders private values (on the contrary to GPV), the estimation of conditional mean, variance or quantiles of a given bidder's private values requires the joint use of the pseudo private values with the associated vector of pseudo probabilities.

We now state our main result. The study of uniform convergence is restricted to inner closed subset of the support to avoid boundary effects.

Proposition 4.4 *Suppose that A1-A6 hold, then* $(\widehat{f}_{\mathbf{X_1}|\mathbf{Z},\mathbf{I}}(.|.,.), \cdots, \widehat{f}_{\mathbf{X_n}|\mathbf{Z},\mathbf{I}}(.|.,.))$ *is uniformly consistent as* $L \to \infty$ *with rate* $(L/\log L)^{R/(2R+d+3)}$ *on any inner compact subset of the support of* $(f_{\mathbf{X_1}|\mathbf{Z},\mathbf{I}}(.|.,.), \cdots, f_{\mathbf{X_n}|\mathbf{Z},\mathbf{I}}(.|.,.))$ *in the first price auction and respectively the rate* $(L/\log L)^{R/(2R+d+1)}$ *in the second price auction.*

In addition to establishing the uniform consistency of our multistep estimator, Proposition 4.4 implies that the upper bounds that have been derived for the first and second price auctions in Proposition 4.2 are in fact the optimal uniform convergence rates for estimators of the conditional density $F_{\mathbf{X}|\mathbf{Z},\mathbf{I}}(.|.,.)$ from observed anonymous bids and that our procedure is asymptotically optimal. On the contrary, if the interest of the econometrician lies only in the estimation of the distributions $F_{\mathbf{B}^*|\mathbf{Z},\mathbf{I}}(.|.,.)$, then, in the first price auction, our bandwidths are suboptimal and the same bandwidths as those for the second price auction should be used.

We present the proof of Proposition 4.4 as it helps identify the additional points relative to GPV's two step procedure and why it does not change the asymptotical rates of convergence.

Proof 1 *We have* $\widehat{f}_{\mathbf{X_i}|\mathbf{Z},\mathbf{I}}(x|z,I) = \widehat{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I)/\widehat{f}_{\mathbf{Z},\mathbf{I}}(z,I)$. *Given the optimal bandwidth choice for* $h_{f_{\mathbf{Z}}}$ *in assumption A(6), we know that* $\widehat{f}_{\mathbf{Z},\mathbf{I}}(z,I)$ *converges uniformly to* $f_{\mathbf{Z},\mathbf{I}}(z,I)$ *at the rate* $(L/\log L)^{(R+1)/(2R+d+1)}$ *on any* *inner compact of its support. Because this rate is faster than that of the theorem (for both the first and second price auction) and* $f_{\mathbf{Z},\mathbf{I}}(z,I)$ *is bounded away from 0 by assumption A3-(ii), it suffices to show that* $\widehat{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I)$ *converges at the rate* $(\frac{\log L}{L})^{R/(2R+d+3)}$ *and* $(\frac{\log L}{L})^{R/(2R+d+1)}$ *in the first and second price auctions respectively.*

We decompose the difference $\widehat{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) - f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x|z,I)$ *into three terms.*

$$\begin{aligned}
&\widehat{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) - f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) \\
&\qquad = \frac{1}{Lh^{d+1}_{f_{\mathbf{X_i},\mathbf{Z}}}} \sum_{l=1}^{L} \sum_{p\in \mathrm{I}_l} (\widehat{P}_{ipl} - P_{ipl}) \cdot K_{f_{\mathbf{X_i},\mathbf{Z}}}(\frac{x - X_{ipl}}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}, \frac{z - Z_l}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}) \cdot \mathbf{1}(I_l = I) \\
&\qquad + \frac{1}{Lh^{d+1}_{f_{\mathbf{X_i},\mathbf{Z}}}} \sum_{l=1}^{L} \sum_{p\in \mathrm{I}_l} \widehat{P}_{ipl} \cdot \mathbf{1}(I_l = I) \\
&\qquad \times \left(K_{f_{\mathbf{X_i},\mathbf{Z}}}(\frac{x - \widehat{X}_{ipl}}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}, \frac{z - Z_l}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}) - K_{f_{\mathbf{X_i},\mathbf{Z}}}(\frac{x - X_{ipl}}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}, \frac{z - Z_l}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}) \right) \\
&\qquad + \widetilde{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) - f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I),
\end{aligned} \tag{26}$$

where $\widetilde{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}$ *is the (infeasible) nonparametric estimator of the density of* (X_i, Z, I) *using the unobserved values* X_{ipl} *and the unobserved probabilities* P_{ipl}*:*

$$\widetilde{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) = \frac{1}{Lh^{d+1}_{f_{\mathbf{X_i},\mathbf{Z}}}} \sum_{l=1}^{L} \sum_{p\in \mathrm{I}_l} P_{ipl} \cdot K_{f_{\mathbf{X_i},\mathbf{Z}}}(\frac{x - X_{ipl}}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}, \frac{z - Z_l}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}) \cdot \mathbf{1}(I_l = I).$$

In the second price auction, the bandwidth $h_{\mathbf{f_{X_i},Z}}$ *is optimal and thus leads to a uniform convergence of* $\widetilde{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I)$ *to* $f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I)$ *at the rate* $(L/\log L)^{R/(2R+d+1)}$ *in any inner compact of its support. In the first price auction, the suboptimal bandwidth leads to the rate* $(L/\log L)^{R/(2R+d+3)}$ *as in GPV. Thus we are left with the first two terms, the first one resulting explicitly from the anonymous nature of the bids is new, whereas the second term appears already in GPV.*

First consider the second price auction. Since $\widehat{X}_{ipl} = X_{ipl}$*, the second term vanishes and we are left with the first term which is bounded by:*

$$\sup_{p,l} \mathbf{1}_{\mathcal{C}}(X_{ipl}, Z_l, \mathrm{I}_l)|\widehat{P}_{ipl} - P_{ipl}| \cdot \left[\frac{1}{Lh^{d+1}_{f_{\mathbf{X_i},\mathbf{Z}}}} \sum_{l=1}^{L} \sum_{p\in \mathrm{I}_l} K_{f_{\mathbf{X_i},\mathbf{Z}}}(\frac{x - X_{ipl}}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}, \frac{z - Z_l}{h_{f_{\mathbf{X_i},\mathbf{Z}}}}) \cdot \mathbf{1}(I_l = I) \right]$$

The above term appearing in the bracket may be viewed as a kernel estimator, and hence converges uniformly on $\mathcal{C}$ *to*

$$\sum_{p\in I} f_{\mathbf{X_{ip}},\mathbf{Z},\mathbf{I}}(x,z,I)\cdot\int f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I)dxdz$$

Thus this term stays bounded almost surely. Finally the difference $\widehat{f}_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) - f_{\mathbf{X_i},\mathbf{Z},\mathbf{I}}(x,z,I) = O(\log L/L)^{R/(2R+d+1)}$.

In the first price auction, similarly to GPV, a first-order Taylor expansion establishes that the second term has the order $O(\log L/L)^{R/(2R+d+3)}$*, whereas the same argument as above establishes that the first term has the order* $O(\log L/L)^{(R+1)/(2R+d+3)}$*. Thus with anonymity, it is still the second term that results from the gap between estimated and real private values that is the 'binding' term relative to the uniform convergence rate.*

5 Conclusion

Our identification methodology has been limited to the independent private value framework with risk neutral bidders, no reserve price and a complete set of bids. For the second price auction, we can be reluctant to propose identification and estimation methods that are relying on the observation of the complete set of bids, in particular on the observation of the highest bid which may remain unobserved. Moreover, this excludes any direct application for the ascending (English) auction.

All our analysis of the first-price auction can be adapted to risk averse bidders under a conditional quantile restriction and a parametrization of bidders' utility function following Campo et al. [6]. As in GPV, our analysis can also be adapted to a binding reserve price provided that we are prepared to assume that the number of potential bidders is constant. Naturally, identification is obtained only for the truncated distribution of types that are above the reserve price. More involved is the extension of our methodology with incomplete sets of bids, whose developments are left for further research. Let us briefly precise the issues. Each ordered statistic leads to an equation leading thus to an n equations system, whereas we face n unknowns. Thus the least unobserved bidding statistic leads to unidentification. There are two routes to restore identification. First, to impose more symmetry by assuming that some bidders are symmetric: it corresponds to a reduction of the number of unknowns. Second, to exploit some exogenous variations in the number of bidders: it corresponds to an expansion of the number of equations.

Note that the symmetric APV model is not identified if we do not observe the highest bid (Theorem 4 in Athey and Haile [2]) and that identification could not be tackled even if we do observe some exogenous variation in the number of bidders. Thus the way we exploit independence could be developed in further research to obtain identification with an incomplete set of anonymous bids and which goes beyond the symmetric IPV framework that is currently used in such a case.

Our approach can also be used for alternative asymmetric auction models with independent private signals as the one developed by Landsberger et al. [15] where the ranking of bidders' private valuations is common knowledge among bidders, but not to the econometrician. A promising avenue for research, which was the initial motivation of this work, is the structural analysis of models with shill bidding as developed by Lamy [14]. In a private value framework, models with shill bidding are strategically equivalent to models with a secret reserve price. It differs only from the econometrician point of view: in the latter, she distinguishes a submitted bid from the reserve price which facilitates the estimation as in [8, 17], whereas, in the former, the strategic bidding activity of the seller is indistinguishable from any other bid. Nevertheless, our methodology can be adapted.

References

[1] S. Athey. Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69(4):861–890, 2001.

[2] S. Athey and P. Haile. Identification of standard auction models. *Econometrica*, 70(6):2107–2140, 2002.

[3] L. H. Baldwin, R. C. Marshall, and J.-F. Richard. Bidder collusion at forest service timber sales. *J. Polit. Economy*, 105:657–699, 1997.

[4] C. L. Benkard and S. Berry. On the nonparametric identification of nonlinear simultaneous equations models: Comment on brown (1983) and roehrig (1988). *Econometrica*, 74(5):1429–1440, 2006.

[5] J. Blum, J. Kiefer, and M. Rosenblatt. Distribution free tests of indenpendence based on the sample distributions functions. *Annals of Mathematical Statistics*, 32:485–498, 1961.

[6] S. Campo, E. Guerre, I. Perrigne, and Q. Vuong. Semiparametric estimation of first-price auctions with risk averse bidders. *mimeo*, 2002.

[7] S. Campo, I. Perrigne, and Q. Vuong. Asymmetry in first-price auctions with affiliated private values. *J. Appl. Econ.*, 18:179–207, 2003.

[8] B. Elyakime, J.-J. Laffont, P. Loisel, and Q. Vuong. First-price sealed-bid auctions with secret reservation prices. *Annales d'Economie et de Statistique*, 34:115–141, 1994.

[9] V. Flambard and I. Perrigne. Asymmetry in procurement auctions: Evidence from snow removal contracts. *The Economic Journal*, 116:1014–1036, 2006.

[10] E. Guerre, I. Perrigne, and Q. Vuong. Optimal nonparametric estimation of first price auctions. *Econometrica*, 68:525–574, 2000.

[11] K. Hendricks and R. H. Porter. An empirical study of an auction with asymmetric information. *Amer. Econ. Rev.*, 78:865–83, 1988.

[12] R. Khas'minskii. A lower bound on the risks of nonparametric estimates of densities. *Theory of Probability and its Applications*, 23:794–798, 1978.

[13] J.-J. Laffont and Q. Vuong. Structural analysis of auction data. *Amer.Econ. Rev. Papers and Proceeding*, 86(2):414–420, 1996.

[14] L. Lamy. The shill bidding effect versus the linkage principle. *mimeo CREST-INSEE*, 2006.

[15] M. Landsberger, J. Rubinstein, E. Wolfstetter, and S. Zamir. First price auctions when the ranking of valuations is common knowledge. *Review of Economic Design*, 3(4):461–480, 2001.

[16] B. Lebrun. Uniqueness of the equilibrium in first-price auctions. *Games Econ. Behav.*, 55:131–151, 2006.

[17] T. Li and I. Perrigne. Timber sale auctions with random reserve price. *Rev. Econ. Statist.*, 85:189–200, 2003.

[18] T. Li, I. Perrigne, and Q. Vuong. Structural estimation of the affiliated private value auction model. *RAND J. Econ.*, 33:171–193, 2002.

[19] E. Maskin and J. Riley. Asymmetric auctions. *Rev. Econ. Stud.*, 67(3):413–438, 2000.

[20] R. B. Myerson. Optimal auction design. *Mathematics of Operation Research*, 6(1):58–73, 1981.

[21] H. Paarsch and H. Hong. *An Introduction to the Structural Econometrics of Auction Data.* The MIT Press, Cambridge, Massachusetts, 2006.

[22] M. Pesendorfer. A study of collusion in first-price auctions. *Rev. Econ. Stud.*, 67(3):381–411, 2000.

[23] R. H. Porter and D. J. Zona. Detection of bid rigging in procurement auctions. *J. Polit. Economy*, 101:518–538, 1993.

[24] R. H. Porter and D. J. Zona. Ohio school milk markets: an analysis of bidding. *RAND J. Econ.*, 30:263–288, 1999.

[25] C. Roehring. Conditions for identification in nonparametric and parametric models. *Econometrica*, 56(2):433–447, 1988.

[26] K. Sailer. Searching the ebay marketplace. *CESifo Working Paper*, 2006.

[27] U. Song. Nonparametric estimation of an ebay auction model with an unknown number of bidders. *mimeo*, 2004.

[28] C. Stone. Optimal global rates of convergence for nonparametric estimators. *Annals of Statistics*, 10:1040–53, 1982.

[29] D. Uherka and A. M. Sergott. On the continuous dependence of the roots of a polynomial on its coefficients. *The American Mathematical Monthly*, 84(5):368–370, 1977.

[30] M. Yokoo, Y. Sakurai, and S. Matsubara. The effect of false-name bids in combinatorial auctions: new fraud in internet auctions. *Games and Economic Behavior*, 46:174–188, 2004.

A Proofs of Mathematical Properties

A.1 Proof of Proposition [3.1]

We write the proof for the first price auction, the most difficult case where the correspondance between bids and private signals is not immediate. In the second price auction, bids are equal to private values and the following proof can be easily adapted.

Remind that under observability of bidders' identities, Li, Perrigne and Vuong [18] show that the symmetric APV model is identified whereas Campo, Perrigne and Vuong [7] extend this result to the asymmetric APV model. Let us see why [18]'s proof remains valid under anonymity whereas [7]'s proof does not.

The main step to obtain identification is the equilibrium equation (7) that allows to express bidder i's private value x_i as the function of his bid b_i, the CDF $G_{\mathbf{B}_{-i}|\mathbf{b}_i}(.|.)$ and the PDF $g_{\mathbf{B}_{-i}|\mathbf{b}_i}(.|.)$ of his opponents bids conditional on his bid. Under observed identities, it is possible to obtain the full distribution of the vector of private valuations X since $G_{\mathbf{B}_{-i}|\mathbf{b}_i}(y|x)$ is observed. Under anonymity, only $\frac{1}{n} \cdot \sum_{i=1}^{n} G_{\mathbf{B}_{-i}|\mathbf{b}_i}(y|x)$ is observed, which prevents the use of the above equation except in the symmetric case where $G_{\mathbf{B}_{-i}|\mathbf{b}_i}(y|x) = \frac{1}{n} \cdot \sum_{i=1}^{n} G_{\mathbf{B}_{-i}|\mathbf{b}_i}(y|x)$. Therefore the symmetric APV model is identified.

For the asymmetric APV model, for any distribution of bids $F_{\mathbf{B}}$ and a given distribution of signals $F_{\mathbf{X}}$ that rationalizes $F_{\mathbf{B}}$, let us construct a distribution of signals $F'_{\mathbf{X}}$ that differs from $F_{\mathbf{X}}$ (up to any permutation) and that leads to $F_{\mathbf{B}}$. Consider two bids $\underline{b}$ and $\overline{b}$, $\overline{b} > \underline{b}$, used by all bidders, take $\epsilon < \frac{\overline{b}-\underline{b}}{2}$ such that bidders are bidding in the intervals $[\underline{b}-\epsilon, \underline{b}+\epsilon]$ and $[\overline{b}-\epsilon, \overline{b}+\epsilon]$. For any bidder i, define $\underline{x}_i^{-\epsilon}$, $\underline{x}_i$ and $\underline{x}_i^{+\epsilon}$ by the equations:

$$\begin{aligned} \underline{b}-\epsilon &= \beta_i(\underline{x}_i^{-\epsilon}) \\ \underline{b} &= \beta_i(\underline{x}_i) \\ \underline{b}+\epsilon &= \beta_i(\underline{x}_i^{+\epsilon}). \end{aligned} \tag{27}$$

We define $\overline{x}_i^{-\epsilon}$, $\overline{x}_i$ and $\overline{x}_i^{+\epsilon}$ in the same way. For a couple of bidders (i, j), define

$$\begin{aligned} c(x_1, \cdots, x_n; \epsilon, i, j) \equiv & \prod_{k \neq i,j} (\mathbf{1}\{x_k \in [\underline{x}, \underline{x}_k]\}) \cdot \mathbf{1}\{x_i \in [\overline{x}_i^{-\epsilon}, \overline{x}_i]\} \\ & \cdot \Big(\mathbf{1}\{x_j \in [\underline{x}_j, \underline{x}_j^{+\epsilon}]\} - \mathbf{1}\{x_j \in [\underline{x}_j^{-\epsilon}, \underline{x}_j]\}\Big) \\ & - \prod_{k \neq i,j} (\mathbf{1}\{x_k \in [\underline{x}, \underline{x}_k]\}) \cdot \mathbf{1}\{x_i \in [\overline{x}_i, \overline{x}_i^{+\epsilon}]\} \\ & \cdot \Big(\mathbf{1}\{x_j \in [\underline{x}_j, \underline{x}_j^{+\epsilon}]\} - \mathbf{1}\{x_j \in [\underline{x}_j^{-\epsilon}, \underline{x}_j]\}\Big) \end{aligned} \tag{28}$$

For sufficiently small $\gamma > 0$, $f'_{\mathbf{X}}(.) \equiv f_{\mathbf{X}}(.) + \gamma \cdot (c(.; \epsilon, i, j) - c(.; \epsilon, j, i))$ is a PDF with the functions c shifting probability weight from some regions to others.

In a first step, we prove that the first order conditions that characterizes equilibrium bidding functions do not change when the map $c(.; \epsilon, i, j)$ is added to the original PDF $f_{\mathbf{X}}(.)$. It results from the fact that both $G_{\mathbf{B}_{-s}|\mathbf{b}_s}(x|x)$ and $g_{\mathbf{B}_{-s}|\mathbf{b}_s}(x|x)$ do not change. On the one hand, for $s \neq j$, it is easy to see that the term $\mathbf{1}\{x_k \leq \underline{x}_k, \forall k \neq i, j\} \cdot \mathbf{1}\{x_i \in [\overline{x}_i^{-\epsilon}, \overline{x}_i]\} \cdot \Big(\mathbf{1}\{x_j \in [\underline{x}_j, \underline{x}_j^{+\epsilon}]\} - \mathbf{1}\{x_j \in [\underline{x}_j^{-\epsilon}, \underline{x}_j]\}\Big)$ does not modify $G_{\mathbf{B}_{-s}|\mathbf{b}_s}(x|x)$ and $g_{\mathbf{B}_{-s}|\mathbf{b}_s}(x|x)$. On the other hand, the second term in expression (28) has been explicitly added to guarantee that $G_{\mathbf{B}_{-s}|\mathbf{b}_s}(x|x)$ and $g_{\mathbf{B}_{-s}|\mathbf{b}_s}(x|x)$ do not change for $s = j$. Indeed, if γ is small enough $\gamma < \gamma_1$, then the original equilibrium bid functions still satisfy the global equilibrium conditions.

In a second step, we have to check that the full distribution of the complete set of bids remains the same. To understand our construction, first remark that the perturbation $\gamma \cdot c(.; \epsilon, i, j)$ alone changes the final distribution of bids. It shifts probability weight from regions where two bids are respectively in the intervals $[\underline{b}-\epsilon, \underline{b}]$ and $[\overline{b}-\epsilon, \overline{b}]$ (respectively in the intervals $[\underline{b}, \underline{b}+\epsilon]$ and $[\overline{b}, \overline{b}+\epsilon]$) to regions where two bids are respectively in the intervals $[\underline{b}-\epsilon, \underline{b}]$ and $[\overline{b}, \overline{b}+\epsilon]$ (respectively in the intervals $[\underline{b}, \underline{b}+\epsilon]$ and $[\overline{b}-\epsilon, \overline{b}]$). Subtracting the (symmetric) permutation $\gamma \cdot c(.; \epsilon, j, i)$ allows to restore those shifts in the bids joint distribution making it identical to the original one.

Finally, we have to check that $f'_{\mathbf{X}}(.)$ and $f_{\mathbf{X}}(.)$ do not coincide up to a permutation. By coincidence, for a given γ, there may exist a permutation π such that $f'_{\mathbf{X}}(x_1, \cdots, x_n) = f_{\mathbf{X}}(x_{\pi(1)}, \cdots, x_{\pi(n)})$. Our construction is valid for any γ which is sufficient small, thus an infinite number of γ are potential candidates. On the other hand, there exists only a finite number of permutation. Now suppose that for any $\gamma < \gamma_1$ there exists a permutation π_γ such that $f'_{\gamma,\mathbf{X}}(x_1, \cdots, x_n) = f_{\mathbf{X}}(x_{\pi_\gamma(1)}, \cdots, x_{\pi_\gamma(n)})$. Then there exists γ^a and γ^b such that $\pi_{\gamma^a} = \pi_{\gamma^b}$, which implies that the function $(c(.;\epsilon, i, j) - c(.;\epsilon, j, i))$ should be null, which raises a contradiction.

For instance, we have only proved that, for any asymmetric PV model, there exists a local perturbation which corresponds to an asymmetric PV model and that leads to the same distribution of bids. Indeed the above perturbation may break affiliation due to the non-smoothness of the indicator function. Let $\phi(.)$ be a smoothed version of the indicator function on the interval $[0,1]$: $\phi(x) > 0$ if and only if $x \in [0,1]$ and $\int \phi = 1$. Then in the above perturbations, replace the expressions of the kind $\mathbf{1}\{a \in [\underline{a}, \overline{a}]\}$ by $\phi(\frac{a-\underline{a}}{\overline{a}-\underline{a}})$. The resulting modified perturbations are still shifting probability weight from some regions to others for γ sufficiently small. Moreover, by setting γ sufficiently small, the expressions

$$\partial^2 log(1 + \frac{\gamma \cdot (c(x_1, \cdots, x_n; \epsilon, i, j) - c(x_1, \cdots, x_n; \epsilon, j, i))}{f_{\mathbf{X}}(x_1, \cdots, x_n)})/\partial x_i \partial x_j,$$

for any (i, j), $i \neq j$, can be made arbitrarily small, which guarantees that strict affiliation is preserved if γ is small enough.

A.2 Proof of Proposition [4.1]

In their proposition 1, GPV obtain the same properties for the CDFs $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}$ instead of $F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}$. From (3) and (4), we obtain that any CDF $F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(b, z, I)$ can be expressed as a linear combination of terms which are product of $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}(b, z, I)$, i.e. as a continuous function of the CDFs $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}$. The CDF $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}$ have the desired smoothness properties on the set $S^0(F_{\mathbf{B_n}|\mathbf{Z},\mathbf{I}}) \setminus \{\overline{b}(z, I, i)\}$: on the set $S^0(F_{\mathbf{B_i}|\mathbf{Z},\mathbf{I}})$, it comes from GPV, whereas $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}$ is equal to 1 above $\overline{b}(z, I, i)$ and is thus C^∞. Thus all the regularity properties (iii-v) that are valid for $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}$ are still valid for $F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}$ if the points $\{\overline{b}(z, I, i)\}$ have been appropriately removed. The image of a closed interval by a continuous function is a closed interval. Thus (i) holds also for $F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}$. Finally we are left with (ii). Note the difference between the similar point in GPV which holds for the whole support and not only for a closed subset of the $S^o(F_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}})$ as above. By deriving (4) and (3), we obtain an another expression of $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(b, z, I)$ as a function of $F_{\mathbf{B_i^*}|\mathbf{Z},\mathbf{I}}(b, z, I)$ and

$f_{\mathbf{B}_\mathbf{i}^*|\mathbf{Z},\mathbf{I}}(b,z,I)$:

$$f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(b,z,I) = \frac{1}{(p-1)!(N_I-p-1)!} \cdot \sum_{\pi\in\Sigma_I}\left[\prod_{i=1}^{p-1} F_{\mathbf{B}^*_{\pi(\mathbf{i})}|\mathbf{Z},\mathbf{I}}(b,z,I)\cdot f_{\mathbf{B}^*_{\pi(\mathbf{p})}|\mathbf{Z},\mathbf{I}}(b,z,I)\cdot \prod_{i=p+1}^{n_I}(1-F_{\mathbf{B}^*_{\pi(\mathbf{i})}|\mathbf{Z},\mathbf{I}}(b,z,I))\right]$$

Thus we obtain that $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(b,z,I)$ is strictly positive on $S^o(F_{\mathbf{B_n}|\mathbf{Z},\mathbf{I}})$. Remark that $f_{\mathbf{B_p}|\mathbf{Z},\mathbf{I}}(b,z,I)$ is null at the lower bound $b=\underline{b}(z,I,p)$ for $p>1$ (respectively at the upper bound $b=\overline{b}(z,I,p)$ for $p<n$).

B Proofs of Statistical Properties

The proofs of the statistical properties are very closely related to GPV. The proof for the derivation of the asymptotic uniform rate of convergence of bidders' private values uses intensively the rates derived previously by GPV. Less obvious is the adaption of GPV's proof for the upper bound on the uniform convergence rate. We follow their proof very carefully and focus only on the new ingredients.

B.1 Optimal Uniform Convergence Rate

We adapt GPV's proof to the asymmetric framework. To ease the exposition, we consider the case where there is a positive probability that $n_I=2$. Without loss of generality, this set is $\{1,2\}$ and is denoted by I_2. The first step is identical to GPV: it is sufficient to prove the proposition by replacing $f_{\mathbf{X}|\mathbf{Z},\mathbf{I}}$ by $f_{\mathbf{X},\mathbf{Z},\mathbf{I}}$. The set $U_\epsilon(f^o_{\mathbf{X},\mathbf{Z},\mathbf{I}})$ can also be replaced by any subset $U\subset U_\epsilon(f^o_{\mathbf{X},\mathbf{Z},\mathbf{I}})$. Then the second step consists in the construction of a discrete subset U of the form $\{f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.), k=1,\cdots,m^{d+1}\}$, where m is increasing with the sample size L, that are suitable perturbations of $f^o_{\mathbf{X},\mathbf{Z},\mathbf{I}}$.

We consider a nonconstant and odd C_∞-function ϕ, with support $[-1,1]^{d+1}$, such that

$$\int_{[-1,0]}\phi(b,z)db=0,\quad \phi(0,0)=0,\quad \phi'(0,0)\neq 0, \tag{29}$$

where ϕ' denotes the derivative of ϕ according to its first component.

Let $\mathcal{C}_{I_2}(B^*)$ be the image of $\mathcal{C}(X)$ by the function that maps bidders' types into observed bids and conditionally on $I=I_2$. It is a nonempty inner compact subset of $S(f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}})$. Let $(b_k,z_k), k=1,\cdots,m^{d+1}$ be distinct points in the interior of $\mathcal{C}_{I_2}(B^*)$ such that the distance between (b_k,z_k) and (b_j,z_j), $j\neq k$, and the distance between (b_k,z_k) and any point outside $\mathcal{C}_{I_2}(B^*)$ are larger than λ_1/m. Thus, one can choose a constant $\lambda_2>1/\lambda_1$ such that the m^{d+1} functions

$$\phi_{mk}(b,z) = \frac{1}{m^{R+1}}\phi(m\lambda_2(b-b_k), m\lambda_2(z-z_k)) \quad (k=1,\cdots,m^{d+1})$$

have disjoint hypercube supports. Let C_3 be a positive constant (chosen below), for each $k = 1,\cdots,m^{d+1}$ define:

$$f_{\mathbf{B}_\mathbf{i}^*,\mathbf{Z},\mathbf{I},mk}(b,z,I_2) \equiv \begin{cases} f^o_{\mathbf{B}_\mathbf{1}^*,\mathbf{Z},\mathbf{I}}(b,z,I_2) & \text{if } i=1 \\ f^o_{\mathbf{B}_\mathbf{2}^*,\mathbf{Z},\mathbf{I}}(b,z,I_2) - C_3\phi_{mk}(b,z) & \text{if } i=2, \end{cases} \tag{30}$$

whereas define $f_{\mathbf{B}^*,\mathbf{Z},\mathbf{I},mk}(b,z,I) \equiv f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}(b,z,I)$ for $I \neq I_2$. That is $f_{\mathbf{B}^*,\mathbf{Z},\mathbf{I},mk}$ differs from $f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}$ only for $I = I_2$ and in the neighborhood of (b_k, z_k). The function $f_{\mathbf{B}^*,\mathbf{Z},\mathbf{I},mk}(b,z,I)$ is a density if C_3 is small enough (integrates to 1 from (29) and is bounded away from 0) with the same support as $f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}(b,z,I)$. Now consider the functions $\xi_{i,mk}(b,z) = b + \frac{F_{\mathbf{B}^*_{\mathbf{3-i}},\mathbf{Z},\mathbf{I},mk}(b,z,I_2)}{f_{\mathbf{B}^*_{\mathbf{3-i}},\mathbf{Z},\mathbf{I},mk}(b,z,I_2)}$ for $i = 1,2$. If C_3 is small enough, then $\xi_{i,mk}(b,z), i = 1,2$ is increasing in b with a differentiable inverse denoted by $\xi^{-1}_{i,mk}(x,z)$. Then we define for $I = I_2$ and $i = 1,2$

$$f_{\mathbf{X_i},\mathbf{Z},\mathbf{I},mk}(x,z,I_2) = f_{\mathbf{B}_\mathbf{i}^*,\mathbf{Z},\mathbf{I},mk}(\xi^{-1}_{i,mk}(x,z),z,I_2)/\xi'_{i,mk}(\xi^{-1}_{i,mk}(x,z),z) \tag{31}$$

$$= \frac{f_{\mathbf{B}_\mathbf{i}^*,\mathbf{Z},\mathbf{I},mk}(\xi^{-1}_{i,mk}(x,z),z,I_2)\cdot(f_{\mathbf{B}^*_{\mathbf{3-i}},\mathbf{Z},\mathbf{I},mk}(\xi^{-1}_{3-i,mk}(x,z),z,I_2))^2}{2(f_{\mathbf{B}^*_{\mathbf{3-i}},\mathbf{Z},\mathbf{I},mk}(\xi^{-1}_{i,mk}(x,z),z,I_2))^2 - F_{\mathbf{B}^*_{\mathbf{3-i}},\mathbf{Z},\mathbf{I},mk}(\xi^{-1}_{i,mk}(x,z),z,I_2)f'_{\mathbf{B}^*_{\mathbf{3-i}},\mathbf{Z},\mathbf{I},mk}(\xi^{-1}_{i,mk}(x,z),z,I_2)}$$

For $I \neq I_2$, let $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.) = f^o_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.)$. From the above expression, $f_{\mathbf{X_i},\mathbf{Z},\mathbf{I},mk}(x,z,I_2) > 0$ if and only if $f_{\mathbf{B}_\mathbf{i}^*,\mathbf{Z},\mathbf{I},mk}(b,z,I_2) > 0$, where $b = \xi^{-1}_{i,mk}(x,z)$. This completes the construction of the densities $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.)$, $k = 1,\cdots,m^{d+1}$, which composes the set U. Note that the supports of $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.)$ and $f_{\mathbf{B}^*,\mathbf{Z},\mathbf{I},mk}(.,.,.)$ coincide respectively with the supports of $f^o_{\mathbf{X},\mathbf{Z},\mathbf{I}}(.,.,.)$ and $f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}(.,.,.)$.

Then to adapt GPV's proof, we need the analog of their lemma B1 where the notation $f_{mk}(.,.,.)$ should be replaced by $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.)$, where the first argument x is now the vector of bidders' private values instead of a single unidimensional private value. The analog of Lemma B1 gives two points. First, an appropriate asymptotic lower bound is given for the uniform distance between two elements, i.e. the norm $||.||_{0,\mathcal{C}(X)}$, in the set U as a function of λ_2, m and R. With this bound we can apply Fano's lemma exactly in the same way as in GPV: the step 3 in their proof is unchanged. Second, an asymptotic approximation is given for the distance between $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}$ and $f^0_{\mathbf{X},\mathbf{Z},\mathbf{I}}$ in the norm $||.||_{r,\mathcal{C}(X)}$, which guarantees that $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}$ belongs to the set U if m is large enough.

Lemma B.1 (Analog of lemma B1 in GPV) *Given A2-A3, the following properties hold for m large enough:*

(i) *For any $k = 1, \cdots, m^{d+1}$, the supports of $f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(.,.,.)$ and $f_{\mathbf{B}^*,\mathbf{Z},\mathbf{I},mk}(.,.,.)$ are $S(f^o_{\mathbf{X},\mathbf{Z},\mathbf{I}}(.,.,.))$ and $S(f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}(.,.,.))$.*

(ii) *There is a positive constant C_4 depending upon ϕ, $f^o_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}(.,.,.)$ and $\mathcal{C}(X)$ such that for $j \neq k$,*

$$||f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk} - f_{\mathbf{X},\mathbf{Z},\mathbf{I},mj}||_{0,\mathcal{C}(X)} \geq C_4 \cdot \frac{C_3 \lambda_2}{m^R}.$$

(iii) *Uniformly in $k = 1, \cdots, m^{d+1}$, we have*

$$||f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk} - f^0_{\mathbf{X},\mathbf{Z},\mathbf{I}}||_r = C_3 \lambda_2^{r+1} O(\frac{1}{m^{R-r}}), r = 0 \cdots R-1$$

$$||f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk} - f^0_{\mathbf{X},\mathbf{Z},\mathbf{I}}||_R = C_3 \lambda_2^{R+1} \cdot O(1) + o(1).$$

where the big $O(.)$ depends upon ϕ and $f^0_{\mathbf{B}^*,\mathbf{Z},\mathbf{I}}$

Let us detail the proof of (ii) and what has changed relative to GPV's framework. Remind that $(b_k, z_k) \in \mathcal{C}_{I_2}(B^*)$ implies $(x_k, z_k) \in \mathcal{C}(X)$. As in GPV, it then suffices to prove that $|f_{\mathbf{X},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) - f_{\mathbf{X},\mathbf{Z},\mathbf{I},mj}(x_k, z_k, I_2)| \geq C_4 \cdot \frac{C_3 \lambda_2}{m^R}$, where $x_k = \xi^0(b_k, z_k, I_2)$.

From (29), we have: $F_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) = F^0_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(x_k, z_k, I_2)$ and $f_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) = f^0_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(x_k, z_k, I_2)$ for $i = 1, 2$. The difference is for the expression of $f'_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) - f'^0_{\mathbf{B_i^*},\mathbf{Z},\mathbf{I}}(x_k, z_k, I_2)$ which equals to 0 for $i = 1$ and to $-C_3 \frac{\lambda_2}{m^R} \phi'(0,0) \neq 0$ for $i = 2$. Thus $f_{\mathbf{X_2},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) = f_{\mathbf{X_2},\mathbf{Z},\mathbf{I},\mathbf{mj}}(x_k, z_k, I_2)$ which is bounded away from zero and we are left with the term $f_{\mathbf{X_1},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) - f_{\mathbf{X_1},\mathbf{Z},\mathbf{I},\mathbf{mj}}(x_k, z_k, I_2)$.

Then, from equation (31), we have:

$$f_{\mathbf{X_1},\mathbf{Z},\mathbf{I},mk}(x_k, z_k, I_2) = \frac{f^0_{\mathbf{B_1^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2) \cdot (f^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2))^2}{2(f^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2))^2 - F^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2)(f'^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2) - C_3 \lambda_2 \phi'(0,0)/m^R)} \tag{32}$$

and

$$f_{\mathbf{X_2},\mathbf{Z},\mathbf{I},mj}(x_k, z_k, I_2) = \frac{f^0_{\mathbf{B_1^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2) \cdot (f^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2))^2}{2(f^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2))^2 - F^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2)(f'^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}(b_k, z_k, I_2))} \tag{33}$$

Now compare (32) and (33). As $\phi'(0,0) \neq 0$ and $F^0_{\mathbf{B_2^*},\mathbf{Z},\mathbf{I}}$ is bounded away from zero since (b_k, z_k) are far enough from the boundaries, the desired result (ii) follows. The proof of (iii) is more involved and follows GPV's proof with the same modification as above by carefully separating the cases $i = 1$ and $i = 2$. More precisely, we have $||f_{\mathbf{X_1},\mathbf{Z},\mathbf{I},mk} - f^0_{\mathbf{X_1},\mathbf{Z},\mathbf{I}}||_r = C_3 \lambda_2^{r+1} O(\frac{1}{m^{R-r}})$ and $||f_{\mathbf{X_2},\mathbf{Z},\mathbf{I},mk} - f^0_{\mathbf{X_2},\mathbf{Z},\mathbf{I}}||_r = O(1)$ and the result follows for the product.

B.2 Proof of Proposition [4.3]

There are two new points relative to GPV's analysis. First, their proof is based on the uniform rates of convergence for the CDF, the PDF and also the boundaries estimators of the variable B^* that is observed by the econometrician. Here we do not observe B^* but only the vector of order statistics B. Second, the pseudo probabilities are a new ingredient that do not appear in GPV.

The first issue is then to prove that the same uniform rates of convergence are still valid for B^* though it is not observed. Nevertheless, the uniform rates of convergence they obtained for B^* are still valid under anonymity for the variable B that is observed and with our similar choices for the kernels and the bandwidth parameters. On the contrary to GPV, note that the observed variable B is multidimensional: it does not modify their analysis which immediately adapts.

First the bidding support of the bidders are coinciding with the support of the order statistics. Thus all the results for the estimator of the support of B are immediately converted into results for B^*. From GPV (lemma B2), we obtain the following uniform rate of convergence for the kernel estimators of $\widehat{F}_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})$ and $\widehat{f}_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})$ on any inner closed compact subset of the bidding support $\mathcal{C}(B)$.

$$\sup_{(b,z,I)\subset\mathcal{C}(B)} ||\widehat{F}_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}) - F_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})||_0 = O(\frac{\log L}{L})^{\frac{R+1}{2R+d+2}}$$

$$\sup_{(b,z,I)\subset\mathcal{C}(B)} ||\widehat{f}_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}) - f_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})||_0 = O(\frac{\log L}{L})^{\frac{R+1}{2R+d+3}}$$

In GPV, the corresponding uniform rates of convergence are obtained for the bidding distributions and densities $\widehat{F}_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})$ and $\widehat{f}_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})$ since bidders' identities are observed. However, note that the function mapping the vector of the order statistics CDF $(F_{\mathbf{B_p},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}))_{p=1,\cdots,n_I}$ into $(F_{\mathbf{B^*_i},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}))_{i=1,\cdots,n_I}$ is uniformly continuous. Thus the uniform rate of convergence that holds for $(F_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}))_{p=1,\cdots,n_I}$ remains valid for $(F_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}))_{i=1,\cdots,n_I}$.

Furthermore, from equation (12) and (14), we have on $\mathcal{C}(B)$ where the terms $\widehat{F}_{\mathbf{B^*_i},\mathbf{Z},\mathbf{I}}(b,z,I) - \widehat{F}_{\mathbf{B^*_j},\mathbf{Z},\mathbf{I}}(b,z,I), j \in I \setminus \{i\}$ are bounded away from zero:

$$\begin{aligned} ||\widehat{f}_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}) - f_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})||_0 &\leq C_1 \cdot ||\widehat{f}_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}) - f_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})||_0 \\ &+ C_2 \cdot ||\widehat{F}_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I}) - F_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})||_0 \end{aligned}$$

Thus the uniform convergence rate that holds for $f_{\mathbf{B},\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})$ remains also valid for $f_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}(b,z,\mathrm{I})$.

In any inner compact subset of the support, the pseudo values can be expressed as a continuous function of $f_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}$ and $F_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}$. Furthermore, it is the rate of convergence of $f_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}$ which sets the rate of convergence of $\widehat{X}_{ipl}$ to X_{ipl} in any inner compact subset of the support whereas the estimator for $F_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}$ is converging at a faster rate.

The second issue are the results concerning the uniform rates of convergence of $\widehat{P}_{ipl}$. From equations (22) and (23), the pseudo probabilities can be expressed as a continuous function of $f_{\mathbf{B}^*,\mathbf{Z},\mathrm{I}}$ in any inner compact subset of the support (the denominator stays bounded away from zero). Then uniform the rate of convergence proved by GPV for $\widehat{X}_{ipl}$ also applies for $\widehat{P}_{ipl}$.